APPLYING GENOMIC AND PROTEOMIC MICROARRAY TECHNOLOGY IN DRUG DISCOVERY

APPLYING GENOMIC AND PROTEOMIC MICROARRAY TECHNOLOGY IN DRUG DISCOVERY

Robert S. Matson

CRC PRESS

Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Matson, Robert S.
Applying genomic and proteomic microarray technology in drug discovery / Robert S. Matson.
p. ; cm.
Includes bibliographical references.
ISBN 0-8493-1469-0
1. High throughput screening (Drug development) 2. DNA microarrays. 3. Protein microarrays. 4. Pharmacogenomics. 5. Proteomics.
[DNLM: 1. Drug Design. 2. Genomics--methods. 3. Oligonucleotide Array Sequence Analysis--methods. 4. Proteomics--methods. QV 744 M434a 2004] I. Title.

RS419.5.M38 2004
615'.19--dc22 2004057042

International Standard Book Number 0-8493-1469-0
Library of Congress Card Number 2004057042
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

Preface

Array technology, much like polymerase chain reaction (PCR) technique, was created to satisfy an existing need in molecular biology. PCR provided a means to amplify enough DNA to sequence genes. The first applications for arrays involved gene sequencing by hybridization (SBH) and genotyping. However, gel-based sequencing quickly supplanted the emerging SBH approach, while genotyping and mutation analysis have been slow in development. The challenge for those involved in array technology then became finding that elusive application niche, one that would demonstrate a clear, unmitigated, and thereby sustained need for the technology.

This book picks up the array technology journey from the mid-1990s with the introduction of microarray-based gene expression analysis. The global analysis of genes by microarrays has provided a fresh and exciting view of the cellular process. More importantly, it enabled others to consider similar utility in various "omic" fields. Hence, we have witnessed the emergence of protein arrays to address proteomics.

In writing this book, my aim was first to provide a detailed description and offer insight into present and future utilities for microarray technology. While arguably array-based technologies are now being adopted in diverse fields, I have placed emphasis on applications related to drug discovery. Microarrays continue to play significant and increasingly important roles in the drug discovery process.

Chapter 1 considers the respective roles as well as the many issues surrounding the future adoption of gene expression and protein microarrays for pharmacogenomic and pharmacoproteomic applications. For acceptance by the pharmaceutical and diagnostic industries, commercially validated array technology is required. Chapter 2 details the commercial microarray landscape. Chapter 3 describes alternative substrates and the preparation of various surface chemistries along with their suitability for immobilization of nucleic acids and proteins. In Chapter 4, the mechanics of microarraying are described in detail including environmental conditions, printer and pin performance, and instructions for setting up a print run. Protocols for printing nucleic acids and proteins are provided along with in-depth discussion of other important parameters such as print buffers (inks) and factors influencing print quality. I also set out to discuss the importance and provide a critical assessment of studies that helped to define applications in genomics

and proteomics. In Chapter 5, gene expression microarray applications are described; Chapter 6 examines the utility of protein microarrays.

Finally, an understanding of the making of a microarray is fundamentally important to those interested in producing "spotted" arrays and properly using them. While complementary (cDNA) microarray fabrication on glass slides has been well studied, we have less experience with the attachment of oligonucleotides and the preparation of protein arrays. Moreover, additional substrates and surface chemistries that may be better suited for printing proteins are now available.

It is my hope that this book will provide you with the knowledge and confidence to embrace microarraying in your future.

Robert S. Matson, Ph.D.
Orange, California

Author

Robert (Bob) Matson, Ph.D., is a senior staff scientist in the Advanced Technology Center at Beckman Coulter, Inc., Fullerton, California. He has been involved in the development of both nucleic acid and protein array-based technology for the past 13 years. His initial introduction to array technology began in collaboration with Sir Edwin Southern in developing an *in situ* oligonucleotide array synthesis platform for the corporation. Later work by Dr. Matson and his research team produced some of the first plastic microplate-based microarrays. Beckman Coulter recently launched the A^{2TM} plate based upon the microplate "array of arrays" concept.

Prior to joining Beckman Coulter, he served in several technical management roles including: R&D director at BioProbe International, R&D director at Costar-Nuclepore, and R&D group leader, chemistry, at BioRad Laboratories.

Dr. Matson currently holds seven United States patents and has contributed numerous papers in peer-reviewed journals as well as chapters in several books on microarrays. He has also made many presentations in the United States and abroad on the development of microarray technology. His current interest is in automated approaches to multiplexed assay development.

Dr. Matson grew up in the San Juan Islands of Washington State and attended Western Washington University, Bellingham, where he earned his B.A. and M.S. in chemistry. He received his Ph.D. in biochemistry from Wayne State University. Following postdoctoral studies at the medical school of the University of California at Los Angeles, he served as a principal investigator with the Veterans Administration Medical Center and as an adjunct professor of biological chemistry at the medical school of the University of California at Davis. Dr. Matson also held a faculty lectureship in the department of chemistry at University of Southern California and was an assistant professor of chemistry at the University of Southern Maine, Portland. He served on the editorial boards of *Applied Biochemistry and Biotechnology* and the *Journal of Preparative Chromatography,* and is a member of the Scientific Advisory and Organizing Board of International Business Communications' "Chips to Hits" conferences.

Dr. Matson is a member of the Science Education Advisory Board of the University of California at Irvine. He also volunteers in numerous science education outreach programs including Project Tomorrow and the Orange County Science Fair. He is a past board member of the Orange County Science Education Association (OCSEA).

Contents

chapter one

Quantitative biology: The "Omics" Era

> Now biology has come, in a significant part, a high-throughput, industrialized operation.
>
> **John N. Weinstein***

Introduction

The manner in which we approach the study of cell biology has changed dramatically over the past decade. We can speak of a paradigm shift — one in which the biologist has been greatly enabled by a rather simple tool, the *microarray*. The microarray in its most elementary form is a collection of small spots of biological capture agents (DNA, antibody, carbohydrate, etc.) organized on a planar substrate such as a glass microscope slide. This tool allows us to move beyond our rather myopic views of the cell in favor of a more global assessment of the cellular process. Moreover, the microarray provides us with digital information that we can assemble and use to quantify biological events and relationships on a scale that was unimaginable a few decades ago. Finally, the microarray is a parallel processor that provides the researcher with a rapid response to a biological query. In fact, so much data can be obtained in such a short time that it can be overwhelming and often require the aid of sophisticated bioinformatics analysis software. Thus, the microarray has become a formidable instrument by which to quantify biology. The purpose of this book is to assess the progress on the utility of microarray technology to solve important biological problems.

* 'Omic' and hypothesis-driven research in the molecular pharmacology of cancer, *Curr. Opinion Pharmacol.*, 2, 361–365, 2002.

Microarray format

Terms and definitions

Array technology finds its origins within molecular biology, a scientific field of study notoriously irreverent about the rules governing systematic nomenclature (Figure 1.1). Moreover, microarray applications have crossed over into other scientific fields, each with its own unique terminology. Some general terms that apply to the microarray field are described below. Other terms and further elaboration with illustrations can be found, for example, in the concise A to Z guide titled *Microchips: The Illustrated Hitchhiker's Guide to Analytical Microchips* (Kricka, 2002).

DNA arrays have been categorized into different formats based upon what is immobilized to the surface (also known as the *solid phase*, *substrate*, or *chip*) and what is captured from the sample solution. Definitions change depending upon the format. For the classic *Southern dot blot*, the sample was first spotted down on the surface, cross-linked, and then bathed with a radio-labeled oligonucleotide under *hybridization* (complementary nucleic acid strand base-pairing) conditions to detect the presence of a particular sequence within the sample. This was called probing. The oligonucleotide

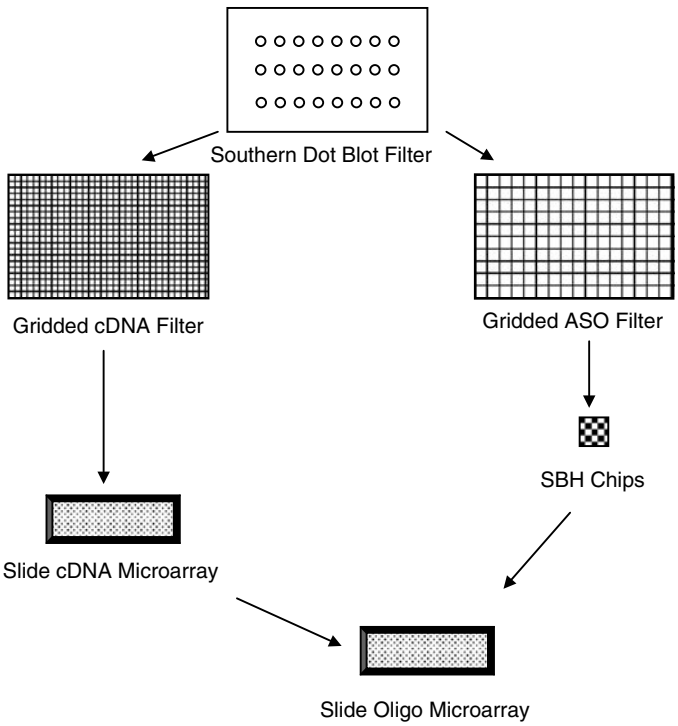


Figure 1.1 Origin of the microarray.

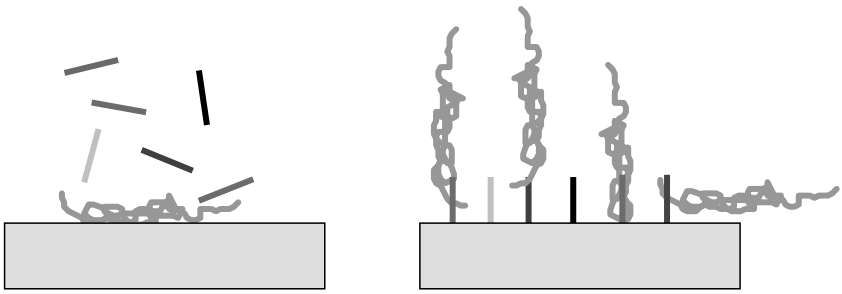


Figure 1.2 Hybridization blot formats.

was the *probe* and the DNA in the sample that was immobilized onto the surface became known as the *target*. Subsequently, the Southern dot blot hybridization process was classified as Format I (or *forward blot*) and was associated primarily with membrane blots. The invention of the Format II (or *reverse blot*) led to some confusion in the early days of microarrays. Here many different probes are immobilized onto the substrate and the sample (target) is labeled for hybridization (Figure 1.2). Additional formats will be described later.

In most instances, however, we will be discussing Format II microarrays for both nucleic acid and protein capture. The generally accepted meaning is that a probe serves as the solid-phase capture agent (Figure 1.3). In drug discovery, the target is usually the protein (or protein complex) that chemically interacts with a compound (drug candidate).

Other terms also relate to the anatomy of the microarray (Figure 1.4). The probes are immobilized on the substrate at discrete (x, y) locations or spatially addressable sites. The probe spots [measured in microns (diameters) for a circular spotted array or as a side of a square for an *in situ* array] are often referred to as features or elements of an array. Thus, an array containing 10,000 features would have 10,000 probes arranged as an array on a substrate.

Typical spotted arrays would have 100 to 150-micron (μm) diameter features while photolithographically prepared *in situ* arrays may have features on the order of 2 to 20 μm. The separation between elements is usually measured in terms of a *center-to-center distance*, *spacing*, or *pitch*. Thus, for a printed array, two adjacent spots in the array, e.g., each at 100-μm spot diameter, might have a center-to-center distance of 150 μm, or the spots would be separated by 50 μm from their edges. The number of spots per square centimeter usually defines the *spot density*. As an example, an array manufactured by Affymetrix (Santa Clara, CA) at >280,000 elements per

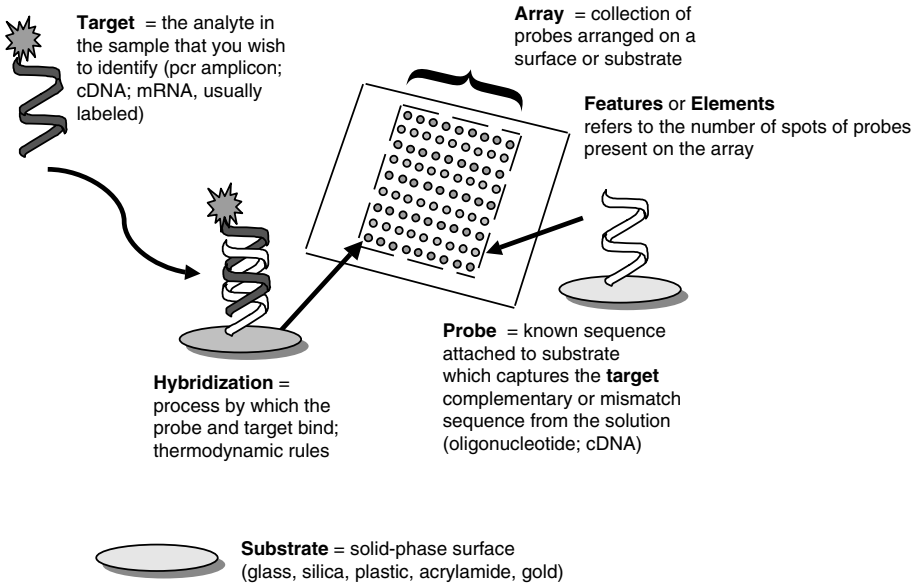


Figure 1.3 Microarray anatomy.

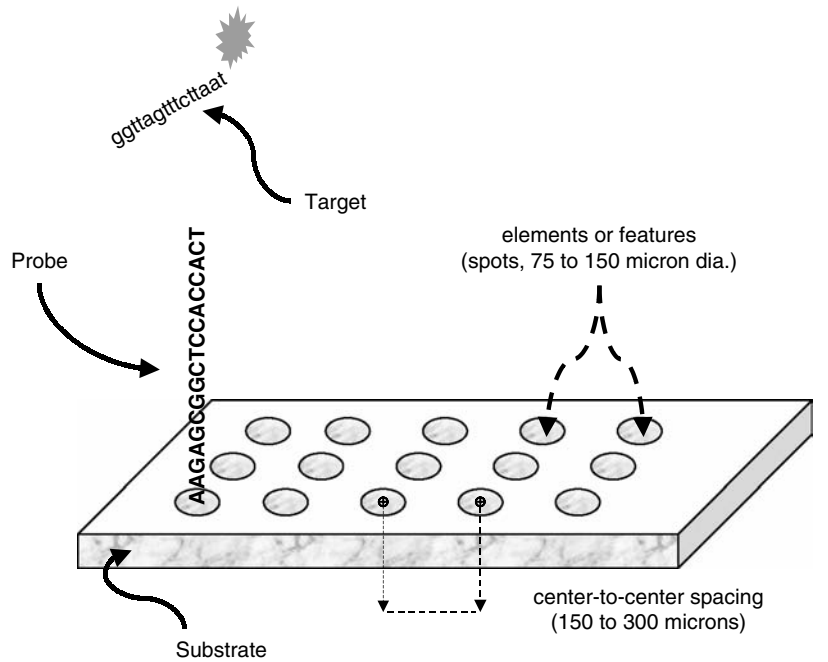


Figure 1.4 Microarray terms.

1.28-cm × 1.28-cm chip would have a spot density on a chip of 170,000 elements/cm^2.

We can also classify arrays in terms of high, medium, and low probe density. There may be some argument regarding the precise boundary limits distinguishing microarrays on a density scale. Nevertheless, a high density array would contain >10,000 probes/cm^2; medium density, 1000 to 10,000 probes/cm^2; and low density, <1000 probes/cm^2. Arrays are also defined as macroarrays and microarrays. A macroarray can be regarded as having larger and fewer spots than a microarray. For example, the Southern blot on a standard sized membrane (~8 cm × 12 cm; nylon or nitrocellulose) with spot diameters of 500 μm would be considered a macroarray by most researchers. However, macroarrays may have thousands of printed spots per membrane and thus functionally perform at a level similar to that of a slide microarray.

The number of probe molecules per square millimeter defines the probe density of a spot. Probes within an element or spot may have densities on the order of 10^9 to ~10^{12} molecules/mm^2 depending upon the molecular size of the nucleic acid (e.g., short oligonucleotide vs. cDNA).

General utility

As we will soon discover, microarray-based technologies have found utility in a number of fields. While DNA arrays are the most technically mature and have the broadest application portfolio, we have witnessed the ever-increasing generation of new kinds of probe arrays: antibody, antigen, enzyme, aptamer, carbohydrate, tissue, cell, and small molecule microarrays. The list undoubtedly will continue to expand. We can also describe microarrays in terms of prognostic, diagnostic, and predictive roles. A few examples that examine these applications are provided.

Biomedical testing

Applications for array technology have broadened to include such fields as forensics and plant and animal genotyping. However, the primary focus for microarrays is on biomedical-related analysis. Biomedical testing can be divided into four major areas: discovery testing, clinical trial testing, specialty testing, and patient care testing (Figure 1.5).

Most emerging technologies are likely to have been invented or developed during the discovery process. In the area of biomedical testing, new technologies will pass from research (academic, biomedical, pharmaceutical) into specialized testing arenas such as core facilities or outsourced (esoteric testing) laboratories, and hopefully be adopted from such uses into the mainstream of diagnostics. Indeed, array technology finds its roots in academic circles as a means to sequence genomes. The reverse blot was originally constructed at a diagnostic research center as a new nucleic acid assay

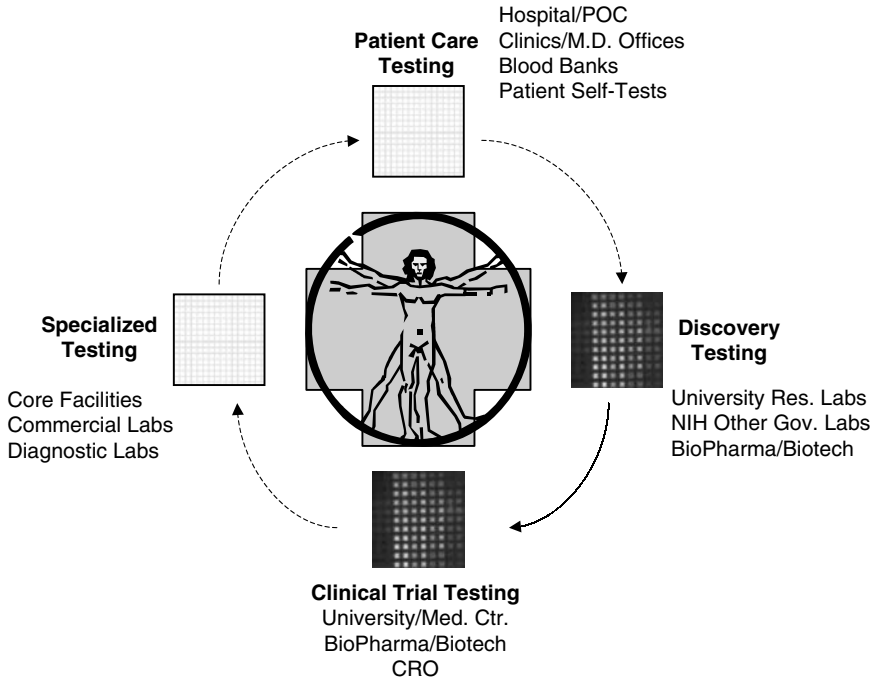


Figure 1.5 Biochip migration.

format for the simultaneous detection of infectious agents. To date success in moving array technology directly into point-of-care or centralized laboratory testing has been limited. Rather, microarray technology is now slowly migrating from discovery into specialized testing arenas such as reference or esoteric testing laboratories (Figure 1.5).

Biotechnology sector trends

We can characterize the application of array technology as riding upon the "omic" wave (Figure 1.6). We are currently experiencing the technological and commercial maturation of microarrays on gene expression and single nucleotide polymorphism (SNP) discovery platforms. These have benefited from the successful completion of the human genome sequencing efforts (Genomics I). Whole genome chips are now or will soon be available for use by a number of vendors. At the same time, we are also at the beginning of an advancing proteomic wave and, behind that, we can speculate that perhaps a second genomics wave based upon the excitement surrounding the small interfering RNA (siRNA) discoveries will follow. The availability of these high-content chips has also increased our thirst for bioinformatics. Information technology (IT) is projected to grow as these waves of progress appear (Figure 1.7).

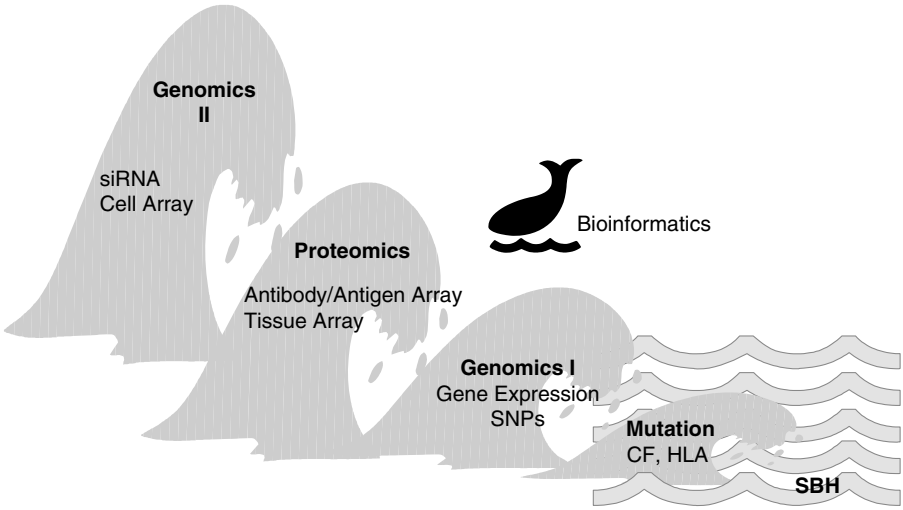


Figure 1.6 Microarray technology waves.

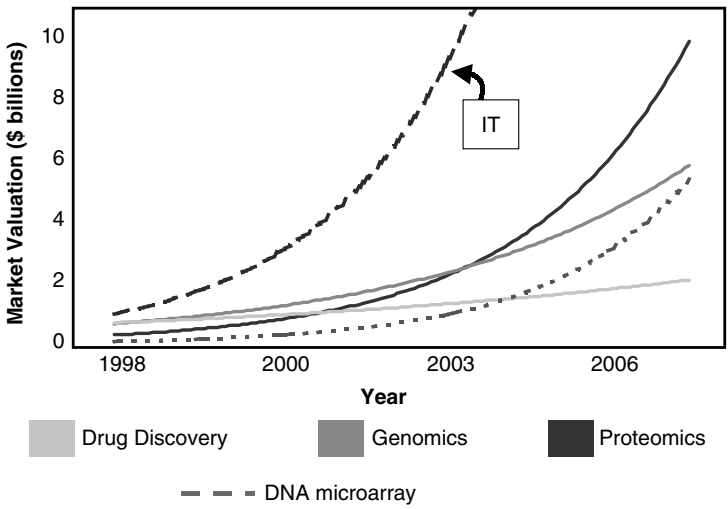


Figure 1.7 Biotechnology sector market growth trends.

The Omics Era

A major driving force in the advancement of microarrays has been the Human Genome Project and the development of the genomics field. Oddly enough, the origins of DNA array technology began with efforts at sequencing by hybridization (SBH) of the human genome. With the exception of a few private ventures such as the work at HySeq Pharmaceuticals, the SBH approach was quickly supplanted by dideoxynucleotide dye-terminated

gel-based sequencing. The other early use of DNA arrays was in mutation screening such as for cystic fibrosis (CF) mutations or in human lymphocyte antigen (HLA) typing.

The problem with screening for mutations associated with genetic disease remains a lack of therapy or cure. Without the availability of treatment there is very little incentive to produce a diagnostic test. The other issue with mutation detection is that the prognosis is often a statistical inference. For example, possessing the BRCA1 gene mutation only increases the probability risk or likelihood that a woman may develop familial breast cancer. The presence of the mutation alone cannot unequivocally determine whether or not someone will develop breast cancer. Other factors must be weighed — in this case, the family history regarding the occurrence of breast cancer. Thus, mutation detection based upon microarray technology has not grown as rapidly as anticipated.

The breakout opportunity for widespread adoption of array technology came largely from gene expression studies in which the expression levels from two cell states (e.g., control vs. drug-induced; normal vs. diseased) were compared. While sequencing efforts defined structural genomics, the gene expression microarray became the tool for functional genomics. In particular, gene expression microarrays were found to be well suited as a new type of differential display technology applied to the drug discovery process in which pharmacogenomics examines the responsiveness of genes (cellular messenger RNA [mRNA] levels) to drug candidates. The microarray measuring mRNA levels within a cell's genome permits identification of potential targets (mRNA serving as the surrogate of translated protein) or alterations to metabolic pathways, thereby implicating the participation of other targets. Metabolic alterations may also lead to so-called off-target effects that are adverse or toxic. Therefore, microarrays have been found to be very useful in toxicogenomic applications. In essence, mRNA profiling using microarrays can potentially reduce the need for extensive animal model or tissue studies to survey off-target toxicity at the onset of the target discovery phase and in the later preclinical phase (Figure 1.8).

Gene expression analysis has recognized limitations (Lillie, 1997; Clarke et al., 2001). The monitoring of transcriptional events serves as an indirect measure of protein (target) expression. Because proteins can undergo post-translational modification leading to subcellular localization and pooling, export, degradation, or complex formation with other proteins, the correlation between mRNA level and putative protein is often lost (Figure 1.9).

It is important to consider what both genome- and proteome-based arrays offer (Figure 1.10). The gene expression microarray monitors relative mRNA abundance between two cell populations, provided control and sample populations are processed in the same manner. This is often spoken of as a "snapshot" of gene activity. That is, unless we also sample these populations often, only a finite view of the changes in cellular activity between the two sets will be recorded. We may need at times many snapshots or time course studies. The great advantage of the gene expression array is that it

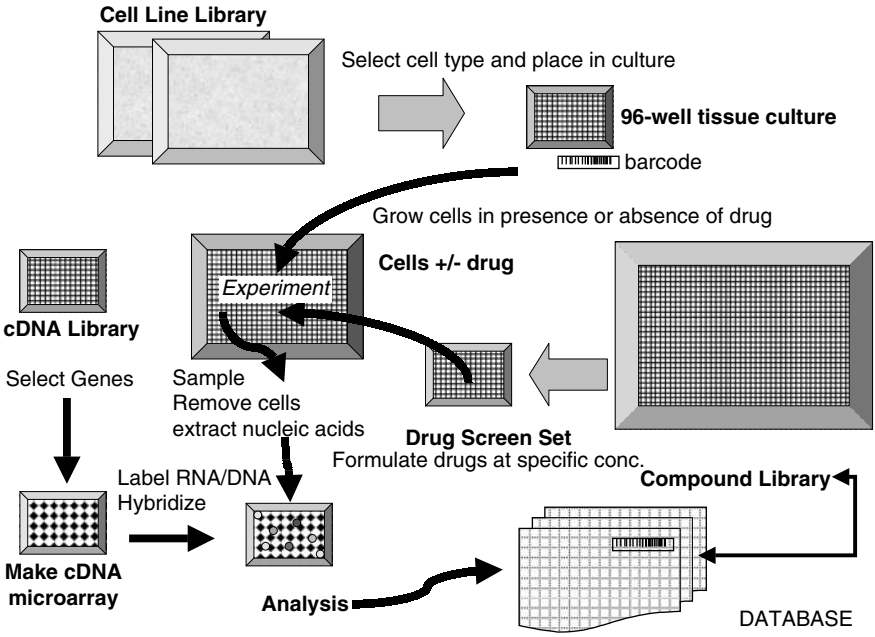


Figure 1.8 Gene expression of off-target analysis using microarrays.

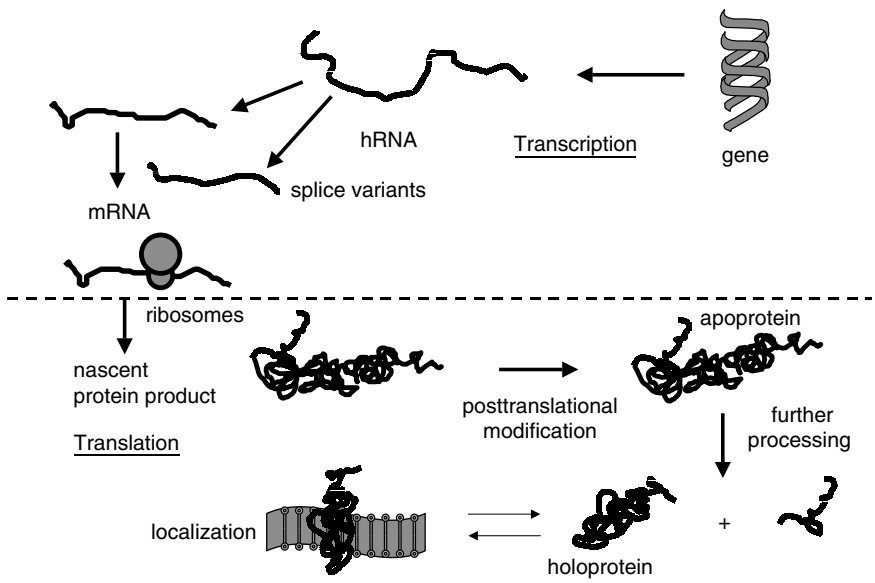


Figure 1.9 Proteins undergoing posttranslational processing.

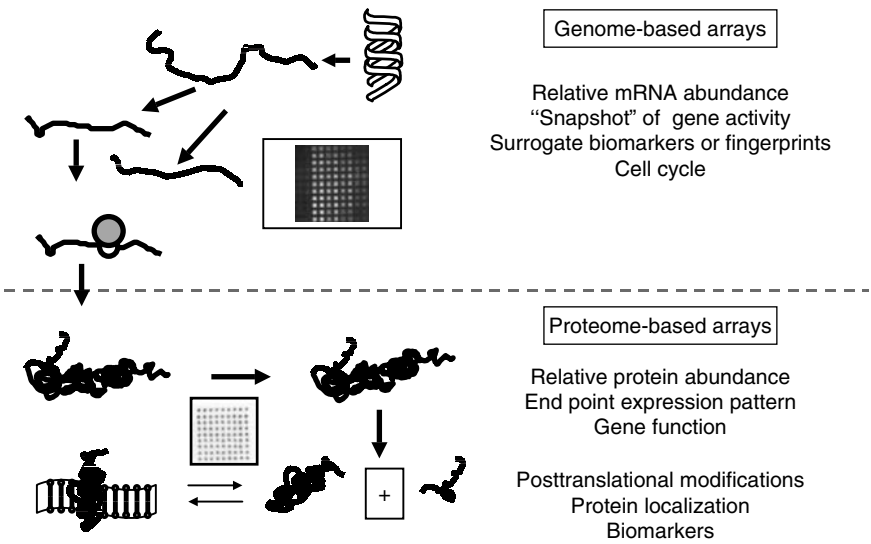


Figure 1.10 Top: Genome-based arrays. Bottom: Proteome-based arrays.

permits a global analysis of the genome. Later, we will relate how genome-based approaches provide good information about cell cycles and pathways, leading to the discovery of potential surrogate biomarkers. Gene expression microarrays remain important tools for several reasons:

1. Availability of full genome representation
2. Ability to amplify for detection of less abundant mRNAs
3. Simple capture process based upon hybridization thermodynamics and primary sequence information

Obviously, for other reasons such as RNA splice variants and the post-translational modification of proteins, one cannot rely on mRNA profiling alone. Protein microarray development is under way and is anticipated to allow global inspection of relative protein abundance. However, this is a future prospectus because currently there are few means of obtaining a highly significant representation of a proteome (i.e., content) on any chip. Antibody arrays have been introduced but are of limited utility at this point for proteome-wide applications. Because microarrays are closed architecture technology platforms, they can only provide information based upon what is contained on the chip; therefore, we must first acquire enough protein content (antibody libraries or mimetic agents such as aptamers) necessary for proteome discovery work.

Some argue that it is a waste of time to monitor gene expression when ultimately we desire the end product, the protein target. More likely both aspects are needed to further expand our understanding of cellular events (Clarke et al., 2001). All of these pro-vs.-con arguments may sound quite

reasonable, yet are we not at the beginning of the "omic" wave? Now, with our list of genes in hand and a few good tools, we may begin to discover the true complexity of our biology. As John Weinstein (2002) suggests,

> Perhaps the most important (and least recognized) aspect of biology that makes it difficult to understand systematically is the fact that biological complexity was not produced by a watch maker or an engineer.

While we can consider the strengths and weaknesses of various technologies and pursue "omic" approaches, we must also not lose sight of the fact that they are tools to aid our studies. J.P. Miller (2002) makes this point:

> These novel and significant tools, which are rapidly becoming indispensable, do not by themselves enlarge the bedrock of basic knowledge that underlies new discoveries. That foundation remains the detailed understanding of the biological basis of health and disease.

Role of gene expression microarrays in drug discovery

Fundamental approaches in the drug discovery process have in themselves undergone a paradigm shift (Neamati and Barchi, 2002). The pharmaceutical industry has embraced molecular biology, adopted robotics for high throughput screening, and supported those efforts aimed at genome sequencing and SNP identification as related to disease. Microarray technology, especially gene expression chips, has been well received in the drug discovery labs of Biopharma. According to Neamati and Barchi (2002), modern drug discovery can be arranged into five major areas (Figure 1.11):

1. Target identification (discovery)
2. Target validation
3. Lead identification (hits)
4. Lead optimization
5. Preclinical pharmacology and toxicology

In the target discovery process, we have witnessed an increased role for genomics. In fact, as much as 25% of new target identification efforts may now be based upon genomic approaches. Differential gene expression pattern analysis (control vs. drug response phenotype; normal vs. diseased state) developed using microarray chip technology continues to play an important role. High-density microarrays such as Affymetrix's GeneChip® were found to be very useful in target identification. Microarrays based upon normalized cDNA libraries have also been successfully used in the discovery of novel genes that are potential candidates for drug targeting (Katsuma and Tsujimoto, 2001).

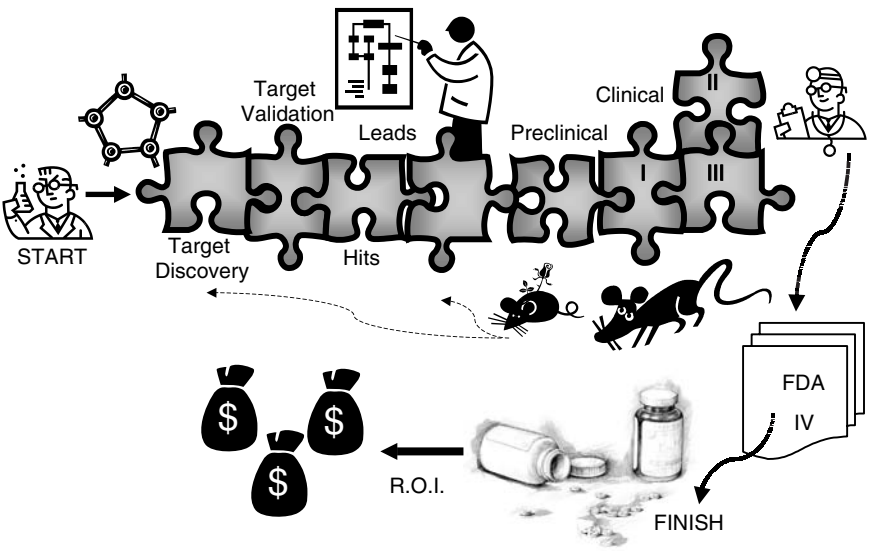


Figure 1.11 Drug discovery process.

Now that the human genome is essentially complete and the sequencing of useful model genomes (mice, rats, yeasts, microbes) is almost finished, the gene expression array should continue to serve in an expanded capacity. While differential gene expression is but a surrogate methodology, the widespread use of the equivalent proteomic tool (i.e., the protein expression microarray) is merely a notion at present. We will see later how gene expression profiling of metabolic pathway enzymes can play a pivotal role in guiding efforts toward identifying new targets.

According to Rininger et al. (2000), pharmacogenomics is rapidly becoming an accepted route in the later stages of the drug discovery process involving both the preclinical and clinical phases. A key factor in the acceptance of the pharmacogenomic (and pharmacogenetic) approach is that both drug efficacy and toxicity are well correlated to changes in gene expression. Microarrays offer both high throughput and sensitivity. These attributes are particularly advantageous in reducing the time and cost in determining drug toxicity during the preclinical stage. During late stage clinical trials, it is anticipated that the profiling of an individual's genetic variation (SNP) correlated to drug response will be an important screening process. Several microarray-based SNP "calling" platform technologies (e.g., Beckman Coulter's SNPstream and Illumina's BeadArray) are involved in defining specific polymorphisms associated with variable drug responses within individuals and various populations.

Without a doubt, one of the most significant testimonies to the power of the microarray has been in the characterization of cancer cell lines and tumor gene expression. The National Cancer Institute's (NCI) study of cDNA

microarray gene expression patterns across 60 human cancer cell lines (with an activity pattern database on each cell line individually challenged by over 70,000 compounds) remains the *tour de force* in microarray-based profiling (Ross et al., 2000). Even more impressive is the assembly of the integrated gene expression–molecular pharmacology database for the NCI60 cell lines (Scherf and Ross, 2000). Although the NCI60 study had some recognized limitations (limited activity assay data; limited numbers of arrayed genes; surrogate relationship of cell line to tumor cells), it also produced encouraging examples indicating that the gene expression approach has unraveled mechanisms of drug resistance.

Rew (2001) considers five areas in cancer research applicable for the use of microarray technologies:

1. **Tumor classification** — Drug treatment regimes often depend upon tumor type. The origins of metastatic tumor cells can be difficult to determine by conventional histopathology. Gene expression profiling may complement the more traditional methodologies where tumors are difficult to classify.
2. **Mutation detection** — Genetic mutations leading to disease states such as familial breast cancer can be detected by arrays, provided that sufficient gene probe sequence content is available in order to make a statistically significant prognosis.
3. **Gene copy number** — Tumors are known to contain variable numbers of genes relative to normal tissue. Comparison of the normal and tumor gene populations is useful in identifying tumors and their differentiated states.
4. **Cancer therapeutics** — Different tumor types and the differentiated states of individual tumors are known to exhibit unique gene expression patterns called *molecular signatures* or fingerprints. The molecular signature may be useful in tumor phenotype classification.
5. **Drug sensitivity** — Microarrays are useful in monitoring both on-target and off-target drug responses. Knowledge of which genes are up- or down-regulated may lead to an understanding of mechanisms of action and new treatments.

In a review by Zanders (2000), a good case is made for the use of gene expression microarrays to monitor changes in signaling pathway activity. The rationale is that under environmental stress such as in tumor growth or during an inflammatory response, signaling pathways are activated or repressed; and these events can be measured by gene activity by mRNA profiling. That is, "transcription of mRNAs could be exploited as a 'surrogate marker' of signaling pathway activation." Cited examples include studies concerning the expression profiling of signaling pathway enzymes in yeasts such as the mitogen-activated protein kinases (MAPKs) during times of growth, differentiation, or under stress (Roberts et al., 2000). Homologous mammalian MAPKs have been reported and are being investigated as

potential drug targets. In a study by Iyer and others (1999), the use of gene expression microarrays revealed the induction of genes involved in wound healing during serum-stimulated human fibroblast growth.

Toxicogenomic applications

As noted by Nuwaysir et al. (1999), "Almost without exception, gene expression is altered during toxicity, as either a direct or indirect result of toxicant exposure." This early "cause-and-effect" revelation most certainly promoted the use of gene expression microarray technology at the National Institutes of Health (NIH) resulting in the creation of the ToxChip v1.0 cDNA microarray (http://ehp.niehs.nih.gov/docs/1999/107-5/innovations.html).

A so-called toxicant signature is derived from gene expression relative fold changes between control and treated cell populations. The ToxChip contains approximately 2000 human gene cDNAs arranged in various functional categories such as apoptosis, cell-cycle, cytochrome P-450s, etc. to detect responses to toxic insults. Hodges et al. (2003) were able to elucidate the mechanism of action for tamoxifen, a drug used in the treatment of breast cancer, using the ToxChip. Lobenhofer et al. (2002) used the ToxChip v1.0 to study the mechanism of estrogen-induced proliferation (mitogenesis) using a hormone-responsive human breast cancer cell line.

While toxicogenomic profiling is very useful in providing an adjunct to animal studies, certain precautions are required to interpret microarray data because "the transcriptome profile is extremely sensitive to any subtle changes surrounding cells, tissues, or individual organisms" (Shoida, 2004). Several such factors may influence the outcome of a microarray experiment and bias the toxicant signature. For example, circadian rhythms are often overlooked. A simple change of culture medium can exert a dramatic effect on cells in a culture.

In conclusion, gene expression array technologies have been largely accepted in the scientific community. They are well suited as tools for drug discovery and the elucidation of drug-target mechanisms of action (Clarke et al., 2001; Cunningham, 2000). In particular, microarrays provide insight into:

1. Discovery and validation of new targets
2. Determination of drug efficacy, resistance, and toxicity
3. Identification of new diagnostic biomarkers

Proteomics today: The great challenge

Riding on the success of the Human Genome Project in cataloging the 30,000 to 40,000 genes of the human genome, the Human Proteome Organization (HUPO; http://www.hupo.org) has set out to determine the entire set of expressed proteins comprising the human proteome(s). As Tyers and Mann

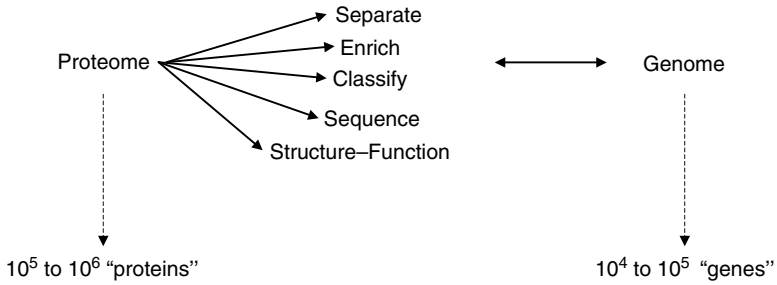


Figure 1.12 Proteomics today: the great challenge.

(2003) relate: "Proteomics would not be possible without the previous achievements of genomics, which provided the 'blueprint' of possible gene products that are the focal point of proteomics studies."

For the development of new therapeutics, the *proteome* (the complete and well-characterized set of human proteins) is an important issue because proteins make up the majority of drug targets, yet the proteome is perhaps greater than an order of magnitude and more complex than the genome (Figure 1.12). While the genome comprises 10^4 to 10^5 genes (including splice variants), the proteome may contain over a million proteins if posttranslational modifications and isotypes (isoforms) are included. Furthermore, proteins vary widely in abundance and occur in multiprotein complexes localized within cellular and suborganelle membranes.

The real challenge could be in the high-throughput, highly parallel, micropreparation of this structurally diverse class of biomolecules in their native states. Low abundance proteins will most likely require enrichment prior to detection (Figure 1.13). Protein complexes will need to be isolated, some along with associated membrane components, in order to preserve activity. It will also be important to be able to reassemble these multiprotein complexes in their native and active states.

Some scientists believe that because of its excellent sensitivity, mass spectrometry (mass spec) may be able to detect proteins in biological fluids such as serum without the need for separation. Petricoin and Liotta (2004) review mass spec differential display pattern profiling of serum proteins and peptides associated with various cancers. In Chapter 6, we will discuss Ciphergen's successful approach to biochip-based mass spec protein profiling. Protein profiling offers an exciting opportunity as a noninvasive (or nearly so, requiring only a prick of blood) diagnostic tool similar in scope to the MRI scan for viewing body tissues.

Potential role for protein microarrays in drug discovery

As we noted earlier, there are some very good reasons to consider proteomic approaches in drug discovery, for example:

1. Most drug targets are proteins (largely receptors and enzymes).
2. Drug-mediated therapy is based upon less than 500 targets (a very small fraction of the proteome's estimated 100,000+ to 10^6 proteins), indicating that finding many more effective targets is likely and highly desirable.
3. mRNA profiling is an indirect assessment of protein expression and cannot detect posttranslational modification, an important signaling and regulatory process for drug targeting.

Because a modest number of well characterized antibodies are readily available, it is no surprise that the first demonstrations of the use of protein microarrays came from work on antibody–antigen arrays. Yet, the fundamental technology is not new. Roger Ekins' reviews (Ekins et al., 1990; Ekins, 1998) introduced "microspot" technology for clinical diagnostics in the 1980s. Later, MacBeath and Schreiber (2000) and Haab et al. (2001), borrowing the tools and know-how from the cDNA microarray world, promoted the use of slide-based antibody microarrays.

Huels and coworkers (2002) identified several areas within the drug development process where protein biochips have potential application. The first

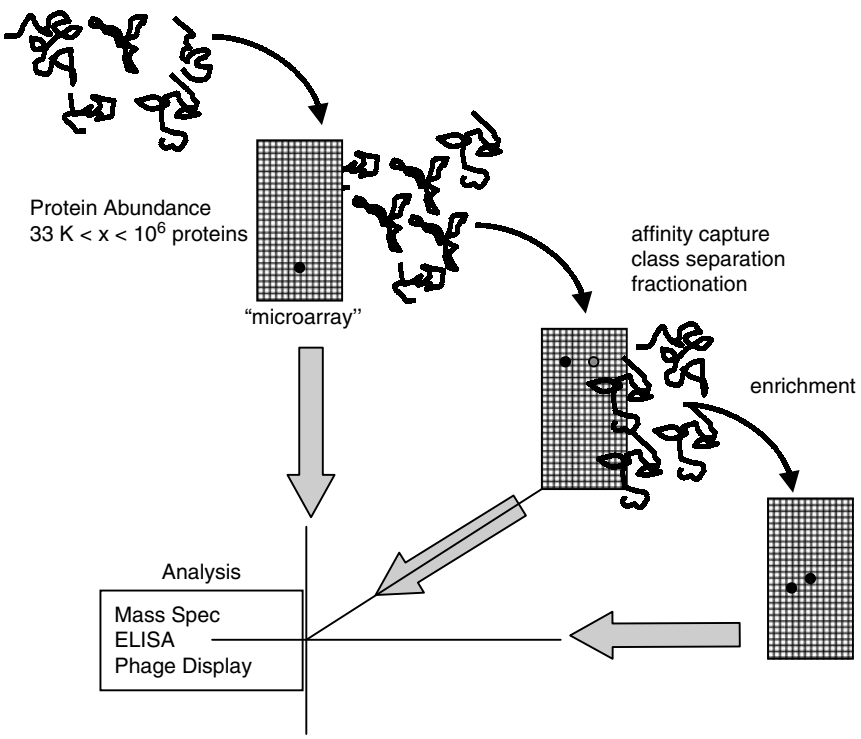


Figure 1.13 Complexity reduction for improved proteomic analysis using protein microarrays.

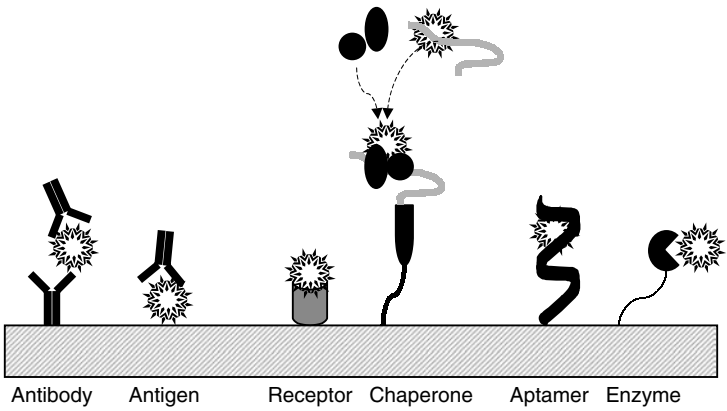


Figure 1.14 Protein microarray formats.

generation protein microarrays, the antibody-based micro-enzyme-linked immunosorbent assay (ELISA) formats, are just now entering the marketplace. Antigen arrays will serve as bait in the capture and characterization of additional antibodies needed to fill in the proteome libraries. It is anticipated that several well characterized antibodies are required in order to cover more than one epitope on each antigen. This is especially important if the sandwich assay is to be optimally employed. Antibody and antigen microarray formats are briefly described below with more detail in Chapter 6 titled "Protein Microarray Applications."

Future generation protein arrays will include, in addition to antibody–antigen binding, other protein–protein or receptor–ligand interactions (Figure 1.14), protein–peptide, substrate–enzyme, aptamer, and protein–small molecule high-throughput assays. Proofs of principles of a variety of these formats have been the subjects of numerous reviews. Essentially, protein microarrays are believed to be able to broadly cover the various phases of the drug discovery process, similar to what we described for gene expression microarrays (Figure 1.15). However, as with the DNA microarray, a few hurdles remain(Figure 1.16).

Critical issues with protein microarrays

Stability and performance

While we can appreciate the first demonstrations of the utility for the protein microarray, in retrospect, they did not perform particularly well relative to standard ELISA results. As reported by Haab et al. (2001), only 50% of the antigens and only 20% of the antibodies performed well (i.e., quantitative detection of the cognate antigen–antibody) in the μg/mL range with some allowing detection into the sub-ng/mL range. Is this really a high degree of sensitivity when a good ELISA can exhibit orders of magnitude with higher sensitivity (pg/mL to fg/mL)?

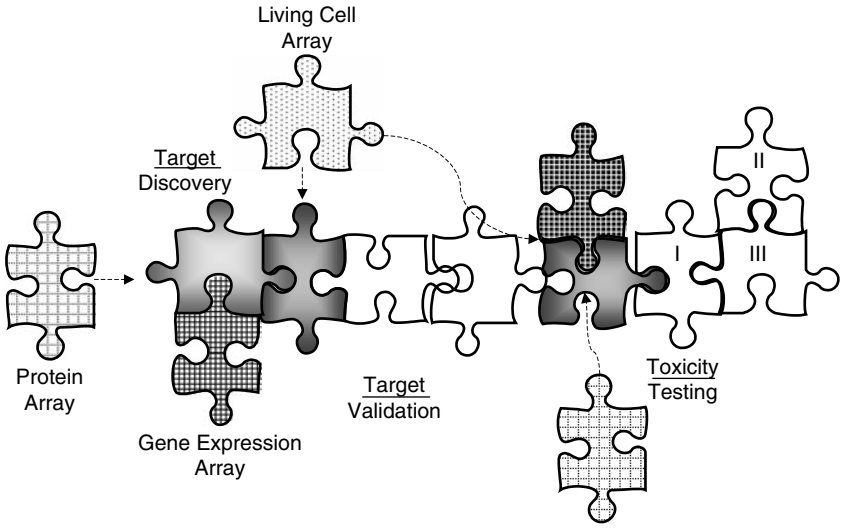


Figure 1.15 Role of microarrays in drug discovery process.

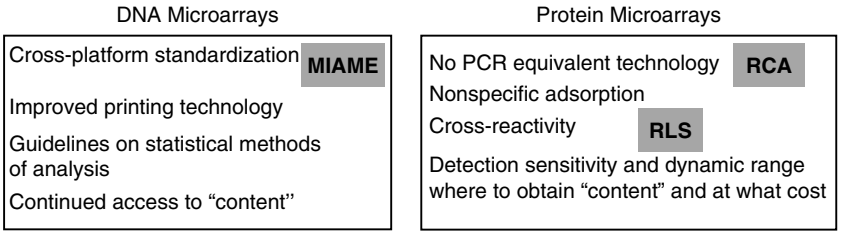


Figure 1.16 Critical issues facing DNA microarrays and protein microarrays. MIAME = minimum information about a microarray experiment, checklist endorsement as of October 14, 2002. RCA = rolling circle amplification. RLS = resonance light scattering.

As others have pointed out, assembling proteins on a substrate is a relatively easy task but much more of a challenge if all are to remain fully functional (Cahill, 2001; Sreekumar and Chinnaiyan, 2002; Valle and Jendoubi, 2003). Unlike the DNA array where the capture of cDNA or cRNA targets is relatively straightforward due to the structural similarities of probes and targets, proteins (even within the same class, e.g., antibodies) can vary considerably in their physical–chemical stability and binding character on surfaces. Other challenges that lie ahead include selection of optimal substrates, immobilization and presentation (orientation) of the ligand, and the development of sensitive assays for different protein classes.

Content

Antibody coverage of the human proteome is estimated to be about 5 to 10% of all human proteins and isoforms (Valle and Jendoubi, 2003). A major bottleneck in the use of protein expression arrays is the lack of such a comprehensive set of these capture agents (Hanash, 2003). Since an equivalent of the polymerase chain reaction (PCR) process for mass amplification of low abundant proteins does not exist, the remaining library of proteome capture ligands will need to be generated by other means such as recombinant protein expression systems (Cahill, 2001).

However, recombinant antibodies may be less stable and have lower binding affinities than monoclonal antibodies (Valle and Jendoubi, 2003). Therefore, in order to fully implement the microarray format, a host of diverse capture agents could be required in addition to antibodies. These include peptides, small molecules, aptamers, ribozymes, or other molecular recognition probes yet to be discovered. However, it is also understandable because of the diverse nature of proteins that additional technologies besides microarrays will be used in proteomics research (Hanash, 2003).

Detection

Thomas Kodadek (2001) stated that: "As difficult as it is going to be to isolate and produce thousands of high affinity and specificity protein ligands, it may be even harder to come up with a good way to monitor binding of proteins to the chip." In his review, he made the following points regarding near-term approaches for detection on protein microarrays:

1. **Dye labeling of proteins in cell extracts** — A well known problem is that different proteins have different labeling efficiencies; generating individual calibration curves for each protein within an array would be impractical, so quantification of protein arrays is a challenge. Another approach would be to generate a labeled standard reference protein mixture or calibration set at different concentrations (Haab et al., 2001). The reference could then be mixed with the cell extract, providing some degree of normalization semiquantitative information from the dye ratios.
2. **Chemical modification of proteins** — This can lead to denaturation and aggregation and reduce both specificity (increased nonspecific binding) and sensitivity (decreased ligand affinity).
3. **Sandwich assay** — This format works well for ELISA. The success and potential shortcomings for microarrays are discussed in the next section.
4. **Mass spectrometry** — It has excellent sensitivity but low throughput and is semiquantitative. The ICAT (isotope-coded affinity tag) labeling method for proteins enables differential display analyses using mass spectrometry (Griffin and Abersold, 2001). This may be

adaptable to work with protein microarrays using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry similar to the original method developed by Sequenom (San Diego, CA) for SNP detection.

Micro-ELISA formats

Assays involving antibody or antigen arrays tend to be micro-ELISA formats. For instance, anticytokine monoclonal–monoclonal or monoclonal–polyclonal antibody pairs are readily available and well characterized for use in ELISA assays. Cytokines are important biological indicators used in drug discovery and toxicity testing. Therefore, it is no surprise that many of the early antibody microarray demonstrations involved the determination of cytokines from cell cultures. They simply miniaturized and multiplexed the familiar sandwich immunoassay. The microarray-based cytokine assays generally claimed sensitivities in the low pg/mL range with linear dynamic ranges from about 10 to 10,000 pg/mL, depending upon the analyte. These are similar in performance to the standard ELISA.

Woodbury et al. (2002) developed a micro-ELISA assay for determination of hepatocyte growth factor (HGF) in human serum. They employed a horseradish peroxidase (HRP)-catalyzed tyramide signal amplification (TSA)–biotin amplification with streptavidin-Cy3 reporter and claimed sensitivity to 0.5 pg/mL or 6 fM HGF with a linear dynamic range of 12 to 4,000 pg/mL in serum. This was good enough because *clinically relevant levels* varied from 199 to 1640 pg/mL (breast cancer patients) and from 153 to 998 pg/mL for age-matched normal controls. In addition, HGF and four other antigens found in serum were simultaneously quantified even though the analytes varied in physiological concentrations from 20 to 60,000 pg/mL. The HGF microarray ELISA correlated ($r^2 = 0.90$) with a standard 96-well plate ELISA.

For the determination of a limited set of analytes, the antibody microarray works reasonably well and is comparable in performance to standard ELISAs for individual analytes. Furthermore, a suitable number of detection technologies are available for use on a variety of supports. The real problem may be determining at what point antibody microarrays "hit the wall" for multiplexing. What is not generally appreciated is that each antibody pair needs to be matched up and cross-reactivity for all pairs determined. That involves not only the cross-reactivity of each capture antibody but also the cross-reactivities of all secondary antibodies, including the extent of secondary-to-secondary interactions. It is not uncommon to sort through several antibody pairs before finding compatible sets. Consider now having to do the same for an antibody array comprising hundreds to thousands of analyte-specific pairs. See Chapter 6 "Protein Microarray Applications" for additional examples and discussion.

Protein profiling formats

Protein expression profiling (protein differential display) using microarrays is considered an important new tool for proteomic discovery. It is similar in concept and approach to the gene expression microarray for mRNA profiling. Sreekumar and Chinnaiyan (2002) describe a general approach for using the microarray to monitor protein expression in cancer and normal tissues. Here are the steps:

1. Extract total protein from cancer and normal cells using a detergent, e.g., 1% NP40.
2. Remove excess detergent from the lysate by adsorption onto a solid phase (e.g., beads).
3. Determine total protein concentration for each lysate.
4. Label equal amounts of protein from cancer and normal cell lysate, e.g., Cy5-NHS labeling of cancer lysate, Cy3-NHS labeling of normal lysate.
5. Remove free dye from dye-labeled proteins by gel filtration chromatography.
6. Mix Cy5-protein with Cy3-protein purified lysates.
7. Concentrate mixture.
8. Apply concentrate to antibody array.
9. Wash array free of unbound antigens, then perform confocal scan.
10. Analyze data.

A number of potential issues should be considered. First, detergent extraction can be problematic. Not all proteins will extract or will extract to the same extent. The amount of protein present can influence the efficiency and stability of detergent micelle formation. Inefficient removal of detergent as well as irreversible partitioning of proteins onto a solid support during purification is likely. If differences in protein abundance of test and control cell lysates existed before processing, significant differential loss could occur following processing.

Individual proteins can have very different labeling efficiencies depending upon concentration, pH, ionic strength, and the number and accessibility of dye-reactive amino acid residues (Kodadek, 2001). As with the labeling of nucleic acids, Cy5 and Cy3 or other dyes may demonstrate different labeling efficiencies for the same protein. Dye-labeled proteins may differentially adsorb onto the solid phase used for purification. Concentrating may do more harm than good if proteins denature and aggregate, forming protein complexes. If such complexes are applied to the antibody array, both false positive and false negative associations are likely.

Near-term biomedical applications

Cytokines

A number of commercial antibody-based microarrays for multiplexed cytokines analysis are now available (Beckman Coulter; BD Biosciences; Panomics; Pierce; S&S; Zyomyx; and others). Cytokines are essentially biomarkers of cell injury, inflammation, and apoptosis. They are released by cells in culture in response to drug action (Turtinen et al., 2004) or are elevated in serum in various disease states. Moreover, numerous cytokines are involved in cellular response and many serve as dual effectors (Asao and Fu, 2000). As a result, anticytokine microarrays are being evaluated in drug discovery for off-target toxicity testing to replace standard ELISA plate formats.

Huang et al. (2002) prepared an antibody array for the simultaneous detection of 43 cytokines. They were able to verify the down-regulation of MCP-1 cytokine in transfected cells (human glioblastoma cells transfected with *cx43* expression vector) relative to control cells. The antibody array is an emerging technology. In at least one study based upon the use of a commercial membrane format, the cytokine microarray failed to accurately determine cytokine levels in bacterial and lipopolysaccharide (LPS)-stimulated whole human blood (Copeland, 2004).

Autoimmune diseases and allergies

Advancement in autoimmune and inflammatory disease treatment and diagnosis represents a critical worldwide need ranking in importance only behind management of cardiovascular disease and cancer to the medical practitioner. The list of related diseases is long; major classes include rheumatoid arthritis, asthma, diabetes type I, multiple sclerosis, and inflammatory bowel disease.

Antigen arrays, also described as *reverse-phase protein arrays* (Paweletz et al., 2001), involve the immobilization of proteins to serve as bait for various protein–protein interactions (Sreekumar and Chinnaiyan, 2002). For example, Joos et al. (2000) printed down various autoantigens present in sera with known associations with various autoimmune diseases such as Graves' disease; lupus; connective tissue disease, and others. The group then screened various sera for the presence of autoantibodies. By immobilizing on the array a serial dilution series for each antigen, the titers for these antibodies could be determined.

Feng et al. (2004) prepared an antigen microarray on a polystyrene support comprising 15 autoantigens useful for the detection of autoantibodies involved in rheumatoid autoimmune diseases. De Vegvar et al. (2003) used antigen microarrays to examine epitope-specific antiviral antibody responses in vaccine trials in an animal model for human immunodeficiency virus (HIV) infection. Hueber et al. (2002) recently reviewed different formats for antigen microarrays. A brief description is also provided in an article by Robinson et al. published on a website maintained by P.J. Utz's laboratory at Stanford University Medical School, http://www.stanford.edu/group/utzlab/autoantibodies.htm.

Mezzasoma and coworkers (2002) used antigen microarrays to determine the levels of infectious agents in human sera. Antigens (*Toxoplasma gondii*, rubella virus, cytomegalovirus, herpes simplex virus, or ToRCH antigens) were printed down in an array format. Serum samples were applied and serodiagnosis determined using a sandwich assay employing fluorescently labeled secondary antibodies directed toward the primary sera antibodies.

The typing of various allergens using the antigen microarray has also met with success due in part to the availability of recombinant allergens (content). Jahn-Schmid and coworkers (2003) examined the analytical performance of an allergen (grass and tree pollen) microarray for the detection of allergen-specific serum IgG in sera of 51 patients. While considerable intra- and interassay variation was observed for some allergens, the sensitivity and specificity of the microarray was comparable to conventional ELISA. Shrefler et al. (2004) constructed an antigen array comprising an overlapping series of peptide probes representing epitopes associated with the major peanut allergens. Examining 77 patient sera, the group found considerable variation in patient IgE epitope profiling suggesting such population heterogeneity might be of prognostic value.

Finally, Nishizuka et al. (2003) have begun the arduous task of proteomic profiling the NCI-60 cancer cell lines based upon high density arraying of cell lysates. The lysates are prepared in a urea denaturing buffer and maintained in a reduced state with dithiothretol (DTT). This allows opportunity for additional assessment from 2D polyacrylamide gel electrophoresis (PAGE) gels. Serial dilutions of each protein lysate were printed down onto the substrate [fluorescent antibody staining technique (FAST®) slides, nitrocellulose-coated glass; Schleicher & Schuell BioScience 2003].

Monoclonal antibodies screened by Western blotting to lysate were used for detection and SYPRO ruby protein stain (Molecular Probes, Eugene, OR) for determination of total protein. A total of 52 proteins (those with high specificity antibody recognition) were analyzed in lysates. Of these, 31 were matched to cDNA and GeneChip microarrays, and 19 of 31 showed significant correlation between the two gene expression formats used earlier in characterizing the NCI-60 cell lines (Ross et al., 2000; Scherf and Ross, 2000).

How do the expressions of these proteins correlate with the corresponding mRNA profiles? The answer is that *structure-related proteins are almost always better correlated with mRNA levels across the 60 cell lines.* These kinds of studies may allow us to eventually unravel the mysteries surrounding the complex relationships of transcriptional and translational events.

Future medicine: Pharmacoproteomics or pharmacogenomics?

The end result for drug development is to successfully supply cost-effective drugs that provide better patient treatment, offering cures from new therapies and improvements in disease management including the development

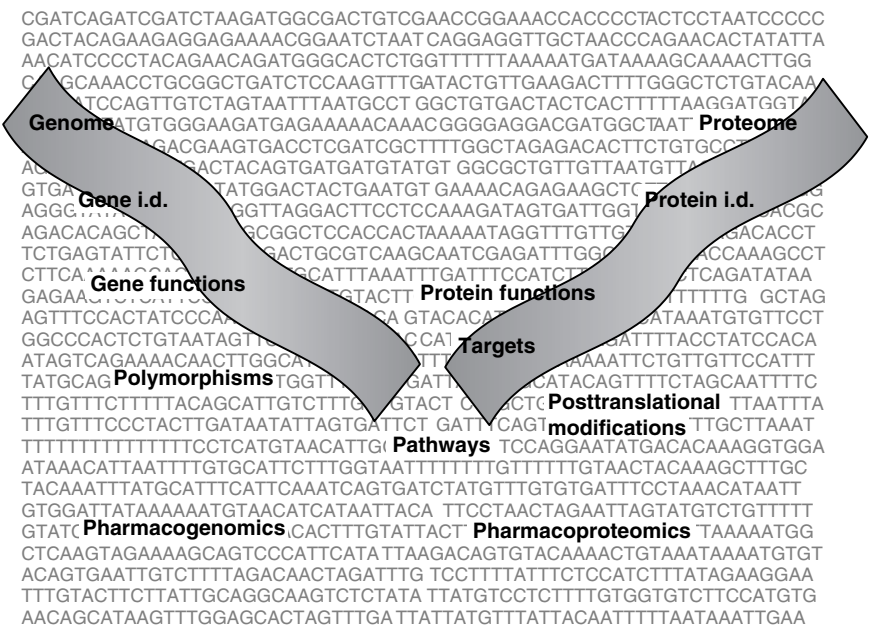


Figure 1.17 Pathways to molecular medicine.

of diagnostic tests that will more rapidly and more accurately determine disease states and monitor treatments. What road should we travel in the future to meet this need? Pharmacogenomics or pharmacoproteomics? See Figure 1.17.

Jain (2004) defines *pharmacoproteomics* as the use of proteomic technologies in drug discovery and development. Jain's contention is that pharmacoproteomics rather than genotyping will take the lead role in promoting the practice of personalized medicine that is anticipated to enter clinical practice by 2010. Key to that success will be the continued application of protein chips, enabling future discovery and development of drugs for personalized therapy and point-of-care diagnostics.

We may also find that Sabatini's reverse transfection method of creating "live cell" microarrays offers even greater advantage for drug discovery and development (Ziauddin and Sabatini, 2001). Their method, relying on arrayed cDNA expression vectors to transform adherent cells, provides localized real-time gene expression analysis of the putative gene product.

The live cell microarray may eventually displace the use of protein expression microarrays for identifying drug targets, provided that more extensive libraries of full-length cDNAs needed to express the complete

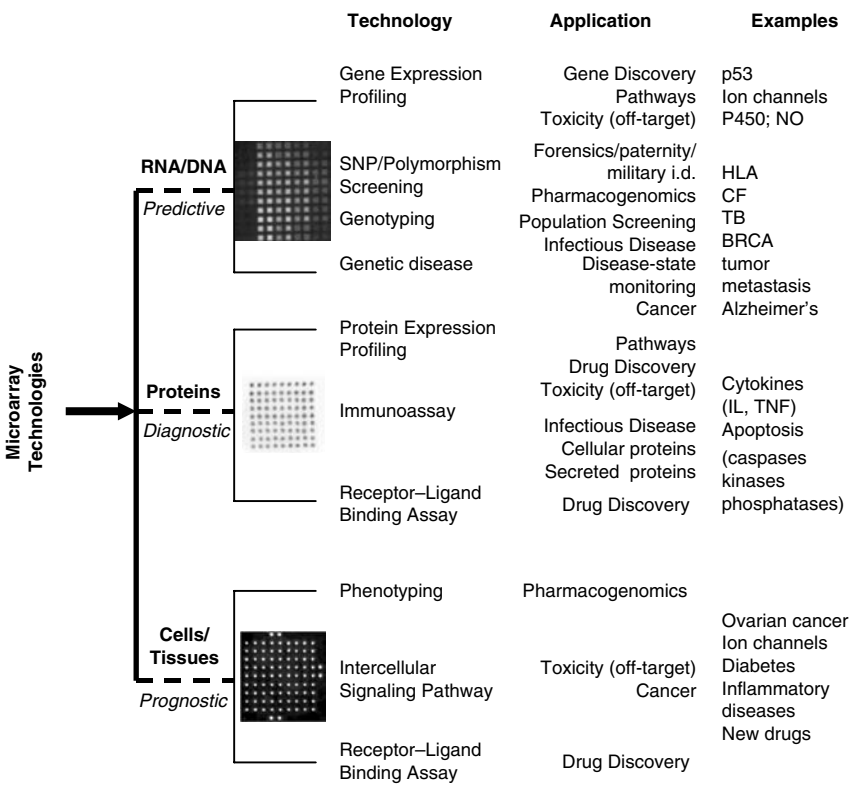


Figure 1.18 (See color insert following page 116.) Applications of microarray technologies.

protein become available. The approach could potentially be utilized for high-throughput screens to uncover genotype–phenotype relationships. Pharmacogenomics should not be ruled out just yet (Ginsburg and McCarthy, 2001).

What appears certain is that microarrays in one omics-based form or another (DNA, protein, cell, tissue, or small molecule) are playing increasingly important roles in drug development and diagnostics (Figure 1.18). Perhaps omics will eventually evolve into an integrated *systems biology* approach (Ideker et al., 2001) in which we monitor metabolic pathways, transport, compartmentalization, degradation, etc. and interrelationships of small molecules, cell surfaces (carbohydrate microarrays; Wang et al., 2002), and biopolymers alike (Figure 1.19). We are only at the beginning of molecular medicine and molecular diagnostics development and live in a truly exciting era for science, technology, and personalized medicine.

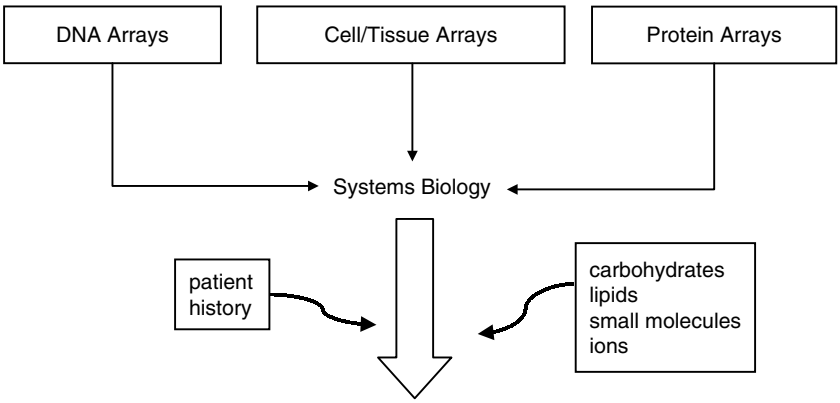


Figure 1.19 Molecular diagnostics. The future of the biochip rests on its ability to serve as a fully integrated device for multiplexed analysis.

References

Asao, H. and Fu, X.Y., Interferon-γ has dual potentials in inhibiting or promoting cell proliferation, *J. Biol. Chem*. 275, 867–874, 2000.

Cahill, D.J., Protein and antibody arrays and their medical applications, *J. Immunol. Meth*., 250, 81–91, 2001.

Clarke, P.A., ter Poele, R., Wooster, R., and Workman, P., Gene expression microarray analysis in cancer biology, pharmacology, and drug development: progress and potential, *Biochem. Pharmacol.* 62, 1311–1336, 2001.

Copeland, S., Siddiqui, J., and Remick, D., Direct comparison of traditional ELISAs and membrane protein arrays for detection and quantification of human cytokines, *Immunol. Meth*., 284, 99–106, 2004.

Cunningham, M.J., Genomics and proteomics: the new millennium of drug discovery and development, *Pharmacol. Toxicol. Meth*., 44, 291–300, 2000.

de Vegvar, H.E.N., Amara, R.R., Steinman, L., Utz, P.J., Robinson, H.L., and Robinson, W.H., Microarray profiling of antibody responses against simian–human immunodeficiency virus: post-challenge convergence of reactivities independent of host histocompatibility type and vaccine regimen, *J. Virol*., 77, 11125–11138, 2003.

Ekins, R., Chu, F., and Biggart, E., Fluorescence spectroscopy and its application to a new generation of high sensitivity, multi-microspot, multianalyte, immunoassay, *Clin. Chim. Acta*, 194, 91–114, 1990.

Ekins, R.P., Ligand assays: from electrophoresis to miniaturized microarrays, *Chemistry*, 44, 2015–2030, 1998.

Feng, Y., Ke, X., Ma, R., Chen, Y., Hu, G., and Liu, F., Parallel detection of autoantibodies with microarrays for rheumatoid diseases., *Clin. Chem*., 50, 416–422, 2004.

Ginsburg, G.S. and McCarthy, J.J., Personalized medicine: revolutionizing drug discovery and patient care, *Trends Biotechnol*., 19, 491–496, 2001.

Griffin, T.J. and Abersold, R., Advances in proteomic analysis by mass spectroscopy, *J. Biol. Chem*., 276, 45497–45500, 2001.

Haab, B.B., Dunham, M.J., and Brown, P.O., Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions, *Genome Biol*. 2, 0004.1–0004.13, 2001.

Hanash, S., Disease proteomics, *Nature*, 422, 226–232, 2003.

Hodges, L.C., Cook, J.D., Lobenhofer, E.K., Li, L., Bennett, L., Bushel, P.R., Aldaz, C.M., Afshari, C.A., and Walker, C.L., Tamoxifen functions as a molecular agonist inducing cell cycle-associated genes in breast cancer cells, *Mol. Cancer Res*., 1, 300–311, 2003.

Huang, R., Lin, Y., Wang, C.C., Gano, J., Lin, B., Shi, O., Boynton, A., Burke, J., and Huang, R.P., Connexin 43 suppresses human glioblastoma cell growth by down-regulation of monocyte chemotactic protein 1, as discovered using protein array technology, *Cancer Res*., 62, 2806–2812, 2002.

Hueber, W., Utz, P.J., Steinman, L., and Robinson, W.H., Autoantibody profiling for the study and treatment of autoimmune disease, *Arthritis Res*., 4, 290–295, 2002.

Huels, C., Muellner, S., Meyer, H.E., and Cahill, D.J., The impact of protein biochips and microarrays on the drug development process, *Drug Discovery Today*, 7, S119–S124, 2002.

Ideker, T., Thorsson, V., Ranish, J.A., Christmas, R., Buhler, J., Eng, J.K., Bumbardner, R., Goodlett, D.R., Aeberold, R., and Hood, L., Integrated genomic and proteomic analysis of a systematically perturbed metabolic network, *Science*, 292, 929–934, 2001.

Iyer, V.R., Eisen, M.B., and Ross, D.T., The transcriptional program in the response of human fibroblasts to serum, *Science*, 283, 83–87, 1999.

Jahn-Schmid, B., Harwanegg, C., Hiller, R., Bohle, B., Ebner, C., Scheiner, O., Muller, M.W., Allergen microarray: comprison of microarray using recombinant allergens with conventional diagnostic methods to detect allergen-specific serum immunoglobulin E, *Clin. Exp. Allergy*, 33, 1443–1449, 2003.

Jain, K.K., Role of pharmacogenomics in the development of personalized medicine, *Pharmacogenomics* 5, 331–336, 2004.

Joos, T.O., Schrenk, M., Hopfl, P., Kroger, K., Chowdhury, U., Stoll, D., Schorner, D., Durr, M., Herick, K., Rupp, S., Shon, K., and Hammerle, H., A microarray enzyme-linked immunosorbent assay for autoimmune diagnostics, *Electrophoresis*, 21, 2641–2650, 2000.

Katsuma, S. and Tsujimoto, G., Genome medicine promised by microarray technology, *Expert Rev. Diagn*., 1, 377–382, 2001.

Kodadek, T., Protein microarrays: prospects and problems, *Chem. Biol*., 8, 105–115, 2001.

Kricka, L.J., *Microchips: The Illustrated Hitchhiker's Guide to Analytical Microchips*, AACC Press, Washington, D.C., 2002.

Lillie, J., Probing the genome for new drugs and targets with DNA arrays, *Drug Development Res*., 41, 160–172, 1997.

Lobenhofer, E.K., Bennett, L., Cable, P.L., Li, L., Bushel, P.R., and Afshari, C.A., Regulation of DNA replication fork genes by 17β-estradiol, *Mol. Endocrinol*., 16, 1215–1229, 2002.

MacBeath, G. and Schreiber, S.L., Printing proteins as microarrays for high- throughput function determination, *Science*, 289, 1760–1763, 2000.

Mezzasoma, L., Bacarese-Hamilton, T., DiCristina, M., Rossi, R., Bistoni, R., and Crisanti, A., Antigen microarrays for serodiagnosis of infectious diseases, *Clin. Chem.*, 48, 121–130, 2002.

Miller, J.P., Back to basics in the "omics" rush, *Targets*, 1, 4–5, 2002.

Neamati, N. and Barchi, J.J., Jr., New paradigms in drug design and discovery, *Curr. Topics Med. Chem.*, 2, 211–227, 2002.

Nishizuka, S., Charboneau, L., Young, L., Major, S., Reinhold, W.C., Waltham, M., Kouros-Mehr, H., Bussey, K.J., Lee, J.K., Espina, V., Munson, P.J., Petricoin, E., Liotta, L.A., and Weinstein, J.N., Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays, *Proc. Natl. Acad. Sci. USA*, 24, 14229–14234, 2003.

Nuwaysir, E.F., Bittner, M., Trent, J., Barrett, C.J., and Afshari, C.A., Microarrays and toxicology: The advent of toxicogenomics, *Mol. Carcinogenesis*, 24, 153–159, 1999.

Paweletz, C.P., Charboneau, L., Bichsel, V.E., Simone, N.L., Chen, T., Gillespie, J.W., Emmert-Buck, M.R., Roth, M.J., Petricoin, E.F., III, and Liotta, L.A., Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front, *Oncogene*, 12, 1981–1989, 2001.

Petricoin, E.F. and Liotta, L.A., SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer, *Curr. Opinion Biotechnol.*, 5, 24–30, 2004.

Rew, D.A., DNA array technology in cancer research, *Eur. J. Surg. Oncol.*, 27, 504–508, 2001.

Rininger, J.A., DiPippo, V.A., and Gould-Rothberg, B.E., Differential gene expression technologies for identifying surrogate markers of drug efficacy and toxicity, *DDT*, 5, 560–568, 2000.

Roberts, C.J., Nelson, B., and Marton, M.J., Signaling and circuitry of multiple MAKP pathways revealed by a matrix of global gene expression profiles, *Science*, 287, 873–880, 2000.

Ross, D.T., Scherf, U., Eisen, M.B., Perou, C.M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S.S., Van de Rijn, M., Waltam, M., Pergamenschikov, A., Lee, J.C.F., Lashkari, D., Shalon, D., Myers, T.G., Weinstein, J.N., Botstein, D., and Brown, P.O., Systematic variation in gene expression patterns in human cancer cell lines, *Nat. Genet.*, 24, 227–235, 2000.

Scherf, U. and Ross, D.T., A gene expression database for the molecular pharmacology of cancer, *Nat. Genet.*, 24, 236–244, 2000.

Shoida T., *J. Environ*. Application of DNA microarray to toxicological research, *Pathol. Toxicol. Oncol.*, 23, 13–31, 2004.

Shreffler, W., Beyer, K., Chu, T., Burks, A., Sampson, H., Microarray immunoassay: association of clinical history, *in vitro* IgG function, and heterogeneity of allergenic peanut epitopes, *J. Allergy Clin. Immunol.*, 113(4), 776–782, 2004.

Sreekumar, A. and Chinnaiyan, A.M., Protein microarrays: a powerful tool to study cancer, *Curr. Opinion Mol. Therap.*, 4, 587–593, 2002.

Turtinen, L.W., Prall, D.N., Bremer, L.A., Nauss, R.E., and Hartsel, S.C., Antibody array-generated profiles of cytokine release from THP-1 leukemic monocytes exposed to different amphotericin B formulations, *Antimicrob. Agents Chemother.*, 48, 396–403, 2004.

Tyers, M. and Mann, M., From genomics to proteomics, *Nature*, 422, 193–197, 2003.

Valle, R.P.C. and Jendoubi, M., Antibody-based technologies for target discovery, *Curr. Opinion Drug Discovery Dev.*, 6, 197–203, 2003.

Wang, D., Liu, S., Trummer, B.J., Deng, C., and Wang, A., Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells, *Nat. Biotechnol.*, 20, 275–281, 2002.

Weinstein, J.N., 'Omic' and hypothesis-driven research in the molecular pharmacology of cancer, *Curr. Opinion Pharmacol.*, 2, 361–365, 2002.

Woodbury, R.L., Varnum, S.M., and Zanger, R.C., Elevated HGF levels in sera from breast cancer patients detected using a protein microarray ELISA, *J. Proteome Res.*, 1, 233–237, 2002.

Zanders, E.D., Gene expression analysis as an aid to the identification of drug targets, *Pharmacogenomics*, 1(4), 375–384, 2000.

Ziauddin, J. and Sabatini, D.M., Microarrays of cells expressing defined cDNAs, *Nature*, 411, 107–110, 2001.

chapter two

Commercial microarrays

Introduction

In this chapter, we will review the commercialization of microarrays. The intent is to look broadly at commercial efforts while recognizing key technological developments. As with any emerging field, development efforts have met with triumphant technological successes and commercial failures. However, all efforts have contributed in some manner to progress along the forward path and made the microarray a commercial reality. While this chapter examines microarray products currently offered in the marketplace, it also discusses certain companies that no longer do business, principally because of their contributions to the development of this field.

Microarrays comprise a number of different formats, all of which have been adopted in one form or another for commercial use. They can be classified as *in situ*, *ex situ*, and *electronically active* or addressable arrays. A *fiberoptic* array is an optically addressable, *ex situ* array (Figure 2.1). While the origin of array-based technology can be traced back several decades (see review by Matson and Rampal, 2003), we examine here the major breakthrough in commercialization that began in the early 1990s.

In situ *arrays*

Affymetrix is credited with the first commercial development of a oligonucleotide array based upon its pioneering work in the photolithographic masking process coupled with standard DNA chemical synthesis (Pease et al., 1994). The Affymetrix microarray (GeneChip brand array) is constructed by synthesis of oligonucleotides from the substrate (surface) of the chip. Hence, the term *in situ* applies to Affymetrix's synthesis process.

It involves photolithographic masking of specific areas on the substrate, followed by removal of the mask from a selected region, thereby exposing light-sensitive deprotecting groups on the previously coupled monomers to ultraviolet (UV) light. With the base fully deprotected and the light turned off, the chip is next flooded with bulk reagents containing the next monomer

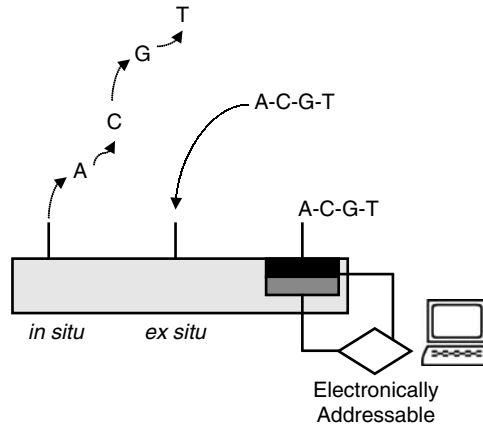


Figure 2.1 DNA array types.

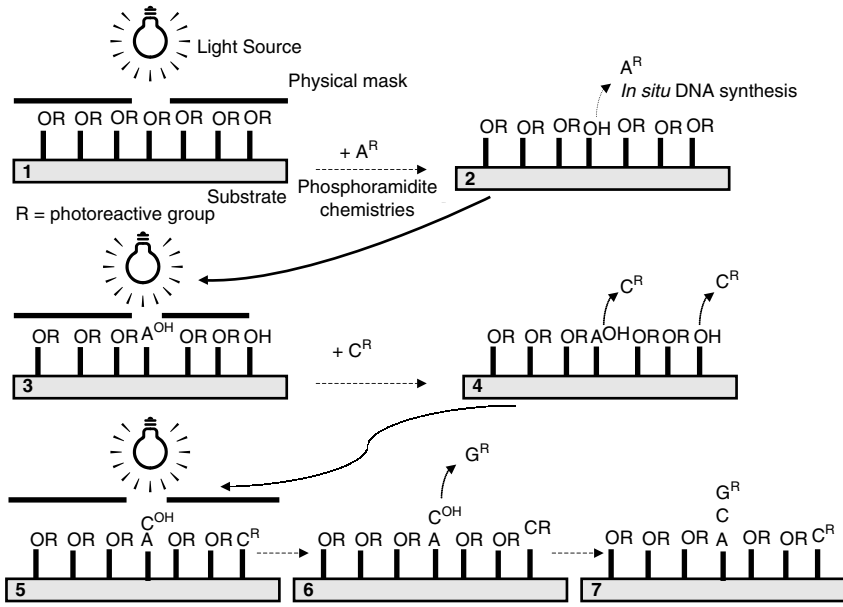


Figure 2.2 Photolithographic-based DNA synthesis. (From Pease, A.C. et al., *Proc. Natl. Acad. Sci. USA*, 91, 5022–5026, 1994. With permission.)

phosphoramidite that can now react. The growing oligonucleotide chain is thus extended by one base and the process is repeated (Figure 2.2). Current GeneChip arrays are produced with 25-mer oligonucleotides and contain over 500,000 features (20 μm^2 each) on a 1.28-cm × 1.28-cm chip. For example, the Human Genome U133A 2.0 chip contains over 22,000 probe sets for interrogation of 14,500 human genes (Figure 2.3).

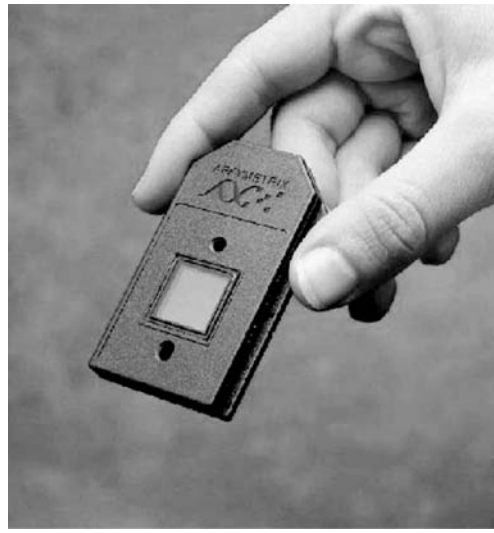


Figure 2.3 Affymetrix's GeneChip. (Photo courtesy of Affymetrix, Inc., Santa Clara, CA.)

While Affymetrix's early entry into the DNA microarray market afforded it a formidable position, the company has competitors. In order to commercialize the *in situ* array, it became clear that access to certain intellectual properties, especially the Southern patent (Oxford Gene Technologies or OGT), was required. Affymetrix obtained a license through a business relationship with Beckman Coulter which originally held the first and exclusive Southern license and later relinquished its exclusivity. Beckman Coulter and Affymetrix entered into a joint venture with Array Automation LLC to automate the processing of Affymetrix chips. Now that license to the Southern technology is available from OGT, others are permitted to commercialize *in situ* microarrays by alternative chemical synthesis approaches.

For example, Rosetta Inpharmatics (now a Merck subsidiary) and Protogene (no longer in business) relied upon the "sequential" dispensing of reactive phosphoramidite monomers by ink-jet printing to discrete locations on a substrate from which to grow the oligonucleotide sequence. Early work on the Rosetta–Agilent microarray (commercialized by Agilent) demonstrated the advantages of 60-mer oligonucleotide probes for gene expression analysis (Hughes et al., 2001). Agilent's SurePrint™ Whole Human Gene microarray slides contain 41,000 elements (135 μ each) comprising 60-mer probes printed *in situ* onto standard 1-in. × 3-in. slides.

Another approach that has attracted considerable attention has been the introduction of the virtual (photolithographic) mask using a digital micromirror device (DMD, Texas Instruments [Note: The Texas Instrument technology used in the DMD chip is referred to as the Digital Light Processing™or DLP™ technology]). Singh-Gasson et al. (1999) of the University of Wisconsin

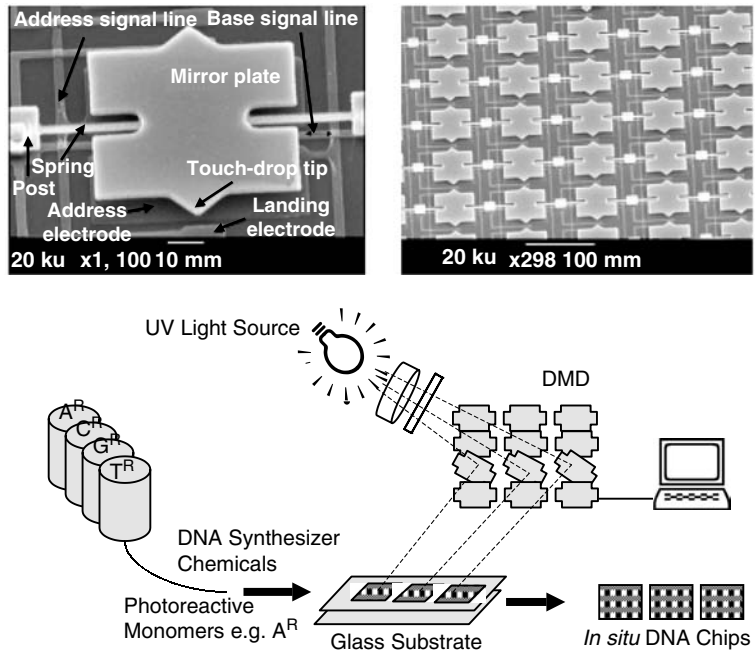


Figure 2.4 Digital micromirror device (DMD). (From Singh-Gasson, S. et al., *Nat. Biotechnol.*, 17, 974–978, 1999 and Lee, P. et al., *J. Micromechanics Microeng.*, 13, 474–481, 2003. With permission.)

built a maskless array synthesizer (MAS) incorporating the DMD processor (Figure 2.4). They were able to produce *in situ* oligonucleotide arrays of 76,800 features (16 μm^2) on standard microscope slides with a stepwise yield of about 95% for 18-mer in 12 hr.

Nimblegen adopted the MAS approach to create arrays containing as many as 786,432 probe features in a 2.3-cm^2 area on a slide. Oligonucleotides in the range of 24 to 90-mer are synthesized with good stepwise yields. In order to accommodate current scanner resolution, the densities have been reduced. Two formats covering the area of a standard slide are provided: 195,000 probes (1:4) and 390,000 probes (1:2). A 24-mer array can be synthesized within 2 to 4 hr. The average stepwise yield is 97.5%.

Febit introduced Geniom® One as a "benchtop microarray facility." The product is a fully integrated instrument providing for DNA design, synthesis, hybridization, and analysis. An eight-channel microfluidic device (DNA processor®) permits the simultaneous construction of eight individual microarrays (Figure 2.5). Each microarray is produced with up to 6000 features (34 μm^2 each) or about 48,000 probes per disposable DNA processor. A complete experiment from design to hybridization analysis can be finished within 24 hr. Baum et al. (2003) used a four-channel device that permitted the synthesis of 25 mer at 4 × 12,880 = 51,520 features. The detailed study describes results comparing the DNA processor to the Affymetrix GeneChip

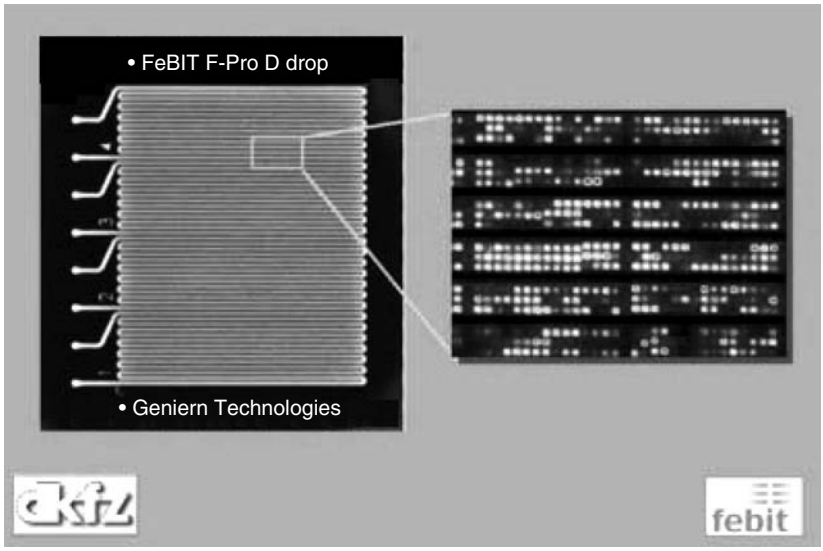


Figure 2.5 *In situ* DNA synthesis in channels using DMD. (Photo courtesy of Febit AG, Mannheim, Germany.)

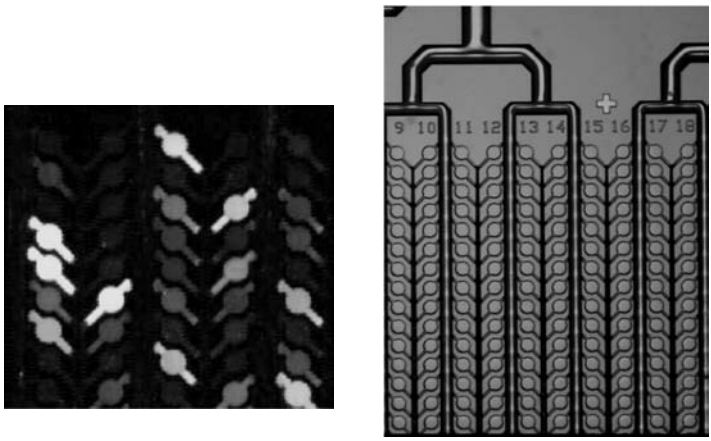


Figure 2.6 XeoChip®. (Photo courtesy of Xeotron Corporation, Houston, TX.)

YG-S98 yeast genome chip. In side-by-side experiments, the Febit device performed in concordance to the Affymetrix chip.

Xeotron also uses a DMD processor to produce arrays in microchannels (Figure 2.6). Standard oligonucleotide phosphoramidite-based synthesis is performed. The virtual masking is directed instead toward deprotection of a photolabile acid. The process is called PGA or photogenerated acid. Once the acid is liberated, the trityl group on the attached base can be removed to allow base extension to occur. Stepwise yields of 98.5% are reported for the XeoChip.

Arguably, a major disadvantage of the *in situ* format is that each coupling cycle is not 100% complete, thereby leading to reduced stepwise synthesis yields such that the final product at each site is a heterogeneous population containing truncated failure sequences. For the GeneChip process, the oligonucleotide probe may only be 82 to 97% pure at each location.

Investigations into the reasons for the reduction in stepwise yield on glass revealed that coupling reactions near the surface are relatively inefficient (LeProust et al., 2001). As Graves (1999) noted, inefficient stepwise couplings reduced the number of probes having error-free sequences at a particular site to somewhere between 1 and 36% of the total probe population. This would be acceptable for gene expression analysis or calling out a genotype, but is less so for single nucleotide polymorphism (SNP) scanning. However, match vs. mismatch ratiometric analysis and the process of "tiling" degenerate bases improves the performance of such arrays (Wang et al., 1998). In specific instances, the performance of GeneChip microarrays has been comparable to that of gel-based sequencing, and improvements in stepwise yields using the virtual masking approach and new coupling chemistries have been reported. Affymetrix recently launched a GeneChip for analysis of 100,000 SNPs. In retrospect, probe purity is not as problematic as once thought.

The other well-known disadvantages of chip-based microarrays are cost and manufacture turnaround time. A 25-mer *in situ* synthesis using the Affymetrix process theoretically requires $4 \times 25 = 100$ cycles. The number of cycles can be further reduced by applying algorithms to the design. However, there are limits even in software for what can be done to reduce the number of physical masks. Thus, the cost to manufacture via the physical photolithographic masking process requires larger scale production and has precluded significant offerings to the custom array market.

On the other hand, the virtual masking (or maskless) approaches have demonstrated rapid *in situ* oligonucleotide synthesis, making them well suited for laboratory scale with the capability to produce 24-mer to 90-mer probes (Nuwaysir et al., 2002). However, one issue regarding implementation of the DMD technology is control of stray light. Garland and Serafinowski (2002) considered a theoretical model to predict the impact of stray light at low photolithographic contrast ratios. This ratio essentially defines the spatial boundary required to reduce adjacent array element illumination.

At high contrast ratios (~10^5), for example, with standard photolithography, very little stray light illumination is observed. At low-to-moderate contrast ratios (10^3 to 10^4) typical of the DMD, stray light may be problematic. For example, the model predicts that the synthesis of 20 mer using DMD would result in the insertion of additional bases such that the major population of probes would be a mixture of 21-mer and 22-mer oligonucleotides. This assumes the use of the direct 5′ photodeprotection process. At a contrast ratio of 400, the fraction of the correct 20-mer sequence would be 0.36, while at a ratio of 200, the fraction is reduced to 0.13. Less of an issue exists with processes involving 5′ trityl deprotection of a base using a photogenerated

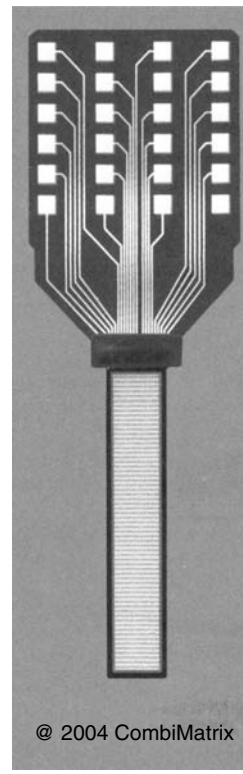


Figure 2.7 (See color insert following page 116.) CustomArray™. (Photo courtesy of CombiMatrix Corporation, Mukilteo, WA.)

acid (e.g., Xeotron's PGA). In this case, stray light-generated acid can be buffered out. Garland and Serafinowski (2000) demonstrated the stray light N + 1 insertion for a T_5 synthesis with 1% stray light. The occurrence of T_6 that resulted from photoacid deprotection could be significantly reduced by the addition of n-octylamine which served as an acid scavenger.

Finally, CombiMatrix has emerged with an *in situ* process for generating oligonucleotide arrays on microchips with 1024 (94-μ diameter) or 13,416 (44-μ diameter) features. The features are actually "digitally addressable" electrode pads covered with porous membrane reaction layers (Figure 2.7). Certain electrodes are turned on so as to generate protons over these electrodes. This permits controlled, site-specific deprotection of the attached base (removal of trityl) and allows the addition of the incoming phosphoramidite monomer. Trityl-phosphoramidite monomers and reagents are flooded over the chip, and the process is repeated through several cycles until the desired oligonucleotide is produced at that particular location.

Roche Diagnostics entered into a 15-yr contract with CombiMatrix in 2001 to supply oligonucleotide arrays for its markets. However, Roche announced a delay in the launch of its matriXarray product line. In the

meantime, CombiMatrix launched its own product called CustomArray in March 2004. The CustomArray product is intended for custom synthesis of up to 40-mer oligonucleotide probe arrays based upon 92-μ features with up to 902 user-defined probe elements. Sensitivity is reported to be 1.5 p*M* with 20% coefficient of variation (CV) (intra- and interchip) and a shelf life of 4 mo. The system provides electrochemical detection. CombiMatrix (www.combimatrix.com) has also initiated work on protein microarrays utilizing the porous reaction layer to link biotin using *in situ* chemistries generated from the electrode pads. Streptavidin is then used to bridge various antigens and antibodies to the surface through the covalently attached biotins (Dill et al., 2001).

Ex situ *or spotted arrays*

Ex situ (also known as spotted or printed) arrays have become very popular formats, especially for the building of custom noncommercial arrays used primarily by academic laboratories [see Association of Biomolecular Resource Facilities (ABRF) surveys on microarrays at www.abrf.org]. The printed cDNA microarray was largely developed from gene expression work originating in the laboratories of P.O. Brown and R.W. Davis at Stanford University (Schena et al., 1995). Plans for the construction of the microarrayer and split pin designs were available at the Brown lab website at http://cmgm.stanford.edu/pbrown/mguide/index.html. This enabled researchers to prepare their own microarrays appropriate for their particular experiments.

However, the demands for microarrays and pin technologies were sufficiently large to initiate a cottage industry arising from a number of precision engineering companies that had the expertise in fine tooling necessary to manufacture split pins. Others followed to produce microarrayers and develop slide scanners. Companies furnishing microscope slides turned their attention to the microarray community as well. Software companies were founded to offer packages for imaging and gene expression analysis. Synteni (founded by Stanford inventor Dari Shalon in 1994) was first to introduce custom microarray analysis based upon the new cDNA-based slide format termed gene expression microarray (GEM) technology. The company was later acquired by Incyte Genomics in 1997 and became Incyte Microarray Systems until the unit was purchased by Quark Biotech in 2002 for internal use for target discovery.

In 1999, an overview of the microarray world titled "The Chipping Forecast" appeared as a supplement to *Nature Genetics*. David Bowtell highlighted the 1999 industry leaders in equipment and services for microarrays. Five years later, the commercial landscape showed significant changes. The most noticeable changes were the consolidations of microarrayer businesses. Three of the five companies (Cartesian, Genomic Solutions, and BioRobotics) listed by Bowtell now exist within a single company known as Genomic Solutions which in turn is owned by Harvard Biosciences. GeneMachines, although not listed in Bowtell's compendium, became a leading producer of microarrayers. It too was purchased by the Harvard Biosciences group and is now part of Genomic Solutions.

Genetic MicroSystems was purchased by Affymetrix. Beecher Instruments is now primarily focused on tissue microarray systems. New microarrayer startups such as GeneXP (BioGridArray™ custom microplate-based microarrays and high-throughput BioGridArrayer™ systems) and Genetix which expanded on the success of its QBot colony pickers with a product line of QArray™ printers continue to appear. TeleChem International, well known for its quill pin technologies, has also been motivated to bring out a line of microarrayers marketed as the NanoPrint™ product line.

About 80% of *all* microarray work involves the use of glass slides (ABRF Microarray Survey, http://www.abrf.org/ResearchGroups/Microarray/EPosters). The most common substrate is a standard glass microscope slide that has been surface treated with a coating such as polylysine that will bind nucleic acids (also proteins). Special graded slides for microarrays (ultraflat, with low intrinsic fluorescence) have been introduced. cDNA is spotted down on the surface using a precision x-y-z plotter (microarrayer) to which are attached either dispenser heads (e.g., Packard's piezoelectric dispenser, Cartesian's solenoid-based dispenser) for noncontact printing or pins (e.g., Telechem's split pin) that touch down on the surface.

The most common nucleic acid spotted method remains the cDNA prepared by reverse transcriptase polymerase chain reaction (RT-PCR) of cloned material. Thus, cDNA libraries can be arrayed as desired. The main advantage of using cDNA is that both known and unknown sequences can be arrayed and used for gene expression analysis. However, following completion of the human genome sequence draft and other genomes, the trend is to shift back to the use of synthetic oligonucleotide probes of well defined sequences. This is evident in the most recent survey by ABRF (2003) covering future directions within core facilities.

For gene expression, oligonucleotides in the range of 40-mer to 70-mer are used, while 15-mer to 25-mer sizes are sufficient for resequencing protocols. Again, the Southern and Affymetrix patents have prevailed, forcing several companies to acquire licenses. Operon (now owned by Qiagen) and Genomic Solutions (now owned by Harvard Biosciences) discontinued sales of printed oligonucleotide arrays in view of Affymetrix's extensive patent portfolio.

The commercial success of the spotted microarray, like that of Affymetrix's GeneChip, is content-driven. In order to provide customers with comprehensive gene expression microarray products, manufacturers must obtain gene-specific annotated sequences covering genomes of major interest to the scientific community (e.g., genomes of humans, yeasts, and mice).

Prior to completion of the sequences of these genomes, such content was largely derived from public databases comprising expressed sequence tags (ESTs). The National Center for Biotechnology Information (NCBI) maintains several databases of interest such as GenBank (Wheeler et al., 2003). GenBank's dbEST contains a database of submitted ESTs that are matched if possible to known genes using a process called UniGene (www.ncbi.nlm.nih.gov/UniGene/index.html). The UniGene clone collections are available from government-licensed vendors. The clones are primarily assembled from the

IMAGE (Integrated Molecular Analysis of Genomes and Their Expression) consortium collection (http://image.llnl.gov/image/html/idistributors.shtml). The current U.S. distributors are:

American Type Culture Collection, Manassas, VA (www.atcc.org)
Open Biosystems, Huntsville, AL (www.openbiosystems.com)
Research Genetics, Carlsbad, CA (now owned by Invitrogen: www.invitrogen.com)

The European suppliers are:

MRCgeneservice, Cambridge, U.K. (www.hgmp.mrc.ac.uk/geneservice/index.shtml)
RZPD German Resource Center for Genome Research, Berlin (www.rzpd.de/services)

Since completion of the Human Genome Project, other extensive privately held clone libraries are now available including Incyte's LifeSeq® sets comprising more than 13 million ESTs and 18,000 genes (www.incyte.com). Invitrogen also offers clone collections including those acquired from its purchase of Research Genetics. In 2001, Invitrogen entered into a worldwide distribution agreement with The Institute for Genomic Research (TIGR) to acquire TIGR's extensive libraries containing over 300,000 cDNA and genomic clones including 70,000 human, 10,000 rat, and 70,000 microorganism clones (http://clones.invitrogen.com).

A number of companies offer spotted microarray products for gene expression analysis. BD Biosciences (formerly Clontech) offers a variety of printed microarrays on nylon, glass, and plastic slides under its *BD* Atlas™ product brand including a 12,000-human gene microarray. Both oligonucleotide (80-mer) and cDNA probe arrays are available, depending upon the genome. SuperArray Biosciences offers thematic cDNA micorarrays on glass-supported nylon under the GEArray™ name (www.superarray.bz). Agilent licensed Incyte's cDNA clones and bioinformatics for gene expression spotted microarray products in 2001. Lists of additional suppliers have been published in *The Scientist* (2003) and *Nature Genetics* (2002).

Several studies have compared the performance of *in situ* and *ex situ* spotted microarrays in gene expression analysis using commercial sources. Tan et al. (2003) evaluated Amersham's CodeLink arrays (30-mer probes), Affymetrix's GeneChip (25-mer probes), and Agilent's cDNA array format using the same cRNA pools. A total of 2009 genes common to all three platforms were analyzed. While intraplatform correlation coefficients were high ($r = >0.9$), the Pearson's correlation coefficients were Affymetrix–Amersham (0.59), Agilent–Amersham (0.48), and Agilent–Affymetrix (0.50). In other words, this study found no significant agreement as to the gene expression profile relationships obtained using these platforms, further suggesting

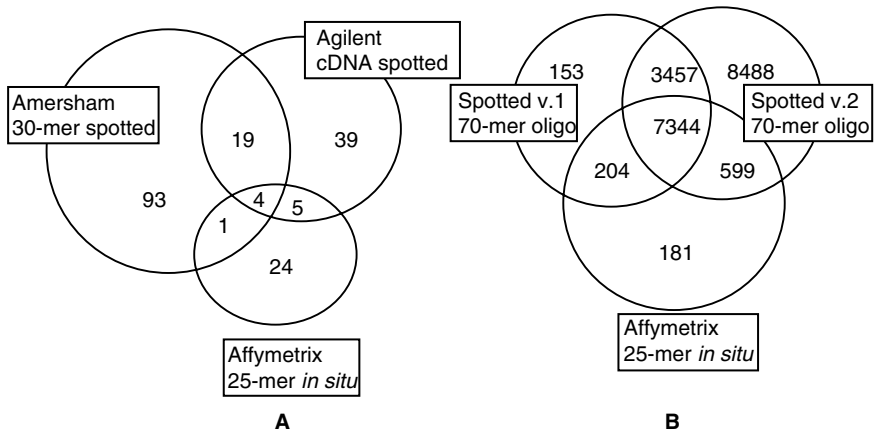


Figure 2.8 Microarray cross-platform showing differentially expressed gene clusters obtained from Amersham, Agilent, and Affymetrix products. (From Barczak, A., et al., *Genome Res.*, 13, 1775–1785, 2003 [Copyright 2003 Cold Spring Harbor Laboratory Press] and Tan, P.K. et al., *Nucleic Acid Res.*, 31(19), 5676–5684, 2003. With permission.)

that "cross-platform differences arise from the intrinsic properties of the microarrays themselves."

In another study, Barczak et al. (2003) compared GeneChip arrays to so-called "long" oligonucleotide arrays. A total of 7344 genes from the human genome were analyzed using the Affymetrix U95 GeneChip along with two spotted arrays comprising 70-mer probes (Operon Human Genome Oligo Set, versions 1 and 2). A good correlation for differential expression was obtained between the spotted 70-mer arrays and the *in situ* 25-mer arrays. However, the long oligonucleotide hybridization intensities were lower overall, leading to less reliable calls; when compared to the *in situ* arrays, the correlation was significantly reduced (r = ~0.6). Excluding 4467 genes with low intensity values improved the correlation (r = ~0.9) for the remaining 2877 common genes. Obviously probe selection is critical.

Another important point was made by this study: if possible, use reference samples with gene expression levels similar to those expected for the test samples. This should improve quantification at the low end. Moreover, the Tan et al., (2003) and Barczak (2003) studies also demonstrate the need for the standardization of commercial microarray products and reference standards so that both interlaboratory and cross-platform performances can be properly assessed. While commercial microarrays from different vendors all claim substantial gene probe "real estate," the number of probes in common is smaller. This makes it difficult to compare experiments across platforms, especially those from major suppliers such as Affymetrix, Agilent, and Amersham (Figure 2.8).

The spotted array is also used for SNP analysis based upon primer extension labeling of oligonucleotide or cDNA probes. The advantage of

using oligonucleotides is that they may be synthesized in good quantities and highly purified (98 to 99%) prior to attachment. For example, Sequenom first introduced chip-based high-throughput SNP screening with mass spectroscopy analysis. The original process involved the creation of arrays by immobilization of single-stranded PCR amplicons onto silicon chips [matrix-assisted laser desorption (MALD) targets] that were surface treated with N-succinimidyl (4-iodoacetyl) aminobenzoate.

Disulfide terminated primers were used to generate amplicons and the reduced template (thiol-DNA) was covalently coupled to the sulfide-reactive chip. A single-base primer extension reaction on the immobilized target was performed followed by the addition of 3-hydroxypicolinic acid (matrix). The extended primer sequence was then determined using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectroscopy (MassArray™) to reveal the SNP (Tang et al., 1999). In later refinements, it became unnecessary to use the anchored primer extension (APEX) approach with the analytical power of mass spectroscopy. Instead, a solution-phase primer extension was employed followed by direct dispensing of the extended primer onto individual elements of a silicon chip was preloaded with matrix (Buetow et al., 2001).

Orchid introduced a "zip code" array approach to SNP analysis by printing down capture oligonucleotides onto 384-gasket well skirted glass plates. The complementary oligonucleotide was incorporated as a tag in the extending primer. Following a solution-based single-base extension with labeled dideoxy nucleotide triphosphate (ddNTP) the labeled primer was captured and the SNP determined from the hybridized tag (Figure 2.9). The SNP genotyping instrumentation business was purchased from Orchid BioSciences by Beckman Coulter in 2002 and marketed as the the SNPstream® Genotyping System.

Low to medium density arrays have also entered the marketplace based upon the so-called array-of-arrays format. Genometrix first commercialized products using this format (printing down nucleic acid or antibody probes in a 96-well pattern on Teflon masked slides) to perform micro-ELISA or genotyping (Mendoza, 1999; Wiese, 2001). The products were marketed as services under the GenoVista Partnership Program name. The 96-well microarray platform was called the VistaArray. Unfortunately, Genometrix's service provider model failed to generate or sustain sufficient revenues.

Following the collapse of Genometrix, High-Throughput Genomics (HTG) purchased Genometrix's intellectual property portfolio and introduced an array within a well microtiter plate (ArrayPlate™) for gene expression analysis. Assays are based upon HTG's proprietary multiplexed nuclease protection assay (m-NPA). The primary advantage of m-NPA is that no amplification or sample processing of RNA need be conducted. Cell lysates may be added directly to the well and the released mRNA species captured and protected from nuclease digestion by sets of sequence-specific oligonucleotide probes. Once the RNA-probe hybrids are formed, other nucleic acids within the sample are digested with S1 nuclease.

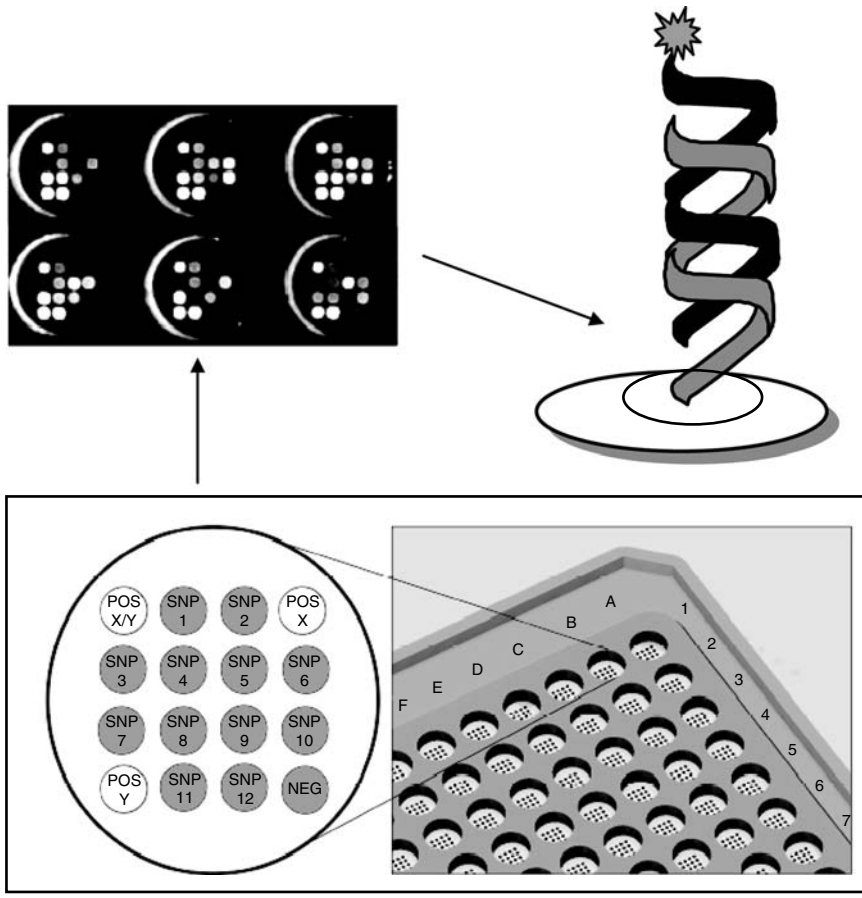


Figure 2.9 SNP detection using Orchid's SNPstream. (From Bell, J. et al., *Biotechniques*, 32, S70–S77, 2002. With permission.)

The nuclease-protected RNA strands can then be destroyed by alkaline hydrolysis, leaving behind the original probes that are now present in the sample in amounts quantitative to the protected RNA species. Probes are then captured within the array and their identities determined by the complementary sequence of the capture oligonucleotide. The amount of captured probe is estimated using a chemiluminescent reporter assay (Martel et al., 2002). This product is primarily focused on high-throughput drug-target gene profiling for drug discovery applications.

A multiplexed immunoassay has likewise been introduced based upon the ArrayPlate format in which antibody–oligonucleotide conjugates are assembled by hybridization to complementary capture probes. Beckman Coulter introduced the A^{2TM} MicroArray System in March 2004. The A squared or A^2 (array of arrays) approach involves the printing of a capture oligonucleotide "zip code" in the bottom of a rounded square 96-well

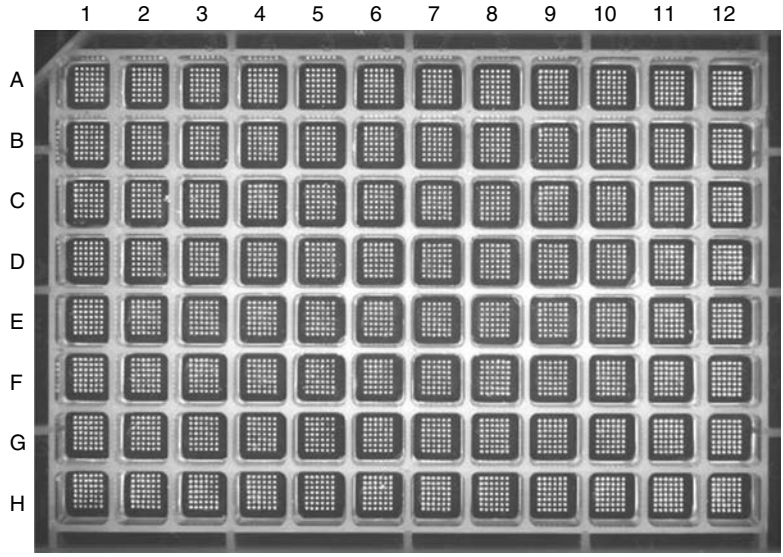


Figure 2.10 A² plate microarray.

polypropylene plate. Complementary oligonucleotides may then be conjugated to the customer's antibody (up to 13 different antibodies per array) and the antibody array assembled by the customer under hybridization conditions (Figure 2.10). The system includes a reader, fluorescent label detection reagents, and fully integrated software to perform simultaneously up to 13 individual quantitative micro-ELISAs (Song et al., 2004). The full assay can be automated using the company's automated liquid handling robots such as the Biomek® FX.

3D and 4D chips

Three dimensional (3D) chips are comprised of hydrogels or membranes that are supported by a planar substrate such as a glass slide. Four dimensional chips (4D) here refer to fabricated flow-through chips with well-defined pore channels such as channel glass. The so-called 3D surfaces are thought to offer higher probe density over microarrays constructed on planar (2D), nonporous surfaces, thereby leading to increased sensitivity and dynamic range. Such coatings may also reduce nonspecific background adsorption problems associated with planar substrates.

Whether such advantages over 2D surfaces are realized is dependent upon the application. In particular, such surfaces may be most beneficial for working with proteins. However, 3D chips have not yet produced the anticipated or sought-after impact in the current market despite the considerable hydrogel development efforts of Mirzabekov, Khrapko and colleagues at the Engelhardt Institute (Moscow) in the late 1980s and in collaboration with

the U.S. Argonne National Laboratory (ANL) in Illinois in 1994. In 1998, ANL entered into a research partnership that began the commercialization forays of Motorola and Packard Biosciences (see Chapter 3).

While large-scale commercialization of the 3D chip has taken considerably more time than expected, several companies now offer such surfaces. Motorola Life Sciences abandoned the Mirzabekov gel (Chapter 3) in favor of SurModic's PhotoLink® surface chemistry (www.surmodics.com), chemistry which it then brought to market as the CodeLink™. Amersham Biosciences acquired CodeLink from Motorola in 2002 and obtained a license to the Southern patents in 2003 to expand the product into the clinical diagnostic arena.

In addition to CodeLink, several other 3D microarrays have entered the marketplace including Perkin Elmer's HydroGel™ polyacrylamide gel microarray (obtained via the acquisition of Packard Biosciences) for proteins. Biocept prints down droplets of a polyethylene glycol hydrogel onto glass slides offering a DNA microarray product called 3D HydroArray™ (Gurevitch et al., 2001).

Membranes cast upon glass slides also fall into the 3D surface category. Most notable are the nitrocellulose-coated FAST slides offered by Schleicher & Schuell (S&S) BioScience. A cytokine micro-ELISA product under the trade name ProVision™ has been introduced in single-slide and 96-well spacing (64 usable wells) formats. It is called FAST®Quant (Harvey, 2003). Historically, nitrocellulose membrane has been used for the sequestering of both proteins and nucleic acids. The adaptation by S&S of this microporous (0.2-μ pores) material cast into a microarray format has been relatively straightforward.

The disadvantage of the nitrocellulose (NC) slide relates to resolution because biomolecules more easily diffuse from the surface than with planar arrays. Also, membranes suffer from considerable light scatter and higher intrinsic fluorescence, which are problematic for increased sensitivity. Membranes are better suited for detection of colorimetric or chemiluminescence reporters. Pall introduced a similar product, the Vivid™ GeneArray slide, based upon nylon-impregnated slides.

Flow-through biochips

Beattie et al. (1995) described the fabrication and use of a flow-through porous silicon genosensor prepared from channel glass. An array of square wells was first etched into a silicon wafer. The wafer was then bonded to a second wafer with the alignment of the wells to square batches of acid-etched pores. For example, a packed array of 10-μ diameter pores is oriented perpendicular to the wafer face. Light can be transmitted through the porous silicon array, thereby increasing resolution and detection sensitivity. GeneLogic licensed Beattie's invention but later abandoned commercialization and formed a spin-off company, Metrigenix, which later introduced the Flow-Thru Chip™ technology under the MGX™ 4D Array brand (Figure 2.11).

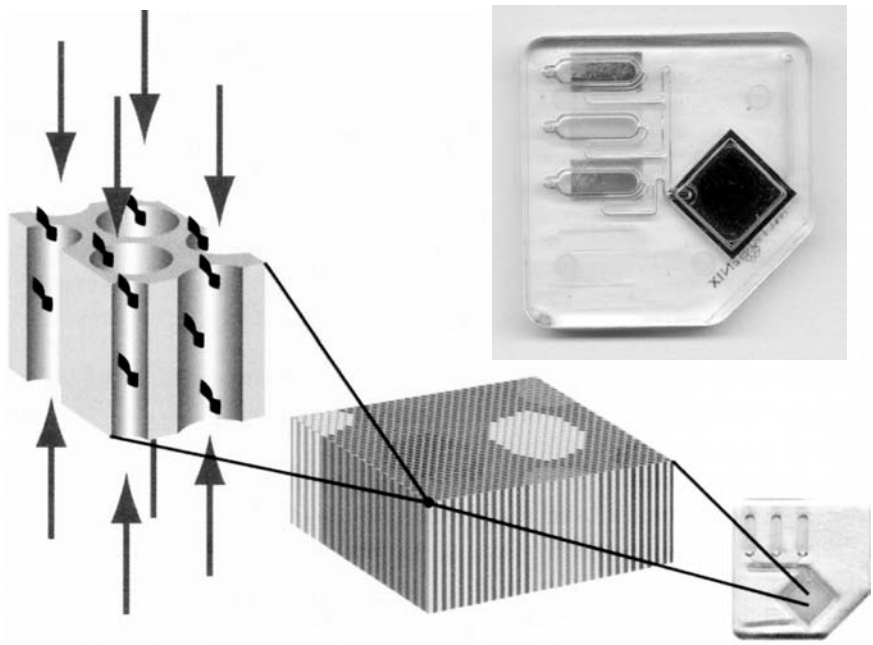


Figure 2.11 Metrigenix's Flow-Thru Chip platform. (From Iyer, M. et al., *IVD Technol.*, July–Aug. 2003, 47–53. With permission.)

PamGene, a spin-off of Organon Teknika (a subsidiary of Akzo Nobel NV, Arnhem, The Netherlands) founded in 1999 has been working with a similar flow-through technology based on porous aluminum oxide. Its product is marketed under the PamChip™ trade name (Figure 2.12). The capillary porous structure (200-nm pore diameter; 10^7 per mm^2) provides a large surface area and small volume flow path. Because of the capillary pore geometry, the printing process leaves sharp, well-defined probe spots as liquid is rapidly transported into the capillaries rather than diffusing across the substrate surface (Chan, 2002).

Electronic biochips

Electronically active chips (e.g., Nanogen's NanoChip® Electronic Microarray) are true microchips in which microelectrodes (pads) become elements of the array (Figure 2.13). The microelectrodes are covered with materials that allow immobilization of probes. Each electrode is individually addressable so that specific probes can be attached to different electrodes. Hybridization is accelerated by electromotive force (emf) on the target. Enhanced stringency is also achieved by modulation of the emf (Heller et al., 2000).

Xanthon (now defunct) planned to introduce a disposable 96-well microplate microelectrode system known as the X²AS (Xanthon Xpression Analysis™ System) for use in high-throughput nucleic acid assays using robotic

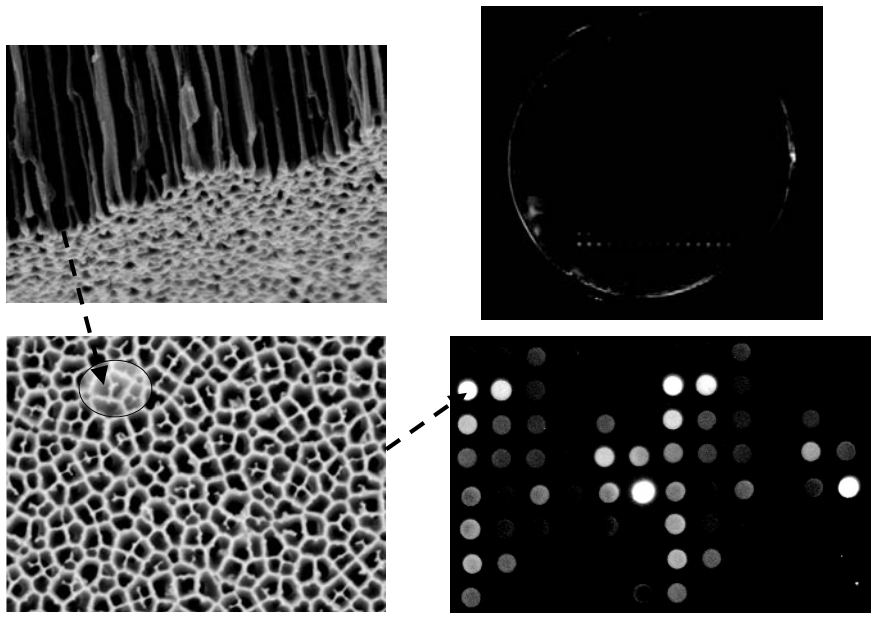


Figure 2.12 PamChip. (Photo courtesy of PamGene International BV, Arnhem, The Netherlands.)

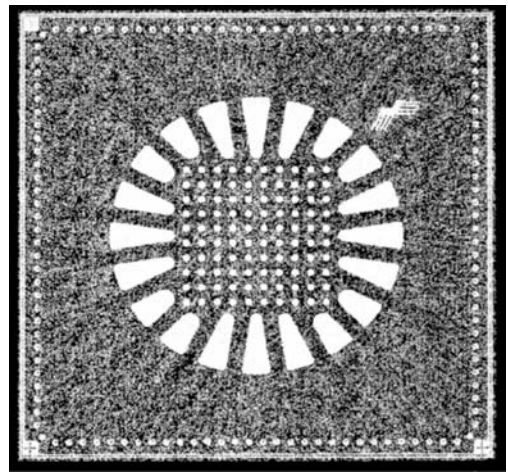


Figure 2.13 NanoChip electronic microarray. (From Heller, M. et al., *Electrophoresis*, 21, 157–164, 2000. With permission.)

workstations. The X^2A plate allowed analysis of 96 samples for 5 specific sequences. The company projected the ability to analyze 27,000 samples per day. The analysis was based upon cyclic voltammetry mediated by ruthenium ion redox in the presence of DNA. Unfortunately, the company failed to achieve mass production and went out of business.

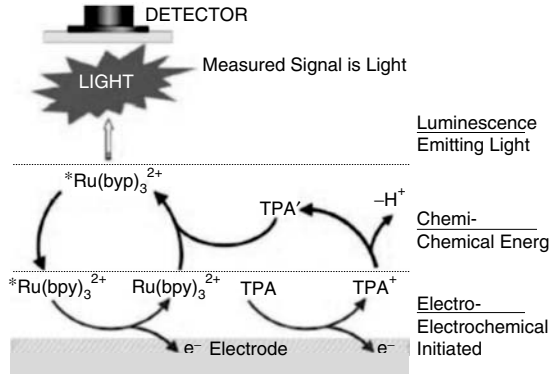


Figure 2.14 Electrochemiluminescence detection. (Courtesy of MesoScale Discovery, Gaithersburg, MD.)

MesoScale Discovery (MSD) succeeded in introducing product with a similar technology approach based upon ruthenium redox-mediated electrochemical detection (Figure 2.14). MSD is a joint venture of its parent company, MesoScale, and IGEN, a company that pioneered much of the work on electrochemical detection based on the ruthenium redox system. MSD's Multi-Spot™ plates contain antibodies immobilized on multiple working electrode pads within each well, allowing each spot within the well to serve as an individual assay. Multiplexed cytokine immunoassays can be performed in 96-well (4, 7, or 10 spots per well) patterns with detection limits of 1 to 10 pg/mL and a linear dynamic range up to 3,000 pg/mL. Both 24- and 384-well electrode systems are available.

Clinical MicroSensors (CMS), now a subsidiary of Motorola Life Sciences, developed the eSensor™ Biochip technology involving DNA capture probes attached to a microelectrode pad through molecular wires of phenylacetylene attached to a gold substrate. The target (unlabeled) is allowed to hybridize under passive conditions. A signaling probe labeled with ferrocene serves as an electron donor that interacts with the gold electrode to produce an electronic signal (Umek et al., 2001). CMS technology may be adapted for point of care (POC) hand-held devices (Figure 2.15) or microtiter plate formats. The major advantage of electronically active or addressable chips is rapid analysis time. Nanogen has the added advantage of emf-driven hybridization and stringency modulation.

Illumina produces *fiberoptic random bead* arrays. Latex beads are encoded using different fluorescent dye mixtures that are either adsorbed into the particles or attached to the surfaces. Presynthesized oligonucleotides are attached to selected bead populations so that a single dye or dye:dye ratio identifies the attached oligonucleotide. Populations are mixed in bulk and then loaded onto the tips of a fiberoptic, one end of which has been acid etched to form microscopic nanowells. The nanowells are filled at random with the mixed bead population to create a BeadArray™. Such bundles can

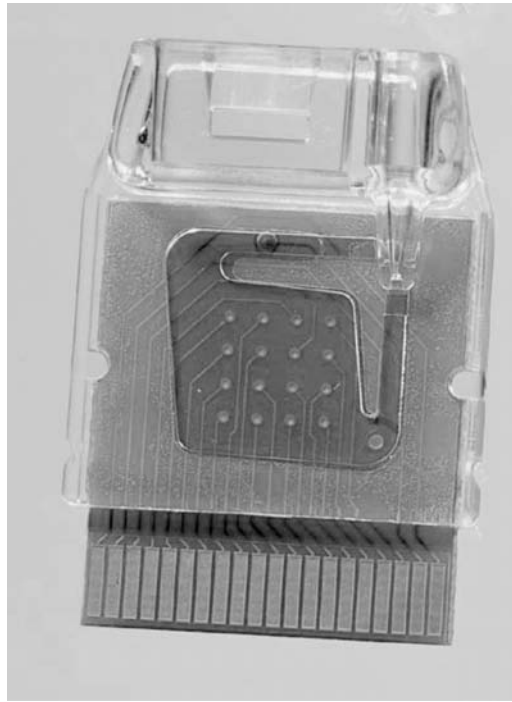


Figure 2.15 eSensor biochip. (Photo courtesy of Motorola Life Sciences, Pasadena, CA.)

be arranged in an Array of Array™ format to match 96-, 384-, or 1536-well plates (Figure 2.16). These products are offered under the trade name of Sentrix™ Array Matrices (Oliphant et al., 2002).

Labeled targets are hybridized and located by imaging a fiber coupled to a light source and charge-coupled device (CCD) camera. After signals and bead positions are mapped, each bead having a positive signal is decoded for the dye ratio and the sequenced determined. A major advantage of this technology is that a high array density can be achieved without printing.

Density is determined by the number of etched wells. Very small volumes of sample can be addressed at high sensitivity. The major application for Illumina's products has been for high-throughput SNP genotyping. With densities up to 50,000 beads per optical fiber, a 96-well system can process 150,000 SNPs in parallel, providing the potential to call more than 1 million genotypes per day.

The Sentrix® BeadChip slide product has recently been introduced for gene expression analysis. Wells are etched in glass slides and in turn filled with 3-μ beads. Each bead contains 50-mer oligonucleotide gene-specific probes. The slide array can screen for 700 genes starting with 50 to 200 ng total RNA. Detection limits to 0.15 p*M* are claimed with fold changes achievable at <1.3, dynamic range at 2.8-fold, and array-to-array imprecision at <10% CV.

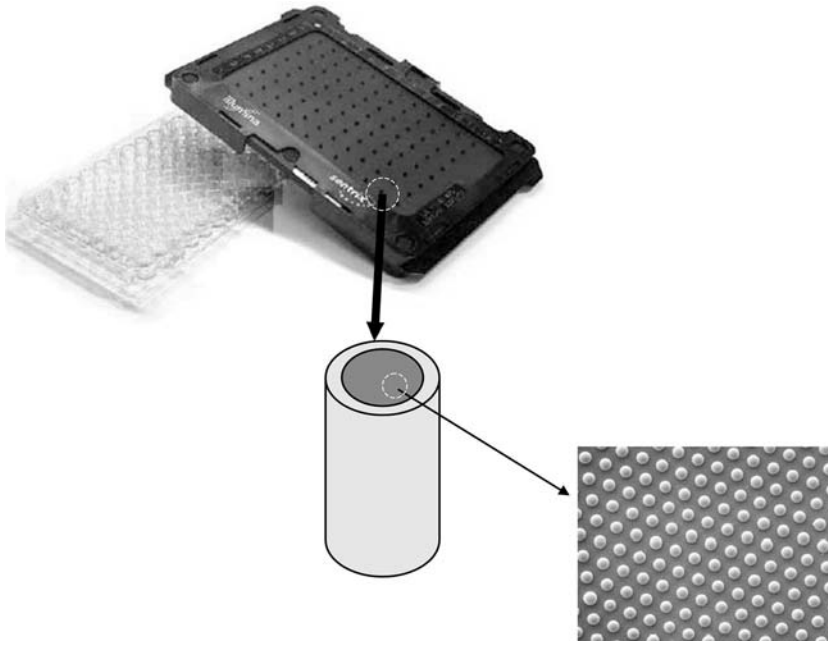


Figure 2.16 Sentrix fiberoptic chip. (From Oliphant, A. et al., *Biotechniques*, 32, S56–S61, 2002. With permission. Photo courtesy of Illumina, Inc., San Diego, CA.)

Future opportunities

DNA microarrays

Microarrays are primarily used for gene expression analysis for pharmacogenomics and drug toxicity testing applications. The drug discovery paradigm is beginning to shift to include microarrays at a number of stages in the process. Early adopters are currently using microscope slides. Automated handling and processing platforms for slide-based arrays are under development for early market entry. The array-of-arrays format in which arrays are constructed within the wells of a microplate are now being introduced for high-throughput, multiplexed assays using robotic liquid handler workstations. A major issue with the use of gene expression microarrays is sensitivity. Sample targets need to be labeled but not mass amplified in order to preserve the differential expression relationships for all genes and especially those in low abundance.

There remains a need to further reduce the amount of sample (mRNA) into the nanogram range. This is achievable through signal amplification based upon improvements to the Eberwine method (Van Gelder et al., 1990), but this is a time-consuming process that has moderate reproducibility. Sample preparation (isolation, purification, and characterization) is also desired for a full solutions approach. Finally, preprinted microarrays with specific

proprietary gene content and associated bioinformatics annotated to disease-specific databases such as those provided by GeneLogic, Genzyme's serial analysis of gene expression (SAGE), and the Phase I tox panel are becoming more available. The "tox" chip holds perhaps the most general utility for preprinted DNA array offerings at present. Moreover, because of the completion of the Human Genome Project, the availability of genomic content has also accelerated commercial product offerings.

Protein microarrays

The proteomics field is currently wide open. Only a few protein microarray products have been commercialized. Unlike DNA microarrays that are commonly constructed on essentially two platforms (GeneChip or slide array) employing oligonucleotides or cDNA probes, the protein microarray appears on a variety of platforms and with different types of probes.

Probes can be antibodies, other binding proteins constructed from protein fusions, or even oligonucleotide aptamers. While completion of the Human Genome Project has enabled access to content for nuclide acid arrays, the content for protein arrays is largely based upon available antibody libraries. Thus, the commercialization of protein microarrays remains largely dependent upon both commercial and institutional providers of protein content. These providers must also permit access to the data-based protein annotations. These are necessary in order for the protein array to be useful as a bioinformatics tool.

Novagen (ProteoPlex™), S&S (FAST Quant), BioSource (Cartesian Array™), and BD Biosciences (BD Clontech™ Ab Microarrays) have introduced or will soon introduce protein microarray slide formatted products in which antibodies are directly immobilized. Beckman Coulter's protein array products for performing micro-ELISAS in standard 96-well plate formats are based upon the self-assembly (by hybridization) of oligonucleotide–antibody conjugates to complementary oligonucleotides arrayed in individual wells. HTG's protein array technology was described previously.

Pierce introduced an array-of-arrays microplate product called SearchLight™ in which antibodies are directly printed into the wells of the microplate. Also, we have reviewed MSD's Multi-Spot plate products having antibodies immobilized onto multiple working electrodes. These products (albeit with some novel approaches to create microarrays and means for detection) utilize the classic immunosorbent sandwich assay but have the advantage of parallel processing using microarrays.

Other available approaches include Ciphergen's ProteinChip™ that acts as a universal platform for protein interactions with mass detection by SELDI (surface enhanced laser desorption/ionization). The device is not really a chip or microarray. Rather, it is a rectangular strip several inches in length (about the length of a standard microtiter plate) containing eight large wells (2 mm in diameter). Various proprietary surface chemistries are available for capturing proteins (e.g., affinity or ion exchange phases). The strip is then fed into the SELDI for analysis (Fung et al., 2001).

Zyomyx chip technology is based upon an atomically flat gold surface to which is attached a proprietary SAM (self-assembled monolayer) surface for optimal protein binding (Peluso et al., 2003). The company launched its Protein Profiling Biochip System and Human Cytokine Biochip products in February 2003.

Another protein expression approach is Phylos's mRNA–protein fusion array, called the PROfile™ chip. The technology allows for the simultaneous monitoring of mRNA and nascent protein expressions (Weng et al., 2002). Thus, high-affinity binding fusion proteins can sequester and select, for example, antibody mimics or be used to determine the expression of related protein families in response to a drug. In 1998, Phylos entered into a collaboration with Hoechst GmbH (now Aventis) to develop the PROfusion™ technology. Aventis provided Phylos with research funding in exchange for certain rights to the technology and specific products resulting from the collaboration. In January 2002, Phylos announced the purchase of the Aventis research unit, thus regaining full rights to the PROfusion technology. To date no product has been commercialized. However, Phylos is also involved in the discovery of binding proteins useful for diagnostic and therapeutic applications. For example, the PROfusion technology led to the development of the Trinectin® binding proteins using fibronectin as a scaffold. Since fibronectin is a naturally occurring protein in humans, it provides an easy route for administration of binding proteins for therapeutics.

While it is true that microarray technology is gearing up for proteomics, it is perhaps still too early to predict what role microarrays will ultimately play. Proteins are much more complex molecules than nucleic acids and the suggestion has been made that additional tools and approaches will be needed. Microfluidic devices (electrophoresis, flow cytometry) with miniaturized detectors may also be applicable.

Tissue and cell microarrays

This technology perhaps represents the most important area from a drug discovery focus. Aurora has pioneered cell-based high throughput screening (HTS); others such as Cellomics and Rosetta have also been actively involved. Rosetta's technology (from the acquisition of Acacia) is yeast genome expression-driven, relying upon cellular events within arrayed yeast colonies as measured by green fluorescent proteins. Cellomics's technology appears to be a more flexible extension of cytology in using differential fluorescent labeling of cellular components.

At the National Human Genome Research Institute, O.P. Kallioniemi (2001) pioneered the use of tissue biopsy microarrays with fluorescence *in situ* hybridization (FISH) analysis and confirmatory cDNA microarrays (Kononen et al., 1998). Stable cell lines or tissues can now be arrayed and stored as "reagents" in a consistent manner, providing researchers with valuable new microarray tools (Kallioniemi, 2001). Commercial sources of tissue microarrays (TMAs) include, among others, Invitrogen's VastArray™

featuring arrays mounted onto standard glass microscope slides containing 600-µm core tissue samples taken from normal human and mouse organs with up to 100 tissues spotted in duplicate. Ambion offers LandMark™ tissue microarrays containing 50 to 200 tissue specimens on a single slide.

Cell microarrays have also been fabricated. Ziauddin and Sabatini (2001) demonstrated the ability to transfect cells cultured onto plasmid DNA arrayed in gelatin on a standard DNA microarray slide. Xu (2002) printed down cells in the form of high density microarrays on permeable membranes and demonstrated phenotypic assay performance with the immobilized cells. The commercialization of viable cell arrays will permit an even closer look at cell-mediated events during the drug discovery process.

References

Anderson, M.L.M. and Young, B.D., Quantitative filter hybridization, in *Nucleic Acid Hybridization: A Practical Approach*, Hames, B.S. and Higgis, S.J., Eds., IRL Press, Oxford, 1985, chap. 4.

Association for Biomolecular Resource Facilities, Microarray Survey 2003.

Barczak, A., Rodrigues, M.W., Hanspers, K., Koth, L.L., Tai, Y.C., Bostad, B.M., Speed, T.P., and Erle, D.J., Spotted long oligonucleotide arrays for human gene expression analysis, *Genome Res*., 13, 1775–1785, 2003.

Baum, M. et al., Validation of a novel, integrated, and flexible microarray benchtop facility for gene expression profiling, *Nucleic Acid Res*., 31(23), 1–13, 2003.

Beattie, K.L., Beattie, W.G., Meng, L., Turner, S.L., Coral-Vazquez, R., Smith, D.D., McIntyre, P.M., and Dao, D.D., Advances in genosensor research, *Clin. Chem*., 41, 700–706, 1995.

Bell, P.A. et al., SNPstream® UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery, *BioTechniques,* 32: S70–S77, 2002.

Bowtell, D.D.L., Options available from start to finish for obtaining expression data by microarray, *Nat. Genetics*, 21 (Suppl.), 25–32, 1999.

Buetow, K.H., Edmondson, M., MacDonald, R., Clifford, R., Yip, P., Kelley, J., Little, D.P., Strausberg, R., Koester, H., Cantor, C.R., and Braun, A., High-throughput development and characterization of a genome-wide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *Proc. Natl. Acad. Sci. USA*, 98, 581–584, 2001.

Chan, A., Coupling different molecular techniques onto a porous microarray for diagnostic applications: Detection of deletions and duplications in Duchenne muscular dystrophy (DMD) using multiplexed amplifiable probe hybridization (MAPH) on PamChip™, First International Conference on Microarrays for Diagnostics, San Diego, CA, March 21–22, 2002.

Dill, K., Montgomery, D.D., Wang, W., and Tsai, J.D., Antigen detection using microelectrode array microchips, *Clin. Chim. Acta*, 444, 66–78, 2001.

Fung, E.T., Thulasiraman, V., Weinberger, S.R., and Dalmasso, E.A., Protein biochips for differential profiling, *Curr. Opinion Biotechnol*., 12, 65–69, 2001.

Garland, P.B. and Serafinowski, P.J., Effects of stray light on the fidelity of photo-directed oligonucleotide array synthesism *Nucleic Acid Res*., 30(19), 1–9, 2002.

Graves, D.J., Powerful tools for genetic analysis come of age, *Trends Biotechnol.*, 17, 127–134, 1999.

Gurevitch, D., Dong, X.F., Pircher, T.J., Matsumoto, S.S., Roycroft, P., Tsinber, P., Falcovitz, Y.H., and Hahn, S., A novel three-dimensional hydrogel-based microarray platform, *JALA*, 6, 87–91, 2001.

Harvey, M.A., Nitrocellulose surfaces for quantitative protein microarrays, Chips-to-Hits Conference, International Business Communications, Boston, October 2003.

Heller, M.J., Forster, A.H., and Tu, E., Active microelectronic chip devices which utilize controlled electrophoretic fields for multiplexed DNA hybridization and other genomic applications, *Electrophoresis*, 21, 157–164, 2000.

Hughes, T.R. et al., Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer, *Nat. Biotechnol.*, 19, 342–347, 2001.

Iyer, M., Philip, R., Matthai, H.E., Eastman, E., and O'Beirne, A.J., Using 4-D diagnostic tools for genetic analysis, *IVD Technology*, July/August 2003, 47–53.

Kallioniemi, O.P., Biochip technologies in cancer research, *Ann. Med.*, 33, 142–147, 2001.

Khrapko, K.R., Lysov, Y.P., Khorlin, A.A., Shick, V.V., Florentiev, V.L., and Mirzabekov, A.D., An oligonucleotide hybridization approach to DNA sequencing, *FEBS Lett.*, 256, 118–122, 1989.

Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M.J., Sauter, G., and Kallioniemi, O.P., Tissue microarrays for high-throughput molecular profiling of tumor specimens, *Nat. Med.*, 4, 844–847, 1998.

Lee, K.N., Shin, D.S., Lee, Y.S., and Kim, Y.S., Micromirror array for protein micro array fabrication, *J. Micromech. & Microeng.*, 13(3), 474–481, 2003.

LeProust, E., Zhang, H., Yu, P., Zhou, X., and Gao, X., Characterization of oligodeoxyribonucleotide synthesis on glass plates, *Nucleic Acid Res.*, 29(10), 2171–2180, 2001.

Martel, R.R., Botros, I.W., Rounseville, M.P., Hinton, J.P., Staples, R.R., Morales, D.A., Farmer, J.B., and Seligmann, B.E., Multiplexed screening assay for mRNA combining nuclease protection with luminescent array detection, *Assay Drug Dev. Technol.*, 1, 61–72, 2002.

Matson, R.S. and Rampal, J.B., DNA arrays: past, present, and future, *Am. Genomic/Proteomic Technol.*, April/May 2003, 37–44.

McCormick, M.,. A high-density platform for the design and synthesis of DNA arrays, *Innov. Pharm. Technol.*, 3, 88–93, 2003.

Mendoza, L.G., McQuary, P., Mongan, A., Gangadharan, R., Brignac, S., and Eggers, M., High-throughput microarray-based enzyme-linked immunosorbent assay (ELISA), *Biotechiques*, 24, 778–788, 1999.

Nuwaysir, E.F. et al., Gene expression analysis using oligonucleotide arrays produced by maskless photolithography, *Genome Res.*, 12, 1749–1755, 2002.

Oliphant, A., Barker, D.L., Stuelpnagel, J.R., and Chee, M.S., BeadArray™ technology: enabling an accurate cost-effective approach to high-throughput genotyping, *Biotechniques*, 32, S56–S61, 2002.

Pease, A.C., Solas, D., Sullivan, E.J., Cronin, M.T., Holmes, C.P., and Fodor, S.P., Light-generated oligonucleotide arrays for rapid DNA sequence analysis, *Proc. Natl. Acad. Sci. USA*, 91, 5022–5026, 1994.

Peluso, P., Wilson, D.S., Do, D., Tran, H., Vankatasubbaiah, M., Quincy, D., Heidecker, B., Poindexter, K., Tolani, N., Phelan, M., Witte, K., Jung, L.S., Wanger, P., and Nock, S., Optimizing antibody immobilization strategies for the construction of protein microarrays, *Anal. Biochem.*, 113–124, 312, 2003.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O., Quantitative monitoring of gene expression patterns with a complementary DNA microarray, *Science*, 270, 467–470, 1995.

Singh-Gasson, S., Green, R.D., Yue, Y., Nelson, C., Blattner, F., Sussman, M.R., and Cerrina, F., Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array, *Nat. Biotechnol.*, 17, 974–978, 1999.

Song, Y., Boyer, D., Leung, I., Moheb, S., Gu, J., Doan, A.T., Khan, A., Flores, I., Yee, F.S., and Wong, J.R., A^2 microarray system: A novel multiplexed assay platform for cytokine profiling, 95th Annual Meeting, American Association for Cancer Research, Orlando, FL, March 27–31, 2004.

The state of the microarray-selected suppliers of microarray chips, spotters, and readers, *The Scientist*, 17(3), 40, 2003.

Table of Suppliers, *Nature*, 416, 894–895, 2002.

Tan, P.K., Downey, T.J., Spitznagel, E.L., Jr., Xu, P., Fu, D., Dimitrov, D.S., Lempicki, R.A., Raaka, B.M., and Cam, M.C., Evaluation of gene expression measurements from commercial microarray platforms, *Nucleic Acid Res.*, 31(19), 5676–5684, 2003.

Tang, K., Fu, D.J., Julien, D., Braun, A., Cantor, C.R., and Koster, H., Chip-based genotyping mass spectroscopy, *Proc. Natl. Acad. Sci. USA*, 96, 10016–10020, 1999.

Umek, R.M., Lin, S.W., Vielmetter, J., Terbrueggen, R.H., Irvine, B., Yu, C.J., Kayyem, J.F., Yowanto, H., Blackburn, G.F., Farkas, D.H., and Chen, Y.P., Electronic detection of nucleic acids: A versatile platform for molecular diagnostics, *Mol. Diagnostics*, 3, 74–84, 2001.

Van Gelder, R.N., Von Zastrow, M.E., Yool, A., Dement, W.C., Barchas, J.D., and Eberwine, J.H., Amplified RNA synthesized from limited quantities of heterogeneous cDNA, *Proc. Natl. Acad. Sci. USA*, 87, 1663–1667, 1990.

Wang, D.G. et al., Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome, *Science*, 280, 1077–1082, 1998.

Weng, S., Gu, K., Hammond, P.W., Lohse, P., Rise, C., Wagner, R.W., Wright, M.C., and Kuimelis, R.G., Generating addressable protein microarrays with PROfusion covalent mRNA-protein fusion technology, *Proteomics*, 2, 48–57, 2002.

Wheeler, D.L. et al., Database resources of the National Center for Biotechnology, *Nucleic Acid Res.*, 31(1), 28–33, 2003.

Wiese, R., Belosludtsev, Y., Powdrill, T., Thompson, P., and Hogan, M., Simultaneous multianalyte ELISA performed on a microarray platform, *Clin. Chem.*, 47, 1451–1457, 2001.

Xu, C.W., High-density cell microarrays for parallel functional determinations, Genome Res., 12, 482–486, 2002.

Ziauddin, J., and Sabatini, D.M., Microarrays of cells expressing defnied cDNAs, *Nature*, 411, 107–110, 2001.

chapter three

Supports and surface chemistries

Introduction

In this chapter, we will survey the kinds of solid supports (substrates) and surface chemistries currently used in the creation of nucleic acid and protein microarrays. Which are the best supports and methods of attachment for nucleic acids or proteins? Does it make sense to use the same attachment chemistry or substrate format for these biomolecules? In order to begin to understand these kinds of questions, it is important to briefly review how such biomolecules were attached in the past to other solid supports such as affinity chromatography media, membranes, and enzyme-linked immunosorbent assay (ELISA) microtiter plates. However, the microarray substrate does not share certain unique properties and metrics with its predecessors. Principal among these are printing, spot morphology, and image analysis; they are the subjects of subsequent chapters.

Substrates

It is interesting to note that while column chromatography and centrifugation were developed for biomolecule purification and separation, many of the early diagnostic substrates for nucleic acids and proteins were membranes. For the Southern transfer process, the membrane provided a convenient way to interrogate sequences in genomic DNA fragments (Southern, 1975).

The advent of the polymerase chain reaction (PCR) made it possible to directly spot down cDNA amplicons onto membranes, giving rise to the Southern dot blot format. In fact, the dot blot should be regarded as one of the earliest, if not the first, array format (albeit a macroarray). Why did many abandon the membrane in favor of the glass substrate for DNA microarrays?

And, how is it that membranes cast upon glass substrates are now used to prepare protein microarrays? We will address these questions in good time.

In the following sections, the major types of substrates currently used for DNA and protein microarrays will be discussed. Much of what is known regarding microarray surface chemistry and the immobilization of biomolecules comes from work with DNA microarrays. Therefore, many of the examples cited here will be from these studies. Zhu and Snyder (2003) in their review provide good insight into the manufacture and utility of protein microarrays. Here are some points to consider when choosing a substrate for protein microarrays:

1. The manufacture and processing of the protein microarray should be conducted in such a manner that the arrayed proteins remain in their native and active state. For most proteins, this usually means the hydrated state in order to avoid surface denaturation. For antibody arrays which are perhaps more forgiving than other proteins, it has been our experience that while these could be stored cold and dry, it is most important to rehydrate them prior to use. This process is in sharp contrast to the preparation of nucleic acid arrays in which strand melting or denaturation is necessary to achieve optimal binding to the solid support. While the hybridization process is well understood and can be controlled under thermodynamic principles, the folding and renaturation of proteins on planar (microarray) surfaces is under study.
2. Hydrogels and other porous 3D matrices that entrap water can offer excellent milieus for maintaining proteins in the hydrated state. The higher surface area may also allow the immobilization of proteins at higher densities than planar or 2D substrates. This in turn can lead to improved sensitivity and dynamic range. On the downside, these polymeric materials can also exclude larger molecular weight biomolecules or slow diffusion and the exchange of buffers and other reagents.
3. Not all substrate materials will be applicable. While affinity chromatography and ion exchange supports have been used quite effectively to purify native proteins, their conversion to a microarray substrate may not work. The printing down of proteins onto these surfaces may be hard to control in terms of spot uniformity and morphology. The exception is nitrocellulose which has been resurrected as a microarray support by casting it down on a glass slide (Schleicher & Schuell Bioscience). Nylon membranes are also prepared in this format (Pall Corporation). Poly-L-lysine (PLL) coated slides used so successfully for creating DNA microarrays were the first to be adopted for protein microarrays (Haab et al., 2001; MacBeath and Schreiber, 2000).

Membrane substrates

Use with nucleic acids

We will begin our discussion with examples from the late 1980s and early 1990s when membranes were first employed in the creation of nucleic acid arrays. It was then that the terminology surrounding DNA analysis began to change. Dattagupta and coworkers first introduced the concept of reverse dot blot hybridization (1989):

> We have developed a simple method of nonisotopically labeling sample nucleic acids, which are then hybridized simultaneously to an array of unlabeled, immobilized probes. This "reversed hybridization" procedure thus provides identification results after a single hybridization reaction.

Instead of spotting down the target DNA (Southern dot blot), unlabeled DNA probes complementary to a target DNA were arrayed onto the membrane. The target DNA (e.g., cDNA) was then labeled and applied to the membrane for hybridization. Reverse hybridization allowed the simultaneous probing of the target against an allele-specific oligonucleotide (ASO) library. One no longer had to strip and reprobe the membrane. The "probe" came to mean the DNA attached to the solid support and the "target," as in the case of the Southern blot, referred to the nucleic acid in the sample to be analyzed.

In this first study, the arrays were constructed by hand spotting of genomic DNA from various bacterial species onto nitrocellulose (NC) membranes. The DNA was denatured with sodium hydroxide and fixed to the NC membranes by baking under vacuum at 80°C. However, the attachment of short oligonucleotides onto NC was not practical.

Nylon membranes were introduced to replace the fragile NC membranes. Single-stranded oligonucleotide probes could then be tethered to the charged nylon membranes by 3′ tailing with poly(dT) followed by ultraviolet (UV) crosslinking (Saiki et al., 1989). In 1991, Zhang and coworkers from Cetus introduced the use of amino linkers in which oligonucleotide probes were terminated at their 5′ ends with amino spacers. The amino-modified probes were then covalently attached to carboxylated nylon membranes (Biodyne C, Pall Corporation, Port Washington, NY) via 1-ethyl-3,3-(3-dimethylaminopropyl)-carbodiimide (EDAC) activation (Zhang et al., 1991). The primary amine was found to be much more reactive than the aromatic secondary amines found in the bases, thus assuring oriented coupling of the majority of probes by the 5′ ends. Probes tethered without the inclusion of spacers were fourfold less efficient in hybridization. Furthermore, the amino linker also provided greater hybridization efficiency over the poly(dT) tailing method.

For longer double-stranded nucleic acids such as cDNA [~200 bp (base pairs) to 1500 bp], the positively charged nylon membrane easily sequestered

the negatively charged strands by adsorption, most probably involving both electrostatic and hydrophobic interactions with the support. The exact mechanisms of baking and UV crosslinking of nucleic acids to nylon or NC are not well understood but presumably involve some covalent interaction of the DNA and the support. Considerable study over the past several decades has been conducted on the use of the nucleic acid dot blot format. The fundamentals are described by Anderson and Young (1985).

Much of the early work relied upon hand spotting or manual application of probes using vacuum filtration devices such as the DotBlot apparatus (BioRad Laboratories) that allowed the formation of more uniform spotting of probes in the form of small dots or rectangular slots. The use of membranes for printed DNA arrays (often referred to as "grid" arrays) was subsequently developed.

For example, Hans Lehrach and coworkers (Nizetic et al., 1991) developed a computer-controlled robotic system for the arraying of bacterial colonies (cosmid library) from 96 wells of a microtiter plate onto nylon membranes. Using this early robotic pin printer, a high density grid of 9216 clones from 96 wells × 96 microtiter plates was constructed on a 22-cm × 22-cm filter. Drmanac and Drmanac (1994) described the high density arraying of PCR samples using a Biomek® 1000 robotic workstation adapted with a 96-pin tool (Bentley et al., 1992). Membrane-based nucleic acid grid arrays continue to be used for genomic analysis (Lane et al., 2001; Hornberg et al., 2002). Grid arrays are also known as macroarrays, a term used to more formally differentiate grid arrays from microarrays in terms of spot size and density.

Use with proteins

Lueking et al. (1999) arrayed recombinant proteins on NC membranes and screened them with different antibodies. Joos and coworkers (2000) printed down autoantigens onto NC membranes and compared performance relative to silylated (aldehyde) and PLL glass slides. Protein arrays could be stored at room temperature for a month without significant loss in activity. Huang (2001) hand spotted down IgG species and antibodies directed toward various cytokines onto membranes. The properties of various commercial membranes were assessed in terms of absorption, background, and sensitivity levels based upon detection by enhanced chemiluminescence (ECL).

Advantages — High binding capacity; multiple hybridization cycles possible.

Disadvantages — Size limitations on immobilization of oligonucleotides; lower spot density due to spot diffusion; large sample and rinse volumes required; blocking agents required to reduce nonspecific binding (NSB); high intrinsic fluorescent background and light scattering issues.

Detection — Radioactivity (film) or phosphorimaging; chemiluminescence; colorimetric reagents. See Figure 3.1 for a comparison of different detection results on membranes. In particular, membranes are well suited for colorimetric detection while glass and plastic are less attractive alternatives.

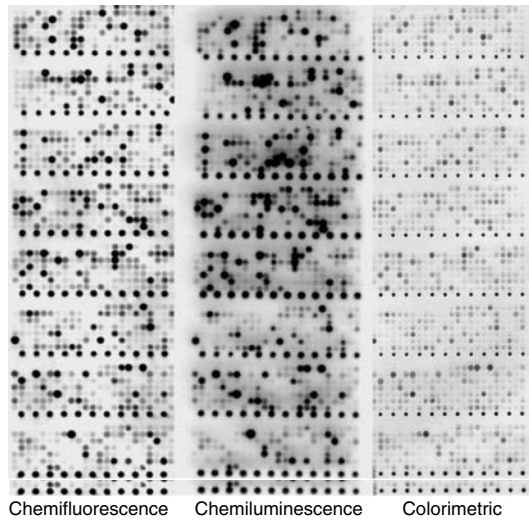


Figure 3.1 Membrane-based signal detection formats. (Courtesy of Andrew Dubitsky, Pall Corporation, Port Washington, NY.)

Glass substrates

Use with nucleic acids

The use of glass as a microarray substrate begins with development of *in situ*-generated oligonucleotide arrays (Southern et al., 1992; Fodor et al., 1991) and later with spotted arrays (Schena et al., 1995). Glass became the substrate of choice for the direct synthesis of oligonucleotides because of its relative inertness for synthesis chemistries; the surface of glass is uniform and impervious, and glasses with good optical properties (flatness, low intrinsic fluorescence) were available. The optical properties of glass later became very important in the development of confocal scanners for reading arrays and achieving fluorescence signals with high sensitivity.

Covalent attachment. Nucleic acids may be attached to glass by noncovalent (electrostatic or hydrophobic interaction) or covalent means. The most common approach has been adsorption to PLL-coated microscope slides or second-generation aminopropyl silane surfaces (APSs). PLL surfaces may be damaged under certain hybridization and stringency washing or stripping conditions, (e.g., high salt levels and elevated temperatures).

For this reason, Zammatteo et al. (2000) examined different covalent coupling conditions for nucleic acids on glass substrates. Silanization of glass with various silane coupling reagents resulted in the grafting of amine (Si–O~NH_2), carboxyl (Si–O~COOH), and aldehyde (Si–O~CHO) surface-reactive groups (Figure 3.2). Likewise, cDNA amplicon (255 bp) probes were synthesized using chemically modified primers with 5′ ends terminating with either phosphate (PO_4–DNA), carboxyl (HOOC–DNA), or amine

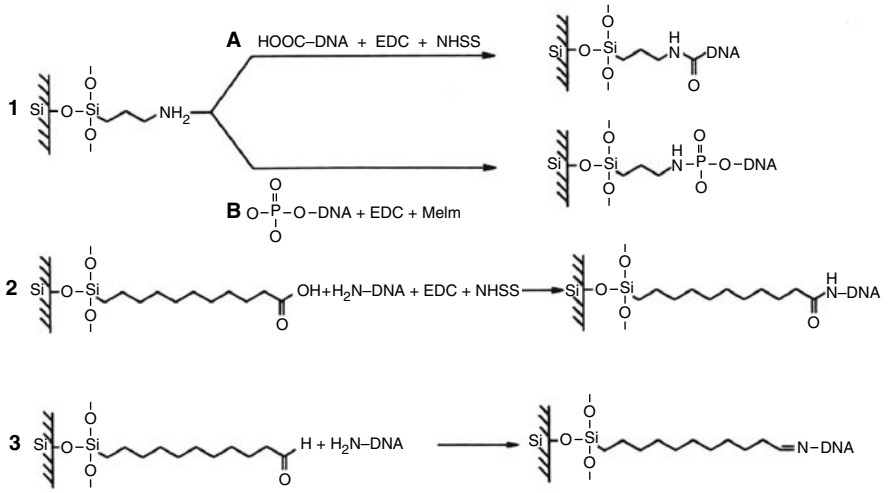


Figure 3.2 Grafting of various functional groups onto silanized glass. (From Zammatteo, N. et al., *Anal. Biochem.*, 280, 143–150, 2000. With permission.)

(H_2N–DNA) groups. Coupling of the modified amplicon probes to the modified glass surfaces was accomplished via carbodiimide-mediated reactions:

$$Si–O\sim NH_2 + HOOC–DNA + EDAC/NHS \rightarrow Si–O\sim NHCO–DNA + H_2O$$

or a Schiff's base reaction:

$$Si–O\sim CHO + H_2N–DNA \rightarrow Si–O\sim C = N–DNA + H_2O$$

The highest efficiency for probe attachment was found for the reaction of carboxylated probe with an aminosilane surface. However, the level obtained by simple adsorption (without the addition of EDAC) was equally as high (mean probe density = ~600 to 700 fmoles/cm^2). Conversely, the attachment of an amine-modified probe to a carboxyl surface was much less efficient (mean probe density = ~300 fmoles/cm^2). The least efficient coupling was obtained using aldehyde surfaces (mean probe density = ~150 fmole/cm^2).

Based upon the loading concentrations of probe DNA (3 μM), the coupling reaction to the aldehyde supports was limiting and required a 60- to 300-fold excess (3000 nM vs. 10 to 50 nM) relative to the other coupling chemistries. Hybridization on the aldehyde surface was higher than on the amine or carboxylated surfaces. While the coupling of phosphorylated DNA was similar to that of the aldehyde surface, it too resulted in lower hybridization efficiency and higher nonspecific adsorption.

Do higher loadings of probe onto the support adversely affect performance or are other factors at play here as well? What about how the DNA probe is tethered to the surface? While there were differences in spacer arm

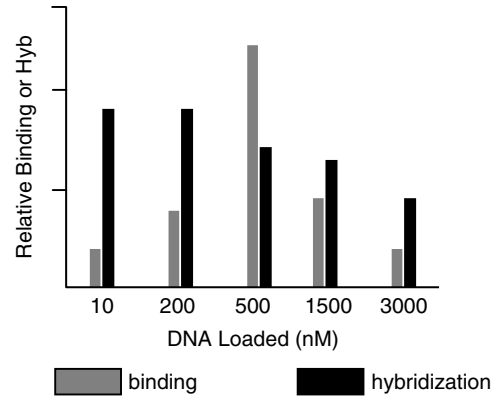


Figure 3.3 Oligonucleotide probe surface loading vs. hybridization efficiency. (From Zammatteo, N. et al., *Anal. Biochem.*, 280, 143–150, 2000. With permission.)

length for the tethering via amine (3 carbon spacer) vs. carboxyl or aldehyde (11 carbon spacer), this most likely had little direct impact because hybridization varied only about twofold for all surfaces, regardless of the surface-reactive groups and probe combinations tested. However, probe density (or probe surface distribution) most surely plays an important role. Zammatteo et al. varied the loading onto the aldehyde surface from 10 nM to 3000 nM and determined that the highest coupling was a 500-nM input probe. The greatest hybridization efficiency was, however, obtained below the optimal loading at about 200 nM, depending upon interpretation of the reported imprecision (signal, mean ± SD); see Figure 3.3.

Following the trend in DNA loading vs. hybridization efficiency resulting from these experiments, one could conclude that higher loadings lead to less efficient hybridization, while minimal loading leads to problems with nonspecific adsorption. One explanation is that the probe distribution may be too tightly packed, thereby preventing hybrid nucleation by steric hindrance of incoming targets. Reducing the loading avoids the issue of steric hindrance. However, the loading should also be titrated to minimize the contribution of nonspecific adsorption (Figure 3.4). Therefore, one should proceed to determine the optimal loading (input probe concentration vs. bound probe) in general, but this does not correlate well with optimal hybridization. It is prudent to carefully quantify that relationship (bound probe vs. hybridization signal strength) and define the limits of nonspecific adsorption.

Adsorptive attachment. PLL surfaces work reasonably well for creating cDNA microarrays, but the suitability of this surface chemistry for immobilization of short oligonucleotides has been questioned. However, as we have learned, covalent attachment chemistries can be problematic as well. In either case, if the oligonucleotide is constrained too close to the surface with multiple points of contact, it may not be able to fully participate in hybridization.

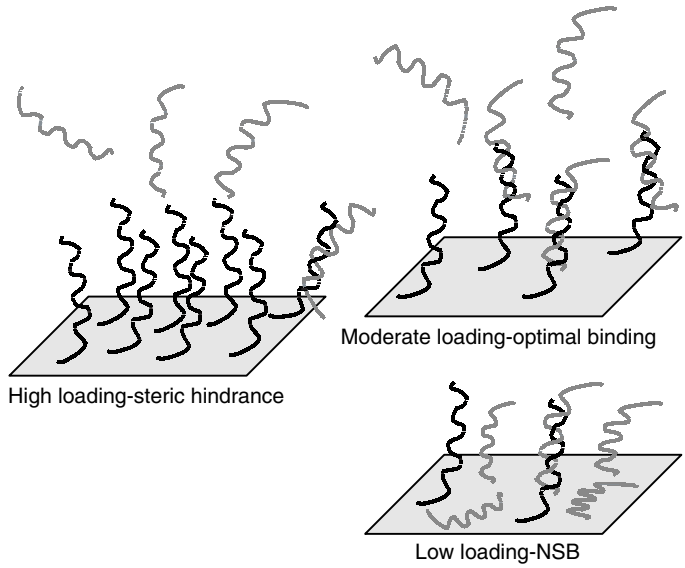


Figure 3.4 Substrate loading effects.

Conversely, too few interactions with the surface may lead to loss during hybridization and washing steps. This is why the covalent attachment of oligonucleotides at their 3′ or 5′ terminus has the advantage, but modified oligonucleotides and slide chemistries can be expensive.

Belosludtsev et al. (2001) described an efficient process for preparing oligonucleotide microarrays based upon the adsorption of unmodified probes to aminosilanized glass slides. This is a two-step process: First, oligonucleotides are presented to the surface dissolved in water so that they remained in a fully denatured state and are then dried down. The second step involves capping of the residual surface amine groups. Recall from the work of Zammatteo et al. (2000) that the nonspecific adsorption of modified oligonucleotides to aminosilane surfaces was regarded as a significant problem.

Capping with succinic anhydride is a common method for reducing backgrounds on amine surfaces (NH_3^+) by creating a negative ($NH–COCH_2CH_2COO^-$) surface charge, thereby repelling nonspecific nucleic acids. However, as these investigators discovered, too high a negative charge density will also repel incoming target DNA. To overcome this issue, a so-called double-capping protocol involving partial capping with acetic anhydride (vapor phase, 50°C, 1 hr, vacuum oven) was followed by succinic anhydride [0.5 M in dimethylformamide (DMF), room temperature, 1 hr] was used. This imparted a slight negative surface charge that was more favorable for hybridization while preventing significant levels of nonspecific adsorption.

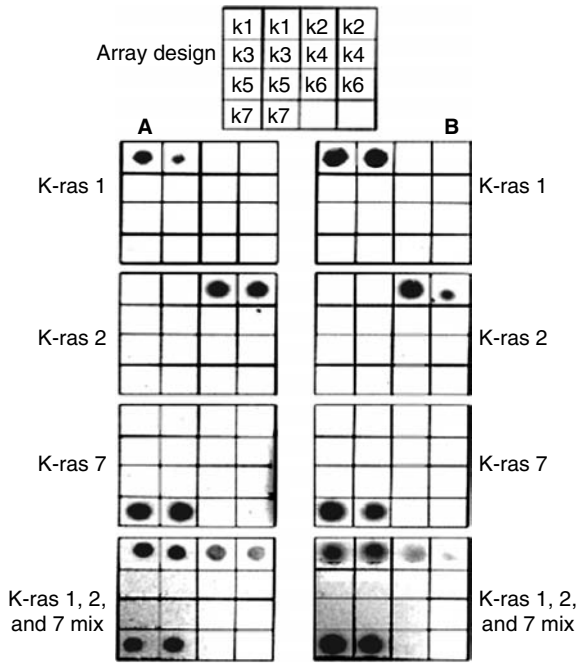


Figure 3.5 Comparison of adsorptive and covalent attachment of capture probes. (From Belosludtsev, Y. et al., *Anal. Biochem.*, 292, 250–256, 2001. With permission.)

The adsorptive process attaching unmodified oligonucleotide onto aminosilane glass was compared to the covalent attachment of amino-oligonucleotides to epoxysilane substrates. A set of 12-mer capture probes for the codon 12-point mutations in K-ras were compared to the hybridization efficiency and specificity of the 152-bp amplicon. Comparable results between adsorption and covalent tethering were obtained (Figure 3.5). Interestingly, the adsorption–chemical capping method required only 20% of the probe loading used for covalent attachment (5 µM unmodified oligonucleotide vs. 25 µM amino-modified oligonucleotide).

Belosludtsev and coworkers (2001) propose that the unmodified probes on the weakly cationic surface, although prevented from diffusing off the surface because of electrostatic interaction, nevertheless are available for hybrid nucleation. They suggest that such probe behavior could be viewed similarly to behavior observed in a liquid crystal matrix. This would be in sharp contrast to models describing the covalent attachment of short oligonucleotides as "oligo lawns" or monolayers of coiled probes (see Figure 3.6).

Call et al. (2001) of the Pacific Northwest National Laboratory in Richland, WA, also studied the immobilization of unmodified oligonucleotides. Amine-modified and unmodified oligonucleotides could be attached to epoxysilane slides (covalent attachment) or acid-washed slides (noncovalent attachment) under the same conditions by printing in an alkaline–sodium

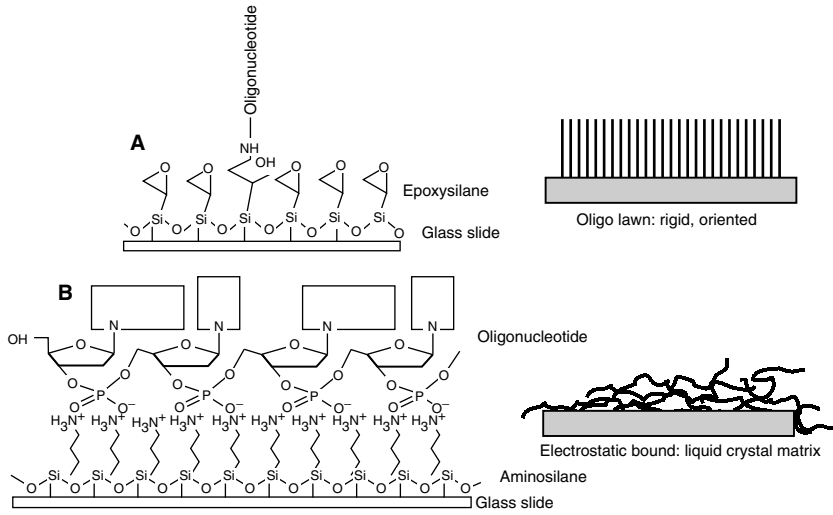


Figure 3.6 Models for oligonucleotide probe behavior on surfaces. (From Belosludtsev, Y. et al., *Anal. Biochem.*, 292, 250–256, 2001. With permission.)

dodecyl surface (SDS) buffer, pH 12, using a microarrayer. The slides were baked in a vacuum oven at 130°C for 30 to 60 min and then stored at 4°C. Optimal hybridization signal was achieved with unmodified oligonucleotides over amino-oligonucleotides on epoxy-activated surfaces using probe loadings of ~100 µM.

When amine-modified oligonucleotides attached to epoxy or hydroxyl slides were subjected to strong alkaline pH conditions, hybridization efficiency rapidly deteriorated on the acid-washed (hydroxyl) slides but not on the epoxy slides (Figure 3.7). PNA (peptide nucleic acid) oligonucleotides that lack the backbone (PO_4^-) charges of standard oligonucleotides exhibited no loss in hybridization on either surface treated with alkaline. These results led the investigators to suggest that hydrogen bonding also plays an important role in nucleic acid attachment to glass surfaces. The effect of baking may further enhance binding by dehydration of the probes on the surface. The authors suggest that, "regardless of which attachment mechanisms are involved, the probes are probably oriented in flat, 'piled' conformation."

Use with proteins

Most protein microarray work has been done using glass slides. Harvard's MacBeath and Schreiber (2000) clearly demonstrated the advantage of printing proteins onto aldehyde slides and the use of bovine serum albumin (BSA) as a scaffold for presenting proteins on the substrate. This was followed by an impressive study from Pat Brown's laboratory at Stanford University investigating the characteristics of 115 antibody and antigen pairs on PLL microarrays (Haab et al., 2001). Wiese et al. (2001) printed down monoclonal antibodies on

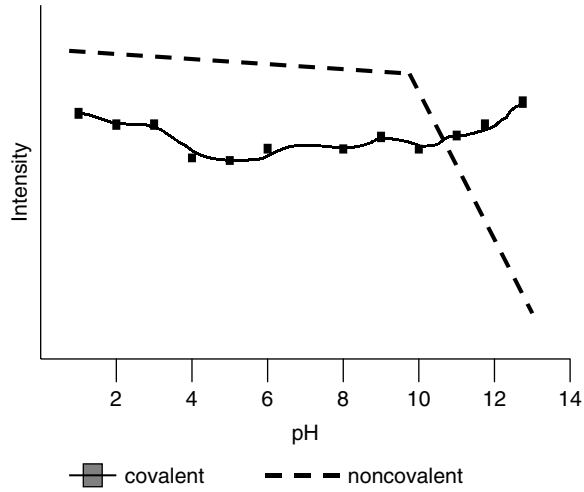


Figure 3.7 Removal of noncovalent-bound probes under alkaline pH conditions. (From Call, D.R. et al., *Biotechniques*, 30, 368–379, 2001. With permission.)

a 96-well Teflon-masked silanized (aminosilane) glass plate and performed micro-ELISAS for PSA and IL-6.

Delehanty and Ligler (2002) sought to preserve biotinylated antibodies immobilized onto avidin-coated slides. For printing of the capture antibodies, a deposition buffer comprising 10 mM phosphate, 10 mM NaCl, 10 mM sucrose, and 0.1% BSA (w/v) was employed. Antibody was dispensed (piezo type; BioChip Arrayer I, Packard Biosciences) at 10 μg/mL. BSA was added as a carrier protein (1 mg/mL) in an effort to reduce the loss of antibody activity by surface denaturation and mass losses due to nonspecific adsorption onto the printing parts and source plates. Sucrose was added to maintain antibody hydration. Glycerol was successfully used to protect proteins from dehydration as well.

Seong (2002) compared silylated (aldehyde) and silanated (amine and epoxy) compounds from several commercial sources to the performance of an antigen (IgG) microarray. In addition, the efficiency of phosphate-buffered saline (PBS) (pH 7.4) and carbonate (pH 9.6) printing buffers were compared. While the various slides and surface chemistries showed differences in their binding isotherms, they ultimately reached similar levels of saturation. Silylated (aldehyde) slides showed comparable loading in both buffer systems. Apparently, tethering of antibody to the surface by Schiff's base formation of the surface aldehyde and lysine residues on the protein was applicable over a broad pH. However, carbonate buffer increased binding of proteins on silanated surfaces.

The interaction on silanated slides is thought to be the result of a combination of electrostatic and hydrogen bonding. Aminosilane surfaces are positively charged at neutral pH. Printing under alkaline conditions shifts

the net charge on the protein, making it more negative and thereby increasing the likelihood of electrostatic interaction with the surface.

Angenendt et al. (2003) reasoned essentially that since surface charge on proteins is variable, it would be unlikely that a single surface chemistry would prove to be universal for the arraying of all proteins while preserving their native active states. The group examined eight different coatings and compared their performances for suitability as antibody microarrays or as protein microarrays. According to the authors, "one surface for both antibody and protein microarray applications could not be found."

Advantages — Optical transparency and flatness; low intrinsic fluorescent backgrounds; light scattering; ultraflatness; impervious surface allows high density array construction.

Disadvantages — Lower probe binding capacity; blocking agents required to reduce NSB; need for expensive scanner for increased sensitivity.

Detection — Primarily via fluorescence-based confocal scanner.

Plastic substrates

Use with nucleic acids

For higher throughput applications, injection-molded plastic microtiter plates have served as the formats of choice for automated assay development. Thermoplastics such as polystyrene, polycarbonate, and polypropylene are used for a variety of purposes including storage and assay plates, lids, pipette tips, and Eppendorf PCR tubes. Polystyrene plates are used for cell culture and ELISAs. Polycarbonate reagent bottles are popular, while polypropylene storage plates and PCR tubes are standards.

The covalent immobilization of DNA into microwells has been described. Rasmussen et al. (1991) used aminated polystyrene plates (CovaLink NH System, Nunc Brand Products) in which a secondary amino group tethered via a spacer arm was grafted to polystyrene. Phosphorylated oligonucleotides or plasmid DNAs were attached to the polystyrene by carbodiimide-mediated coupling of the nucleic acid's 5′-terminal phosphate group to the secondary amine. Hamaguchi and coworkers from the Hitachi Chemical Research Center (Irvine, CA) explored the use of plastic microplates for mRNA capture using immobilized oligo(dT) that allowed them to perform reverse transcriptase polymerase chain reaction (RT-PCR) directly in the wells (1998).

Rigid and transparent polystyrene plates (GenePlate®, Hitachi Chemical) were used initially and found not to be heat stable under PCR thermocycling conditions, especially at the required 94°C denaturation step. For that reason, polypropylene microplates were substituted (GenePlate-PP, initially based upon Nunc's GeNunc PP plate) and became the basis for the GenePlate product offering by RNAture (Irvine, CA) a subsidiary of Hitachi Chemical.

The amount of surface oligonucleotide in each well was approximately 20 pmole and could capture almost 50% of the sample RNA, e.g., well-bound mRNA ~2.0 ± 0.4 ng from 5 ng globin mRNA (40.0%); 27.3 ± 4.0 ng per 50 ng

(54.7%); and 216.0 ± 26.2 ng per 500 ng (43.2%). From 0.5 µg liver, total RNA (5.83 ng mRNA) well capture was 2.8 ± 0.3 ng (48.0%) while at higher sample inputs (5 to 50 µg), recoveries were substantially reduced.

Such thermoplastics have also been used as DNA microarray substrates (Matson et al., 1995; Shchepinov, 1997; Beier and Hoheisel, 1999) and in the construction of protein microarrays in microwells (Matson et al., 2001; Moody, 2001). Pierce (see searchlight@perbio.com) introduced the SearchLight series of microarray-based ELISA assays immobilizing capture antibodies in a low density array format into polystyrene microwells.

Beckman Coulter (www.beckmancoulter.com) now offers the A^2 MicroArray System based upon a 6 × 7 array of capture oligonucleotides printed in the bottom of a 96-well polypropylene plate (A^2 Plate™, Beckman Coulter). Complementary oligonucleotides conjugated to the user's antibody provide a convenient method for the creation of custom microarrays. The plate is Society for Biomolecular Screening (SBS)-compliant for automation using liquid-handling robotics. Greiner Bio-One (Frickenhausen, Germany) also introduced the HTA™ (high-throughput microarraying) plate and HTA slides that print microarrays onto a polystyrene surface (see www.gbo.com/bioscience).

Liu and Rauch (2003) of Motorola investigated oligonucleotide probe attachment onto polystyrene (PS), polycarbonate (PC), polymethyl methacrylate (PMMA), and polypropylene (PP) plastic surfaces. They utilized three different immobilization processes: SurModics' surface modification solution (that allows attachment of adsorbed reactive groups to a surface by photoactivation of polymers at 254 nm), Pierce Reactive-Bind coating solution, and CTAB (cetyltrimethylammonium bromide, a cationic detergent). Not surprisingly, the microarray performances on these plastics varied.

A number of different properties of plastics were studied relative to the utility of these surfaces in the preparation of DNA microarrays and subsequent detection. Intrinsic fluorescence backgrounds were measured at 532 nm (Cy3 excitation) and 635 nm (Cy5 excitation) relative to glass. In terms of the Cy3:Cy5 ratio, PC was 20-fold higher in background fluorescence, as were PS (5.5-fold), PMMA (7.7-fold), and PP (9.3-fold) over the intrinsic background of glass. PMMA had the lowest relative backgrounds (Cy3 = 6.9-fold and Cy5 = 0.9-fold), while PC was highest overall (Cy3 = 107.2-fold and Cy5 = 5.3-fold).

Plastics also varied in their relative hydrophobicity (measured as wetting contact angle), and this property can exert a profound effect on the ability to spot down oligonucleotides of uniform spot diameters and morphologies. The suitabilities of the various surface modifying agents are also dependent upon good wetting in order to uniformly cover the surfaces with the reactive groups.

Reacti-Bind failed to provide a useful surface for printing oligonucleotide microarrays. Reacti-Bind actually increased the contact angle on these plastics, and this may be the reason for its poorer performance. The spotting behavior (spot size and morphology) of 5′-amino-oligonucleotide probes on the various modified surfaces reflected differences in the surface wetting properties among

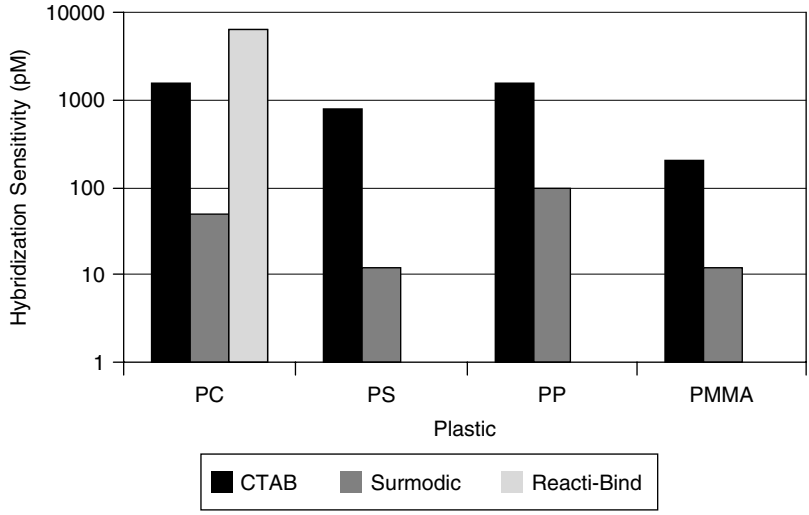


Figure 3.8 Hybridization efficiency on various plastics with reactive surfaces. (From Liu, Y. et al., *Anal. Biochem.*, 317, 76–84, 2003. With permission.)

the plastics. Increased wetting (lower contact angle) resulted in a spreading of the droplet and an increase in spot diameter (Figure 3.8).

The SurModics surface that provided the greatest degree of wetting and spot uniformity also produced microarrays with the highest hybridization efficiencies in terms of the limit of detection (LOD) in sensitivity. Of the plastics studied, PMMA proved to be the best substrate for the immobilization of the 21-mer 5'amino-oligonucleotide probe used in these studies based upon CTAB and the SurModics processes. However, all substrates performed very well with the SurModics process with LODs ranging from 12 to 100 pM. CTAB worked best below its CMC (critical micelle concentration) but was more difficult to control under spotting conditions due to evaporative water loss that could shift the concentration at or near the CMC. This is an important point to consider in utilizing detergents for printing.

Use with proteins

In 1991, Roger Ekins convinced Boehringer-Mannheim to pursue commercial development of his Microspot® technology (Ekins, 1998). Antibody microarrays were constructed on single-well polystyrene carriers by ink-jet printing. Unfortunately, no product was commercialized.

Around the same time, Beckman Instruments (now Beckman Coulter) had begun an array-based product development program focused on the use of modified plastics. Silzel and coworkers (1998) and Matson et al. (2001) of Beckman Coulter were among the first to pursue printing of antibodies onto a plastic surface in a microarray format. Silzel et al. immobilized biotinylated monoclonal antibodies onto an avidin-coated polystyrene surface and performed micro-ELISA-based isotyping of IgG species. Matson et al.

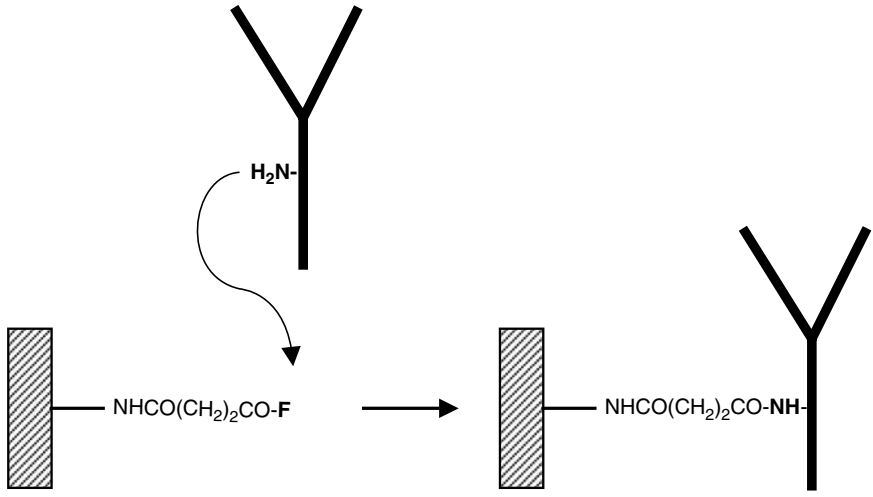


Figure 3.9 Acyl fluoride coupling chemistry.

(2001) printed down monoclonal antibody microarrays in the bottom of an acyl fluoride surface-activated microwell plate and demonstrated a multiplexed micro-ELISA for cytokines (Figure 3.9).

Other early work includes that of Moody et al. (2001) who spotted anti-cytokine monoclonals onto the bottom of polystyrene microtiter plates (Maxisorp, Nalge Nunc) and measured cytokine levels in stimulated peripheral blood mononuclear cells. Finally, although not strictly a microarray, the microwell array system developed by Michael Snyder's group at Yale University to measure kinase activity is a simple and elegant approach (Zhu et al., 2000). The "protein chip" is comprised of microwells fabricated in a flexible elastomer of PDMS [poly(dimethylsiloxane)] substrate by a molding process.

The PDMS microwell array is mounted onto a glass slide and activated using an epoxysilane to which the protein may be attached. The wells were about 1.4 mm in diameter and 300 μm deep allowing a volume of approximately 300 nL. The microwell protein chip should be widely applicable.

Advantages — Inexpensive material; moldable into a variety of shapes; chemical resistance especially toward salts, acids, and bases.

Disadvantages — Flatness and optical clarity issues; largely unsuitable for confocal scanners.

Detection — Primarily via fluorescence-based charge-coupled device (CCD) camera detectors.

Physical features

Hydrogels

Arguments have been made that planar, so-called 2D surfaces are not the optimal physical structures for immobilizing nucleic acids or proteins. The

proposition is that because of its lower surface capacity, a 2D surface has a limitation in providing both high sensitivity and dynamic range. On the other hand, glass slide microarrays appear to work well for gene expression analysis and confocal scanners seem to offer enough sensitivity. Nevertheless, it is very likely that the arraying of proteins in a hydrogel matrix would offer some advantage in preserving the protein's native conformation in the hydrated state. Much of the pioneering work in this area has come from the laboratory of the late Andrei Mirzabekov at the Engelhardt Institute of Molecular Biology, Moscow (Khrapko, 1989; 1991). Later, in collaboration with the U.S. Argonne National Laboratory (ANL) his groups perfected the fabrication processes for hydrogel-based microarrays or MAGIChips™ (microarrays of gel-immobilized compounds on chips) for both nucleic acids and proteins (Guschin et al., 1997; Vasiliskov et al., 1999).

Essentially, various polyacrylamide gels were produced based upon the common polyacrylamide gel electrophoresis (PAGE) casting systems (e.g., acrylamide; bis-acrylamide; sodium persulfate; tetramethylethylenediamine (TEMED)). Following polymerization, the cast gel was exposed to aqueous hydrazine hydrazide to create surface hydrazide reactive groups for attachment of aldehyde terminated oligonucleotide probes. The aldo-oligonucleotide was prepared from oligo containing a 3′ terminal 3-methyluridine. The terminal ribose was oxidized to dialdehyde in the presence of sodium periodate. Reaction of the aldo-oligonucleotide with gel hydrazide resulted in a covalent bond immobilizing the oligonucleotide to the gel (Figure 3.10).

The process was later refined for photo-polymerization of both nucleic acids and proteins. The issues of slowed diffusion and exclusion of targets because of limited porosity were addressed by substituting diallyltartardiamide as the crosslinker in place of bis-acrylamide (Guschin et al., 1997). Finally,

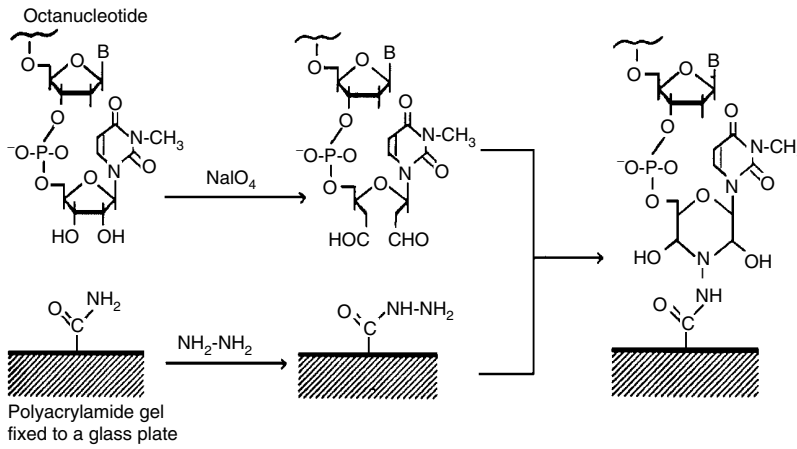


Figure 3.10 Immobilization of aldehyde oligonucleotides in hydrazide gel. (From Khrapko, K.R. et al., *J. DNA Sequencing Mapping*, 1, 375–388, 1991 [www.tandf.co.uk/journals]. With permission.)

photo-induced copolymerization of biomolecules (oligonucleotides and proteins) with gel components that allowed crosslinking of different probes in different gel pads was introduced. Copolymerization was accomplished by the use of a physical mask and movable diaphragm in front of the light source.

Photo-polymerization was conducted at 254 nm under a modified fluorescent microscope. Allyl-oligonucleotides were employed with methylene blue as the free radical initiator for crosslinking. In the case of proteins, acryloyl streptavidin was first immobilized so that biotinylated proteins could be applied to the gel pad (Vasiliskov et al., 1999). One of the major drawbacks with the gel pad approach was that separate pads had to be manufactured instead of coating the entire slide with the gel and then printing down the microarray.

Motorola Life Sciences and Packard Biosciences (now Perkin Elmer Life Sciences) established a development partnership with ANL to commercialize the technology in 1998 but later abandoned the technology in favor of the SurModics hydrogel introduced in 1999 (3D-Link™). Motorola introduced the CodeLink microarray product based upon the SurModics PhotoLink chemistry. Amersham Biosciences acquired Motorola's biochip business in 2002 and now offers CodeLink microarrays. Perkin Elmer sells a similar product under the trade name of HydroGel 3D.

Biocept introduced the 3D HydroArray. The array is produced by mixing the oligonucleotide (or protein) into a polyethylene glycol-based prepolymer solution. The solution is then arrayed onto a glass slide (aminosilane, GAPS2™, Corning) and the droplets cured to form a hydrogel. Biocept estimates 10^{10} to 10^{11} probes per 300-μ diameter spot.

Surface chemistries

Linkers

While linkers (spacers) were added to oligonucleotides to increase the number of available interactions possible with a membrane surface in order to assure attachment, the use of linkers with glass substrates was largely for the opposite reason. That is, the tethering of oligonucleotides by spacers was done to reduce interactions and provide access of the incoming target. This is because of the excessive number of surface silanol groups available to form hydrogen bonds with the oligonucleotide, thereby tightly sequestering the biomolecule to the surface, resulting in steric interference with probe–target hybrid formation.

Lloyd Smith and coworkers (Guo et al., 1994) derivatized aminosilane glass slides with 1,4-phenylene diisothiocyanate in order to covalently attach 5′ amino modified probe oligonucleotides. Direct attachment of the probe to the support through this linkage failed to support hybridization of the PCR targets, while the inclusion of poly(dT) spacer elements between the coupling agent and the oligonucleotide probe resulted in efficient capture of the incoming target (Figure 3.11). The minimal spacer length appeared to be 6 nt and hybridization signal was found to increase linearly up to 15 nt (i.e., dT_{15}); see Figure 3.12.

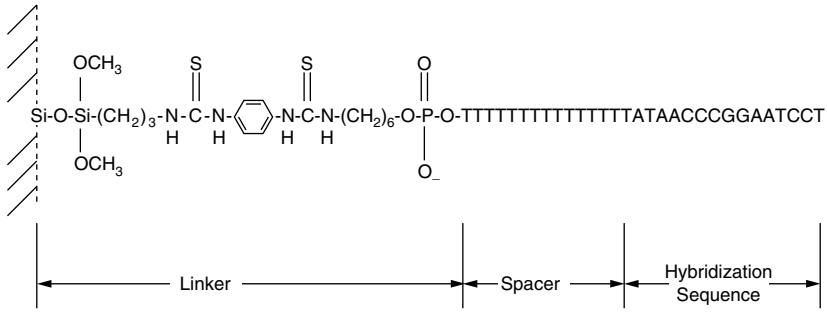


Figure 3.11 Tethering oligonucleotides with a poly(dT) spacer arm. (From Guo, Z. et al., *Nucleic Acid Res.*, 22(24), 5456–5465, 1994. With permission.)

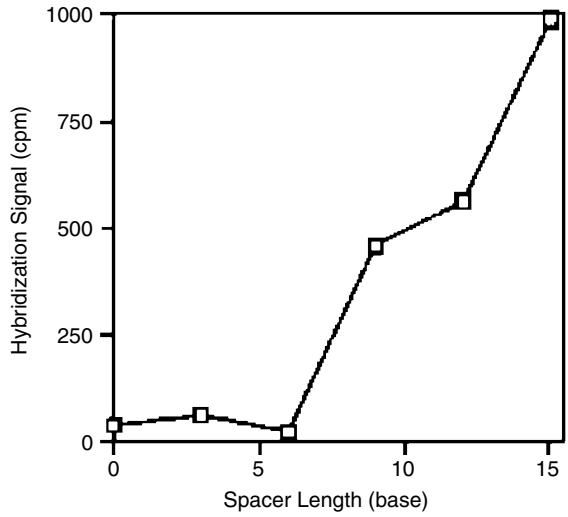


Figure 3.12 Poly(dT) spacer arm effect on hybridization efficiency. . (From Guo, Z. et al., *Nucleic Acid Res.*, 22(24), 5456–5465, 1994. With permission.)

Joos et al. (1997) tethered a 35-mer oligonucleotide to an aminosilane glass coverslip with or without a 15-nt spacer (15-mer hetero-oligonucleotide sequence) and compared hybridization efficiency for the labeled complementary 35-mer target nucleic acid. (Figure 3.13). Under conditions of excess solid phase capture, oligonucleotide hybridization efficiencies ranged from 23 to 64% (no spacer) and from 26 to 74% (with spacer); see Table 3.1. While these results were inconclusive, the authors believed that the spacer offered some advantage.

In an earlier study from Ken Beattie's group (then at the Houston Advanced Research Center), a similar conclusion was reached using a 9-mer oligonucleotide and triethylene glycol phosphoryl repeat unit as the spacer (1995). An optimal probe density leading to the greatest hybridization

Table 3.1 Efficiency of Hybridization to Attached Oligonucleotides

Number of target molecules[a]		Number of probe molecules applied in hybridization[c]		
Applied	Attached[b]	10^{10}	10^{11}	10^{12}
10^{13}	2.7×10^{12}	3.8×10^{9} (38%)	3.7×10^{10} (37%)	6.4×10^{11} (64%)
10^{12}	3.9×10^{11}	3.2×10^{9} (32%)	3.9×10^{10} (39%)	2.0×10^{11} (51%)*
10^{111}	4.1×10^{10}	2.3×10^{9} (23%)	1.3×10^{10} (32%)*	3.7×10^{10} (90%)*
10^{10}	4.1×10^{9}	2.2×10^{8} (5%)	8.9×10^{8} (22%)*	2.2×10^{9} (54%)*

[a] Succinylated target oligonucleotide attached to 12-mm diameter aminopropyl-modified glass coverslips; target = 5′-succinyl-seqβ-seqδ.

[b] Based upon coupling at pH 3.6..

[c] Hybridization with 32p-labeled 70-mer probe = seqβ′-seqδ′ (see Figure 3.13c).

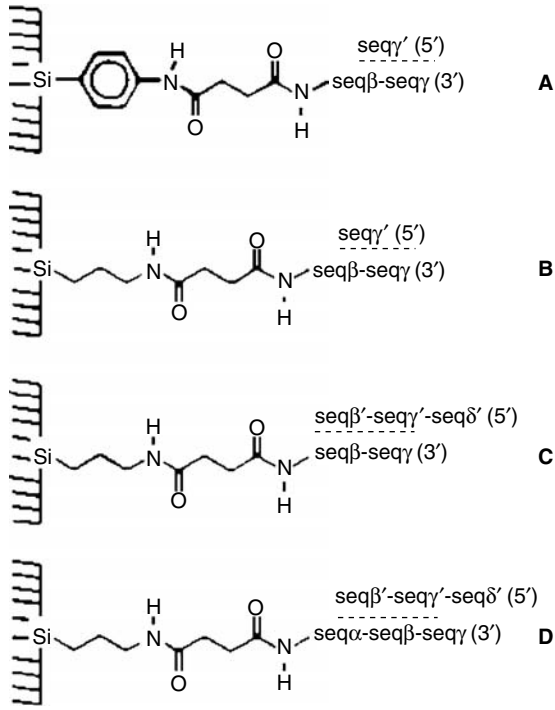


Figure 3.13 Coupling of a 35-mer probe with and without a 15-mer hetero-oligonucleotide spacer. (From Joos, B. et al., *Anal. Biochem.*, 247, 96–101, 1997. With permission.)

efficiency was achieved at a probe loading of about 5 µM. Higher loadings were found to reduce hybridization efficiency to substantially background nonspecific adsorption levels (Figure 3.14).

Poly (dT) units were examined for their effect when distributed between the glycol spacers. PCR products (rendered ssDNA by removal of the biotinylated 5′-primer strand by capture to a streptavidin spin column) ranging

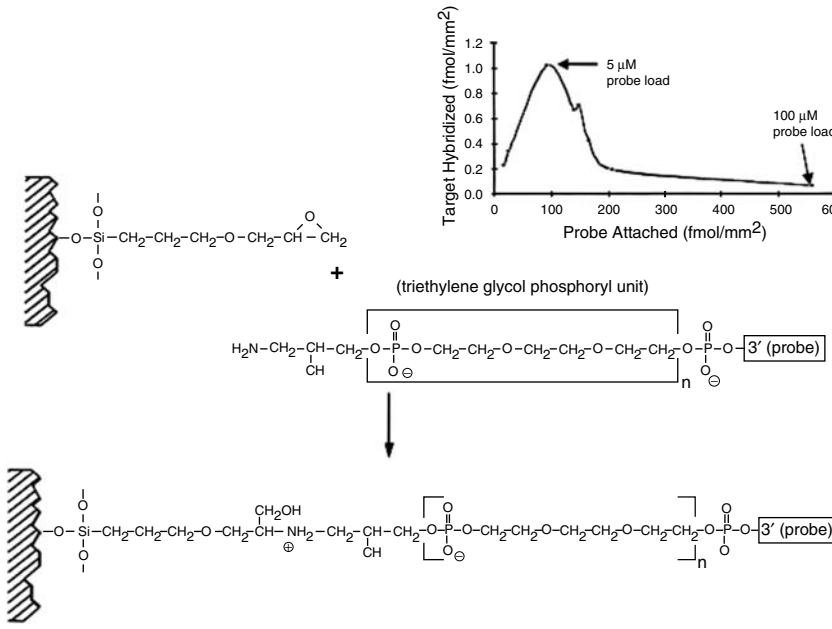


Figure 3.14 Probe density vs. hybridization efficiency. (From Beattie, W.G. et al., *Mol. Biotechnol.*, 4, 213–225, 1995. Copyright, 1995 Humana Press, Inc. With permission.)

between 142 and 1300 nt were hybridized to the 9-mer probe. Although the image quality was poor, including spacer arms for the efficient hybridization of the longer target sequences (624 nt and 1300 nt) appeared to produce some benefit. Nevertheless, even the "directly" attached 9-mer probe tethered to the support via 7 to 10 atom rotatable bonds showed some hybridization to these targets.

The difficulty in assessing the influence of spacers on hybridization efficiency is that factors other than length may be equally as important. Relative hydrophobicity, electrostatic charge effects (spacer arm and surface charge), and wetting of the particular solid support used in the studies may contribute to the overall hybridization event. Keep in mind that the nature of target molecule (e.g., oligonucleotide, cDNA, RNA) and the concentrations of both the tethered probe and incoming target are also important considerations.

Ed Southern's group at Oxford University undertook a systematic study of the influence of spacer molecules (Shchepinov et al., 1997). Instead of aminosilane glass, they chose to use aminated polypropylene, a largely hydrophobic support material (Matson et al., 1995). Hydrophilic, zwitterionic, and uncharged spacer arms were used to tether the probes (Figure 3.15). When the spacer units comprised phospho-propanediol, di-ethylene glycols, or tri-ethylene glycols [i.e., $(OCH_2CH_2)_n$-OPO_2^-], the researchers noted a 50- to 150-fold enhancement in hybridization relative to the oligonucleotide directly coupled to the support.

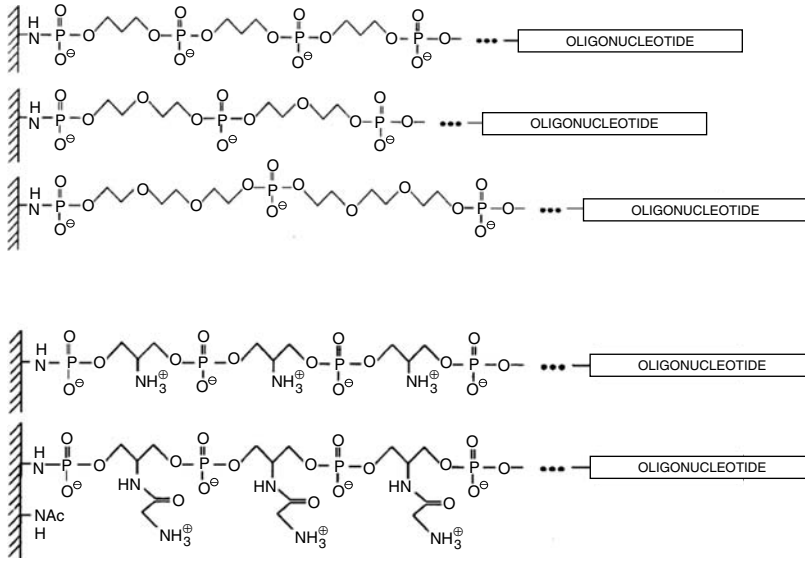


Figure 3.15 Oligonucleotide probe attachment with charged and uncharged spacer arms. (From Shchepinov, M.S. et al., *Nucleic Acid Res.*, 25(6), 1155-1161, 1997. With permission.)

Optimal spacer unit length for the mono-, di-, and tri-glycols was reported to be in the range of 8 to 10 unit repeats, corresponding to 45 to 90 atom lengths, depending upon the repeat unit utilized. Beyond n = 10 units, hybridization decreased to that of directly tethered probes at n = 30 spacer units (Figure 3.16). Longer spacers may in fact fold back onto the surface. Certainly, charge variation among these spacers also produced some effect. The introduction of amphiphilic spacers of both negative (~OPO_3^-) and positive (~NH_3^+) into the spacer arm backbone resulted in less dramatic increases in hybridization efficiency. An optimal spacer unit of n = 3 to 4 was observed while additional n = 5 to 7 units led to a decrease.

Another approach largely borrowed from the extensive work done on chromatographic supports is the use of polymeric coatings. They offer some advantage in masking out the support surface properties while increasing the functional group capacity; and most importantly effectively removing the probe far away from the surface and freely exposing it to the bulk solution.

Beier and Hoheisel (1999) created a dendrimeric linker based upon stepwise condensation of di-amines with acryloyl (chloride) groups generated from a reaction with a surface of either aminosilane glass or aminated polypropylene. This stepwise process resulted in a glass (or polypropylene) surface modified with a mixture of dendrimers with a multiplicity of terminal amine groups (Figure 3.17). The application of homobifunctional crosslinking agents such as disuccinimidylcarbonate (DSC) transformed these end groups into various isothiocyanates, N-hydroxysuccinimide

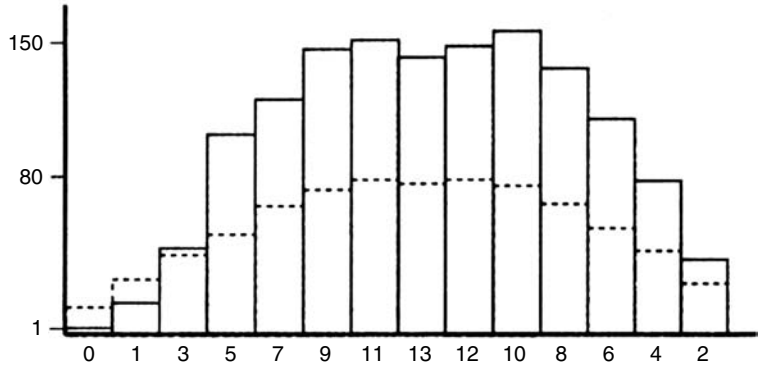


Figure 3.16 Impact of spacer arm length on hybridization efficiency. (From Shchepinov, M.S. et al., Case-Green, S.C. et al., *Nucleic Acid Res.*, 25(6), 1155-1161, 1997. With permission.)

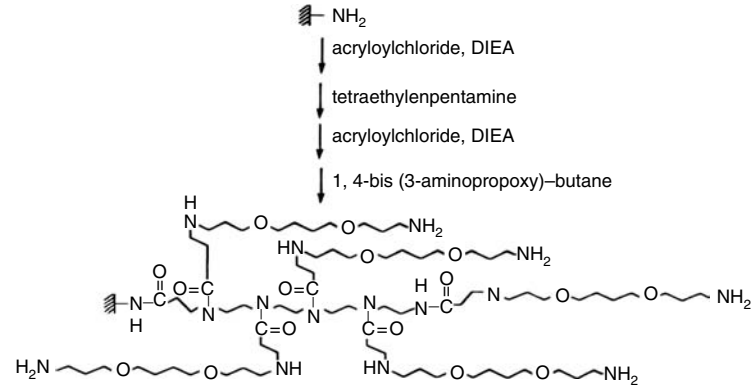


Figure 3.17 Creation of dendrimer spacers. (From Beier, M. and Hoheisel, J.D., *Nucleic Acid Res.*, 27(9), 1970–1977, 1999. With permission.)

(NHS) esters or imidoesters that were in turn reactive toward amino-oligonucleotides.

One particular dendrimer system was reported to have increased the loading capacity 10-fold over direct coupling of the oligonucleotide to the surface. Unfortunately, the presumed improvement in hybridization efficiency was not discussed in quantitative terms. Reusability of the array was a primary objective of the work. Again, quantitative information was not provided. As the authors note, the array withstood >7 actual hybridization stripping cycles and 30 simulated (lacking target) cycles.

Benters et al. (2002) of the University of Bremen in Germany utilized starburst dendrimers to create polymeric surface coatings for glass substrates in an effort to increase sensitivity for DNA microarrays (Figure 3.18). Aminosilane or epoxysilane chemistries were applied to glass substrates. The

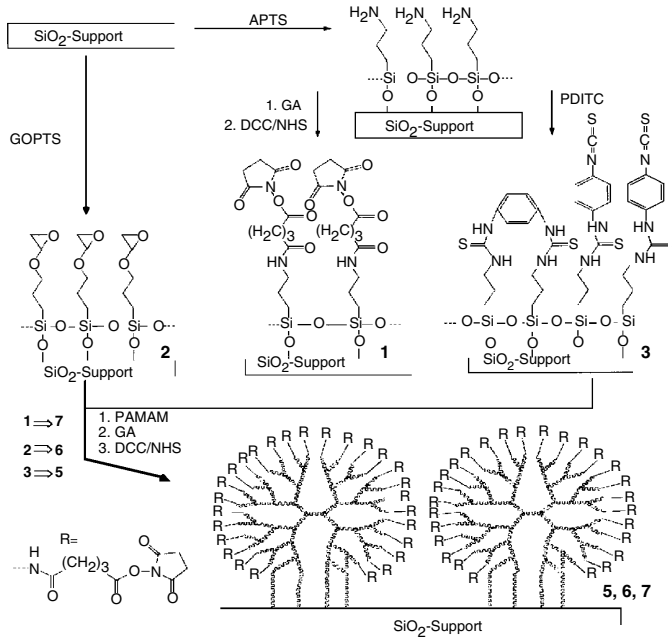


Figure 3.18 Polymeric dendrimer coatings on glass. (From Benters, R. et al., *Nucleic Acid Res.*, 30(2), 1–7, 2002. With permission.)

aminosilanes were further derivatized with terminal NHS or isothiocyanate groups for coupling of the polyamidoamine (PAMAM) starburst dendrimer (Aldrich Chemical, Milwaukee, WI) containing 64 primary amines. The dendrimer could be covalently immobilized to the glass surface by reaction with the three different chemistries: epoxide, NHS, and isothiocyanate linkage. Subsequently, the remaining PAMAM amines could be converted to terminal NHS reactive groups for coupling of 5′-amino oligonucleotides (18 to 24 mer). All PAMAM surfaces were significantly more efficient in hybridization compared to PLL, epoxysilane, or NHS and isothiocyanate-modified aminosilane slides. Up to 10 hybridization–regeneration cycles were accomplished using the optimal starburst surface.

Researchers at Incyte Genomics in collaboration with Samuel Sawan at the University of Massachusetts examined the utility of cross-linked polymer systems for improving oligonucleotide loading capacity in an effort to improve hybridization performance (Lee et al., 2002). Starting with APS slides, the amino groups were activated using cyanuric chloride (CC slides). The APS and CC slides were considered to be 2D surface chemistries. Reactions with polyethylenimine (PEI, branched polymer, average molecular weight ~25,000) produced PEI slides. A final round with cyanuric chloride to produce cyanuric chloride activated PEI (P) slides provided multiple sites for attachment of amino-oligonucleotide probes (Figure 3.19). The PEI and P slides were denoted as 3D surface chemistries.

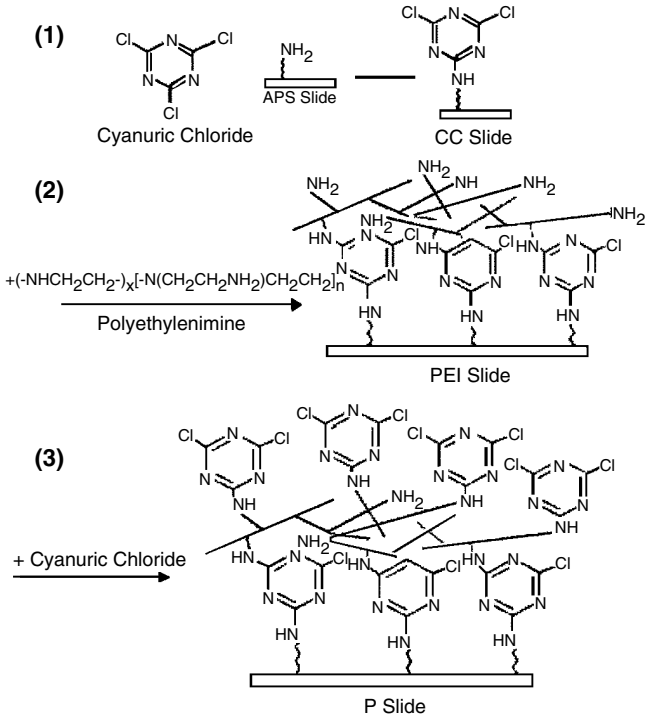


Figure 3.19 Process for preparation of activated polymeric surfaces. (From Lee, P.H. et al., *Bioconjugate Chem.*, 13, 97–103, 2002. With permission.)

All surfaces were characterized in terms of dynamic contact angles, Fourier transform infrared spectroscopy (FTIR) spectra, surface binding uniformity (based upon signal histogram for amino-reactive fluorescent dye), probe oligonucleotide binding capacity, and hybridization. The amount of amino-oligonucleotide probe covalently attached to each surface was directly proportional to the loading concentrations applied. In contrast, hybridization efficiency decreased at the higher loadings. Optimal loading for hybridization appeared to be at 10 μM oligonucleotide. The 3D surfaces performed quantitatively better than the 2D surfaces in terms of hybridization, while both types of activated surface features (CC vs. P) demonstrated similar binding capacities. All surfaces (APS, CC, PEI, and P) bound more oligonucleotide and produced higher hybridization signals than PLL slides (Figure 3.20). Considering the mean signal and variance (SD), the relative rankings for probe immobilization are P ≥ PEI > CC >> APS >> PLL. For hybridization, they are P > PEI > CC > APS > PLL.

Taylor et al. (2003) at Virginia Commonwealth University examined three commonly used glass slide surface chemistries (poly-L-lysine, epoxysilane, and aminopropylsilane) and a dendrimer structure (DAB) similar to that described by Benters et al. (2002); see Figure 3.21. Slides derivatized

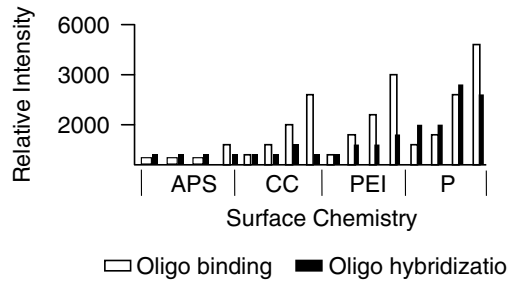
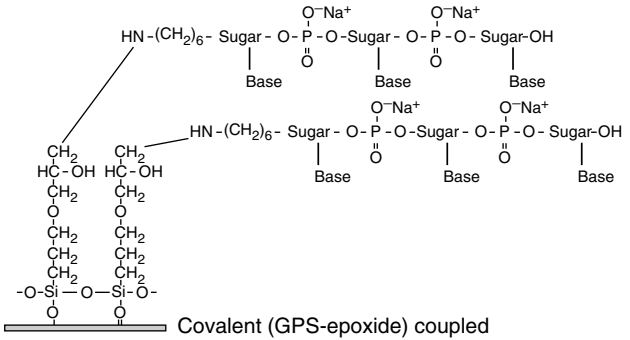
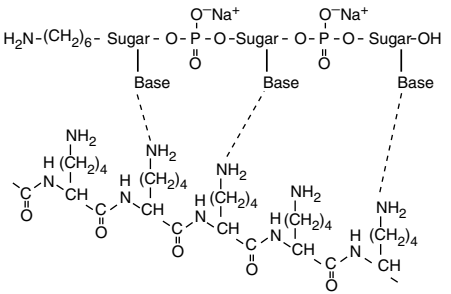


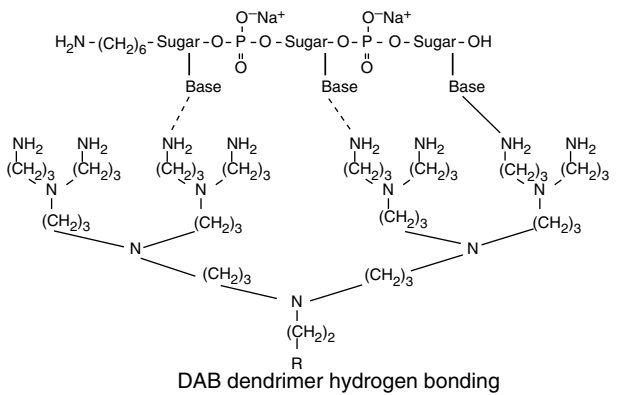
Figure 3.20 Oligonucleotide loading vs. hybridization on activated surfaces. (From Lee, P.H. et al., *Bioconjugate Chem.*, 13, 97–103, 2002. With permission.)

with separate surface chemistries were arrayed with the same cDNA (~600 bp, glyceraldehyde-3-phosphate dehydrogenase [GAPDH] amplicon) and 30-mer GAPDH oligonucleotide probes organized into subgrids for cDNA and probe, respectively (Figure 3.22). Each subgrid contained spots varying in DNA loading concentration in the range of 0.1 to 500 ng/μL (cDNA) and 0.1 to 1000 ng/μL (probe) along with random and nonhomologous control sequences. An extensive series of studies were conducted on these surfaces to measure hybridization performance (Alexa 555-labeled GAPDH amplicon) under different blocking conditions (unblocked vs. BSA vs. succinic anhydride). Spot signal intensity, spot quality, and background fluorescence were evaluated. The following observations were made:

1. cDNA probes were more efficient at hybridization (produced higher spot signal intensity) than oligonucleotide probes, but with greater signal variability. Neither the signal intensity nor its variability could be linked to a particular surface chemistry at a given loading.
2. Oligonucleotide probe signal coefficient of variations (CVs) were in general lower than those of the cDNA probes; however, this was somewhat dependent upon the surface chemistry and blocking condition.
3. Epoxysilane surfaces provided the most efficient hybridization (S:B or signal-to-background ratio) for cDNA or oligonucleotide (Figure 3.23): (a) optimal loading, cDNA, ~0.1 to 0.5 μg/μL; (b) optimal loading, 30-mer oligonucleotide, ~0.01 to 0.1 μg/μL.
4. DAB dendrimer surfaces showed no improvement in hybridization efficiency or spot variability relative to the other surface chemistries evaluated. This is in contrast to the findings of Benters et al. (2002) and Beier and Hoheisel (1999) on dendrimers where such scaffolds worked well.
5. BSA-blocked slides exhibited reduced backgrounds over those blocked with succinic anhydride. However, succinic anhydride-blocked epoxy slides appeared to be less variable. Microarrays created on unblocked epoxysilane glass exhibited the greatest hybridization S:B (signal to background noise) efficiency.







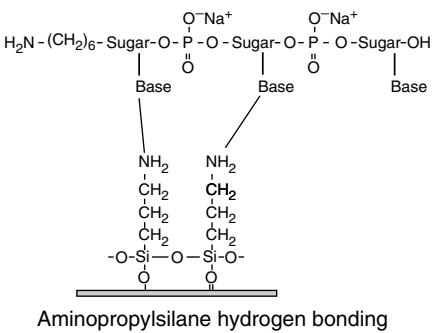


Figure 3.21 Abbreviated surface chemistry structures: PLL, epoxysilane, APS, and DAB. (From Taylor, S. et al., *Nucleic Acid Res.*, 31(16), 1–19, 2003. With permission.)

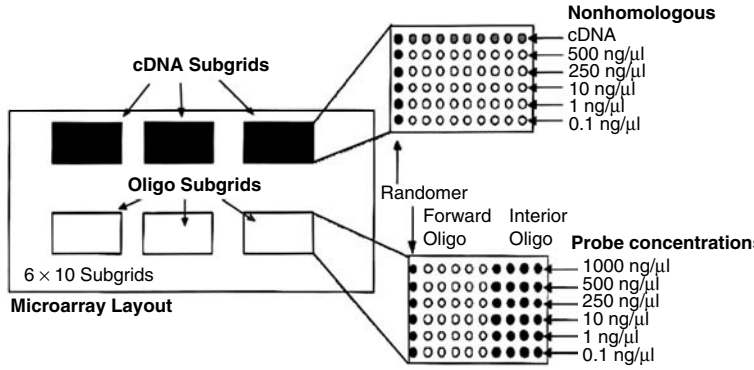


Figure 3.22 Array map for GAPDH DNA and 30-mer probe. (From Taylor, S. et al., *Nucleic Acid Res.*, 31(16), 1–19, 2003. With permission.)

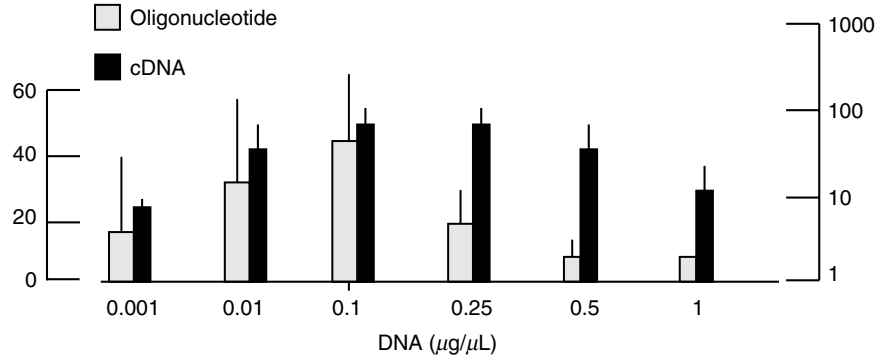


Figure 3.23 Hybridization signal to background (S:B) levels for cDNA vs. oligonucleotide probe. (From Taylor, S. et al., *Nucleic Acid Res.*, 31(16), 1–19, 2003. With permission.)

Generally, the inclusion of a spacer arm between the solid phase and the oligonucleotide probe can result in an improvement in hybridization efficiency. The problem arises in selecting a spacer unit that obviously must be carefully matched to the surface properties and the properties of the probe. Linker (spacer arm) selection is more likely than not an empirical process.

Reactive groups

We can denote and diagram covalent coupling reactions between an activated substrate (S*) and a reactant (R) where R represents the capture probe (oligonucleotide or antibody) by functional groups. Figure 3.24 shows the steps in preparing a glass slide for the attachment of an antibody. The substrate (S) is reacted with aminosilane to form an activated substrate (S*). Oxidized antibody (R) containing reactive aldehyde groups (R-CHO) can be covalently immobilized to the aminosilane slide by a Schiff's base reaction:

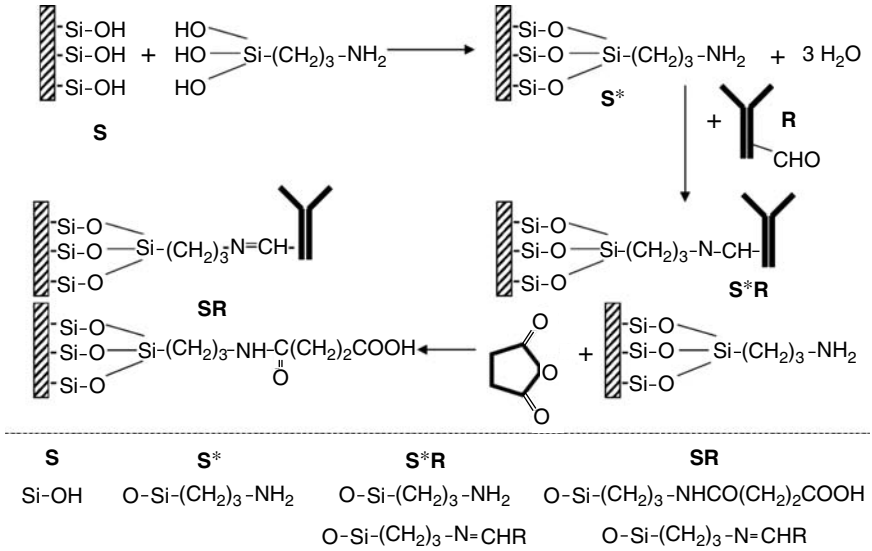


Figure 3.24 Steps in preparation of APS surface, coupling of oxidized antibody, and blocking of residual reactive surface groups.

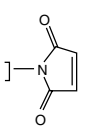
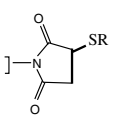
$$S^* - NH_2 + O = CH - R \rightarrow S - NH = CHR + H_2O$$

However, most coupling chemistries do not go to completion so that the substrate will contain a mixture of functional groups capped with attached probe (SR) while others remain free (S*). These residual reactive functional groups must be capped or blocked in some manner to reduce nonspecific binding to the microarray. Residual surface amines may be capped by reaction with succinic anhydride. This renders the support neutral (SR). Using this abbreviated nomenclature, we can describe common surface modifications for microarray substrates.

Most immobilization chemistries for microarrays currently rely upon derivatization of the substrate with amine-reactive functional groups such as aldehydes, epoxides, or NHS esters. While we can choose from many available surface-reactive chemistries, it is important to keep in mind that they must be compatible with a printing process. Ideally, the biomolecule should react completely and rapidly with the substrate in order to achieve good spot formation. It is also critical that the probe remain or be recoverable in its active state following printing. If too reactive a chemistry is employed there is the possibility for excessive crosslinking that can hinder performance by reducing the number of rotatable bonds in the probe.

The best substrate will present the probe to the solution phase with as much rotational freedom as possible so that it can undergo favorable binding with the incoming target molecule. The binding should approximate free solution association. Table 3.2 lists common coupling chemistries employed for probe (nucleic acid and protein) attachments useful for microarrays.

Table 3.2 Common Coupling Chemistries Employed for Nucleic Acid and Protein Probe Attachments

S	S*	R	S–R	Conditions
]–COOH + NHS	]–COO–NHS	RNH_2	]–COO–NHR	pH 6.5–8.5
]–NH_2	Poly-L-lysine	RCHO	]–N=CHR	pH 8–9
]–Si–OH + APS	]–NH_2		]–$BGCG_2R$	Hydrogen reduction
]Si–OH + GPTS	]–CH—CH	RNH	]–CH(OH)CH–NHR	pH 9–10
Zhu (2000)	\ O /	RSH	]–CH(OH)CH–SR	pH 7.5–8.5
		ROH	]–CH(OH)CH–OR	pH 11–12
]–$CONH_2$ + NH_2NH	]–$CONHNH_2$	RCHO	]–CONH=CHR	pH 5–6
Khrapko et al. (1991)	]–CHO	RNH2	]–CH=NR	pH 8–9
]–Si–OH Silane-SH/GMBS Delehanty and Ligler (2002)		RSH		pH 6.5–7.5

Note:] = solid-phase.

Many of these have been discussed in the previous sections. Additional information about specific coupling chemistries for proteins can be found in the protocols and discussion regarding solid phase reagents (Matson, 2000).

Preparation of glass substrates for derivatization

Glass must be thoroughly cleaned to be useful as a microarray support. The cleaning process must remove surface contaminants such as oils, greases, and particulates. Another important reason for cleaning slides is to reactivate the surfaces so that plenty of surface hydroxyl (SiOH) is available for subsequent chemical derivatization. The following protocols have been used to clean glass slides. They include descriptions of the functionalization processes for creating amine, carboxyl, epoxide, aldehyde, NHS, and PEI surface chemistries.

Beattie et al. (1995): Attachment of oligonucleotides by epoxide

1. Soak slide in 1 N nitric acid for 30 to 60 min.
2. Rinse in water; dry.
3. Sonicate for 10 min in each of the following solvents: (a) hexane; (b) acetone; (c) absolute ethanol.
4. Dry slides.
5. Soak slides in 24:8:1 v/v epoxysilane solution comprising anhydrous xylene, glycidoxypropyl trimethoxysilane, and N,N-diisopropyl ethylamine at 80°C for 5 hr.
6. Wash slides three times in tetrahydrofuran (THF).
7. Dry and store slides desiccated under vacuum.

Beier and Hoheisel (1999): Attachment of dendrimer linkers from amine

1. Soak slides overnight in 10% sodium hydroxide.
2. Rinse sequentially with (a) water; (b) 1% hydrochloric acid; (c) water; and (d) methanol.
3. Sonicate 15 min in 3% aminopropyl trimethoxysilane–95% methanol.
4. Rinse in 100% methanol.
5. Rinse with water.
6. Dry slides under nitrogen stream.
7. Bake 15 min at 110°C.
8. The acylation and amination reactions to prepare the dendrimer start with the amine slides prepared according to the above procedure.

Zammatteo et al. (2000): Preparation of carboxylic acid and aldehyde slides

1. Soak slides in 30% hydrogen peroxide:18 M sulfuric acid (33:66 v/v) for 30 min.
2. Rinse with distilled water.

3. Soak 10 min in boiling water.
4. Dry under argon stream.
5. Immerse slides in 1 mM 2′,2′,2′-trifluoroethyl-11-(trichlorosilyl) undecanoate (TETU) in toluene for 1 hr.
6. Sonicate 10 min, three times, in fresh toluene soaks.
7. Dry under argon stream.
8. Preparation of functional carboxylic acid and aldehyde groups: see below.

Carboxylic acid

1. Soak TETU slides in 8 M hydrochloric acid for 2 hr at 95°C.
2. Sonicate three times in distilled water.
3. Dry under argon stream.

Aldehyde

1. Soak TETU slides for 2 hr in 20 mM $LiAlH_4$ in anhydrous ether.
2. Soak for 1 hr in 10% hydrochloric acid; rinse two times in water; rinse in acetone, and dry at 120°C for 10 min.
3. Soak "hydroxyl" slides for 2 hr in 100 mM pyridinium chlorochromate in anhydrous dichloromethane; rinse three times (10 min per time) in dichloromethane; soak in water for 5 min; dry under argon; store aldehyde slides under vacuum.

Belosludtsev et al. (2001): Vacuum amine and epoxy silanization protocols

1. Clean slides in ultrasonic bath with detergent for 2 min.
2. Rinse three times in water.
3. Rinse twice in methanol.
4. Dry for 30 min. at 40°C.
5. Place slides in rack into vacuum oven.
6. Place in a vacuum oven a Petri dish containing 3 mL silane + 3 mL xylene; 3-aminopropyltrimethoxysilane (amine) or -glycidoxypropyltrimethoxysilane (epoxide).
7. Close oven door and apply vacuum.
8. Maintain overnight under vacuum (25 in. Hg) at 70 to 80°C.

Benters et al. (2002): Preparation of carboxyl- and NHS-activated surfaces

1. Clean slides by ultrasonic treatment in chloroform.
2. Soak in fresh "piranha" solution (concentrated sulfuric acid:hydrogen peroxide, 2:1 v/v).

3. Soak slides with stirring in 3-aminopropyltrimethoxysilane (APS) solution (APS:water:ethanol, 2:3:95 v/v) for 2 hr.
4. Soak amine slides in saturated solution of glutaric acid in DMF overnight; rinse thoroughly with DMF. Amine slides have now been converted to carboxyl slides.
5. Soak carboxyl slides in a solution containing 1 M each NHS and DCC (N,N′-dicyclohexylcarbodiimide) in DMF solvent for 1 hr; rinse with DMF. The carboxyl slides have now been converted to the activated NHS slides.

This protocol was described for immediate use. Otherwise, I recommend that the NHS slides be dried under an argon stream, sealed and stored at 4°C until needed for printing. Always allow slides to warm to room temperature before opening to avoid moisture condensation that can inactivate the surface chemistry.

Lee et al. (2002): Preparation of PEI-coated slides

1. Clean slides in a ultrasonic bath with 1% SDS and 4% hydrofluoric acid (HF).
2. Dry slides in an oven at 110°C.
3. Immerse in APS, 0.05% w/v in 95% ethanol; rinse several times with 95% ethanol; oven dry at 110°C.
4. Immerse amine slides in a stirred slurry of cyanuric chloride (12.7 g/L) and sodium carbonate (25g/L) in hexane at 4°C for 1 hr with sonication; rinse with hexane; air dry.
5. Immerse slides in 0.1% aqueous polyethylenimine (25,000 average mol. wt.) at 4°C for 1 hr with sonication.
6. Sonicate in deionized water and allow to air dry.

Summary

Both proteins and nucleic acids may be immobilized to a variety of solid supports. For high density microarrays, glass slides are the preferred substrates because of their flatness and optical properties. Better spot resolution is also possible on nonporous glass as opposed to porous membranes, primarily due to a reduction in diffusion at the surface–liquid interface. However, keep in mind that spot (droplet) diffusion can occur on most substrates by the actions of surfactants and other wetting agents including proteins. Control of spot size and morphology is required in order to achieve reproducible and reliable results with microarrays.

Membranes such as NC supported on glass may be more applicable for protein microarrays than glass substrates. Supported charged nylon membranes for microarrays are currently entering the marketplace as well. The essential ingredient for protein is water. Protein hydration reduces the likelihood for surface denaturation. Hydrophilic membranes allow proteins to

be adsorbed and maintained in their hydrated states. Hydrogels are also important milieus for consideration, especially for proteins, because of their ability retain high water contents. Such 3D surfaces (hydrogels; membranes; other porous substrates) also may provide greater probe density, allowing for increases in sensitivity and dynamic range over 2D surfaces such as planar glass substrates.

Plastic substrates are now utilized for both nucleic acid and protein microarrays. In particular, plastic microarrays have been introduced for high-throughput applications based upon the microtiter plate which is the standard automation platform.

The nucleic acid microarray is now a well-established tool. The attachment of oligonucleotides and cDNA to surfaces has been studied in some detail. Linkers and extended spacer arms allow nucleic acid probes additional degrees of freedom to interact with target molecules in bulk solutions, thereby improving hybridization efficiency and target detection. Linkers can also reduce unwanted interactions of probe and surface. However, complex linking systems such as dendrimers that can potentially increase the probe density and block nonspecific adsorption are not always the optimal choices. Probe density is an important element in the design of a microarray, whether for proteins or nucleic acids. We have learned with nucleic acids that higher probe densities do not necessarily lead to increased hybridization efficiency. Optimal probe surface (distribution) density still needs to be determined.

Substrates for the creation of protein microarrays were initially selected from those used for DNA arrays, for example, PLL glass slides. At first, these substrates proved to be sufficient for antibody microarray studies. However, not all proteins will behave well or similarly on a particular substrate material. New solid phases applicable for protein microarrays need to be found.

References

Anderson, M.L.M. and Young, B.D., Quantitative filter hybridization, in *Nucleic Acid Hybridization—A Practical Approach*, Hames, B.S. and Higgis, S.J., Eds., IRL Press, Oxford, 1985, chap. 4.

Angenendt, P., Glokler, J., Sobek, J., Lehrach, H., and Cahill, D.J., Next generation of protein microarray support materials': *Chromatography* A 1009: 97–104, 2003. evaluation for protein and antibody microarray applications.

Beattie, W.G., Meng, L., Turner, S.L., Varma, R.S., Dao, D.D., and Beattie, K.L., Hybridization of DNA targets to glass-tethered oligonucleotide probes, *Mol. Biotechnol.*, 4, 213–225, 1995.

Beier, M. and Hoheisel, J.D., Versatile derivatization of solid support media for covalent bonding on DNA microchips, *Nucleic Acid Res.*, 27(9), 1970–1977, 1999.

Belosludtsev, Y., Iverson, B., Lemeshko, S., Eggers, R., Wiese, R., Lee, S., Powdrill, T., and Hogan, M., DNA microarrays based on noncovalent oligonucleotide attachment and hybridization in two dimensions, *Anal. Biochem.*, 292, 250–256, 2001.

Benters, R., Niemeyer, C.M., Drytschmann, D., Blohm, D., and Wohrle, D., DNA microarrays with PAMAM dendritic linker systems, *Nucleic Acid Res*., 30(2), 1–7, 2002.

Bentley, D.R., Todd, C., Collins, J., Holland, J., Dunham, I., Hassock, S., Bankier, A., and Giannelli, F., Development and application of automated gridding for efficient screening of yeast and bacterial ordered libraries, *Genomics*, 12, 534–541, 1992.

Call, D.R., Chandler, D.P., and Brockman, F., Fabrication of DNA microarrays using unmodified oligonucleotide probes, *Biotechniques*, 30, 368–379, 2001.

Dattagupta, N., Rae, P.M., Huguenel, E.D., Carlso, E., Lyga, A., Shapiro, J.A., and Albarella, J.P., Rapid identification of microorganisms by nucleic acid hybridization after labeling the test sample, *Anal. Biochem*., 177, 85–89, 1989.

Delehanty, J.B. and Ligler, F.S., Microarray immunoassay for simultaneous detection of proteins and bacteria, *Anal. Chem*., 74: 5681–5687, 2002.

Drmanac, S. and Drmanac, R., Processing of cDNA and genomic kilobase-size clones for massive screening, mapping and sequencing by hybridization, *Biotechniques*, 17, 328–336, 1994.

Ekins, R.P., Ligand assays: from electrophoresis to miniaturized microarrays, *Clin. Chem*., 44, 2015–2030, 1998.

Fodor, S.P.A., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T., and Solas, D., Light-directed, spatially addressable parallel chemical synthesis, *Science*, 251, 767–773, 1991.

Guo, Z., Guilfoyle, R.A., Thiel, A.J., Wang, R., and Smith, L.M., Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports, *Nucleic Acid Res*., 22(24), 5456–5465, 1994.

Guschin, D., Yershov, G., Zaslavsky, A., Gemmell, A., Shick, V., Proudnikov, D., Arenkov, P., and Mirzabekov, A., Manual manufacturing of oligonucleotide, DNA and protein microchips, *Anal. Biochem*., 250, 203–211, 1997.

Haab, B.B., Dunham, M.J., and Brown, P.O., Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions, *Genome Biol*., 2, 4.1–4.13, 2001.

Hamaguchi, Y., Aso, Y., Shimada, H., and Mitsuhashi, M., Direct reverse transcription-PCR on oligo(dT)-immobilized polypropylene microplates after capturing total mRNA from crude cell lysates, *Clin. Chem*., 44, 2256–2263, 1998.

Hornberg, J.J., de Hass, R.R., Dekker, H., and Lankelma, J., Analysis of multiple gene expression array experiments after repetitive hybridizations on nylon membranes, *Biotechniques*, 33, 108–117, 2002.

Huang, R.P., Detection of multiple proteins in an antibody-based protein microarray system, *J. Immunol. Methods*, 255, 1–13, 2001.

Joos, B., Kuster, H., and Cone, R., Covalent attachment of hybridizable oligonucleotides to glass supports, *Anal. Biochem*., 247, 96–101, 1997.

Joos, T.O., Schrenk, M., Hopfl, P., Kroger, K., Chowdhury, U., Stoll, D., Schorner, D., Durr, M., Herick, K., Rupp, S., Shon, K., and Hammerle, H., A microarray enzyme-linked immunosorbent assay for autoimmune diagnostics, *Electrophoresis*, 21, 2641–2650, 2000.

Khrapko, K.R., Lysov, Y.P., Khorlin, A.A., Ivanov, I.B., Yershov, G.M., Vasilenko, S.K., Florentiev, V.L., and Mirzabekov, A.D., A method for DNA sequencing by hybridization with oligonucleotide matrix, *J. DNA Sequencing Mapping*, 1, 375–388, 1991.

Khrapko, K.R., Lysov, Y.P., Khorlin, A.A., Shick, V.V., Florentiev, V.L., and Mirzabekov, A.D., An oligonucleotide hybridization approach to DNA sequencing, *FEBS Lett*., 256, 118–122, 1989.

Lane, S., Birse, C., Zhou, S., Matson, R., and Liu, H., DNA array studies demonstrate convertent regulation of virulence factors cph1, cph2, and efg1 in *Candida albicans*, *J. Biol. Chem*. 276, 48988–48996, 2001.

Lee, P.H., Sawan, S.P., Modrusan, Z., Arnold, L.J., Jr., and Reynolds, M.A., An efficient binding chemistry for glass polynucleotide microarrays, *Bioconjugate Chem*., 13, 97–103, 2002.

Liu, Y. and Rauch, C.B., DNA probe attachment on plastic surfaces and microfluidic hybridization array channel devices with sample oscillation, *Anal. Biochem*., 317, 76–84, 2003.

Lueking, A., Horn, M., Eickhoff, H., Bussow, K., Lehrach, H., and Walter G., Protein microarrays for gene expression and antibody screening, *Anal. Biochem*., 270, 103–111, 1999.

MacBeath, G. and Schreiber, S.L., Printing proteins as microarrays for high-throughput function determination, *Science*, 289, 1760–1763, 2000.

Matson, R.S., Rampal, J.B., Pentoney, S.L., Anderson, P.D., and Coassin, P., Biopolymer synthesis on polypropylene supports: oligonucleotide arrays, *Anal. Biochem*., 224, 110–116, 1995.

Matson, R.S., Solid-phase reagents, in *Immunoassays: A Practical Approach*, Gosling, J.P., Ed., Oxford University Press, New York, 2000, chap. 5.

Matson, R.S., Milton, R.C., Cress, M.D., and Rampal, J.B., Microarray-based cytokine immunosorbent assay, Oak Ridge Conference, 2001, Poster 20.

Moody, M.D., Van Arsdell, S.W., Murphy, K.P., Orencole, S.F., and Burns, C., Array-based ELISAs for high-throughput analysis of human cytokinds, *Biotechniques*, 31, 1–7, 2001.

Nizetic, D., Zehetner, G., Monaco, A.P., Gellen, L., Young, B.D., and Lehrach, H., Construction, arraying, and high-density screening of large insert libraries of human chromosomes X and 21: their potential use as reference libraries, *Proc. Natl. Acad. Sci. USA*, 88, 3233–3237, 1991.

Rasmussen, S.R., Larsen, M.R., and Rasmussen, S.E., Covalent immobilization of DNA onto polystyrene microwells: the molecules are only bound at the 5′ end, *Anal. Biochem*., 198, 138–142, 1991.

Saiki, R.K., Walsh, P.S., Levenson, C.H., and Erlich, H.A., Genetic analysis of amplified DNA immobilized sequence-specific oligonucleotide probes, *Proc. Natl. Acad. Sci. USA*, 86, 6230–6234, 1989.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O., Quantitative monitoring of gene expression patterns with a complementary DNA microarray, *Science*, 270, 467–470, 1995.

Seong, S.Y., Microimmunoassay using a protein chip: optimizing conditions for protein immobilization, *Clin. Diagn. Lab. Immunol*., 9, 927–930, 2002.

Shchepinov, M.S., Case-Green, S.C., and Southern, E.M., Steric factors influencing hybridization of nucleic acids to oligonucleotide arrays, *Nucleic Acid Res*., 25(6), 1155-1161, 1997.

Silzel, J.W., Cerecek, B., Dodson, C., Tsong, T., and Obremski, R.J., Mass-sensing, multianalyte microarray immunoassay with imaging detection, *Clin. Chem*., 44, 2036–2043, 1998.

Southern, E.M., Detection of specific sequences among DNA fragments separated by gel electrophoresis, *J. Mol Biol*., 98, 503–517, 1975.

Southern, E.M., Maskos, U., and Elder, J.K., Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models, *Genomics*, 13, 1008–1017, 1992.

Taylor, S., Smith, S., Windle, B., and Guiseppi-Elie, A., Impact of surface chemistry and blocking strategies on DNA microarrays, *Nucleic Acid Res.*, 31(16), 1–19, 2003.

Vasiliskov, A.V., Timofeev, E.N., Surzhikov, S.A., Drobyshev, A.L., Shick, V.V., and Mirzabekov, A.D., Fabrication of microarray of gel-immobilized compounds on a chip by copolymerization, *Biotechniques*, 27, 592–605, 1999.

Wiese, R., Belosludtsev, Y., Powdrill, T., Thompson, P., and Hogan, M., Simultaneous multianalyte ELISA performed on a microarray platform, *Clin. Chem.*, 47, 1451–1457, 2001.

Zammatteo, N., Jeanmart, L., Hamels, S., Courtois, S., Louette, P., Hevesi, L., and Remacle, J., Comparison between different strategies of covalent attachment of DNA to glass surfaces to build DNA microarrays, *Anal. Biochem.*, 280, 143–150, 2000.

Zhang, Y., Coyne, M.Y., Will, S.G., Levenson, C.H., and Kawasaki, E.S., Single-base mutational analysis of cancer and genetic diseases using membrane bound modified oligonucleotides, *Nucleic Acid Res.*, 19(14), 3929–3933, 1991.

Zhu, H., and Snyder, M., Protein chip technology, *Curr. Opinion Chem. Biol.*, 7, 55–63, 2003.

Zhu H., Klemic, J.B., Chang, S., Bertone, P., Casamayor, A., Klemic, K.G., Smith, D., Gerstein, M., Reed, M.A., and Snyder, M., Analysis of yeast protein kinases using protein chips, *Nat. Genetics*, 26, 283–289, 2000.

chapter four

Arraying processes

> Lady Macbeth: Out, damned spot! out, I say! — One: two: why, then, 'tis time to do't.
>
> **William Shakespeare**
> ***The Tragedy of Macbeth***

Introduction

An *array* is simply a collection of small spots on a surface organized in a particular geometric pattern. All microarrays share this feature. The material placed on the surface is what differentiates their utility. That is, we can create arrays of DNA, proteins, carbohydrates, small molecules, or even cells if we choose and then use these arrays to look at some aspect of biology on a global scale, for example, examining the expression of the entire yeast genome. However, to do this in a quantitative manner, we must rely on high sensitivity labeling schemes and sophisticated detection systems to see the spots and on complicated algorithms to arrange the data in a meaningful way in order to draw conclusions about our experiment. All these functions ultimately depend on the *equality* of the little spots.

What is spot equality? First, ask yourself: What do you wish to accomplish with a microarray? Most likely you want to use the array to measure an event that will be informative of a specific biological response of interest. Two basic levels of data can be obtained. The first is a somewhat qualitative yes-or-no or plus-or-minus examination of an array image: Did the spot light up or not? The second level provides quantitative information via the assignment of numerical values to spot intensity. This allows us to answer the question by determining the quantity: How bright is the spot? Normally, we want to know how bright the spot is relative to other spots in the array in order to ascribe some notion concerning biological relationships.

This is where spot equality becomes important. If we are to compare signal intensities among spots, we want the spots to behave very similarly to one another. The reality is that spot equality is difficult to achieve. At first

glance, it may be difficult to understand this kind of imprecision of a microarray. After all, the spots all look to be of the same uniform size and shape. If we could rely on the use of spot diameter as a principal quantitative measure, then we would find an answer. Unfortunately, while spotted arrays can allow control of spot diameter within 5 to 10% coefficient of variance (CV), this parameter alone has little bearing on the performance of the array.

What lies within the spot must be controlled. Thus, spot morphology and the spatial arrangements of biomolecule capture agents within the spot become important concerns. Ideally, we would like each spot to contain the same number of molecules properly oriented and in their native conformation to assure optimal and equivalent levels of binding to their cognate targets. That would be genuine spot equality. In the sections to follow, we will explore different approaches to producing microarrays that achieve spot equality.

Creating spotted microarrays

It is important to understand the context of "spotted" microarrays. The term refers to the creation of arrays in which the spots are produced by contact printing methods or by the dispensing of droplets onto a surface in which case, the biomolecules within these spots are presynthesized (or of a native form) prior to spotting. This excludes arrays produced by *in situ* means such as by photolithographic or other related processes used to produce tethered biomolecules from monomers or precursors. For example, Affymetrix produces arrays by flooding a photo-deprotected region with active monomers. Agilent arrays are produced by spotting active monomers on acid-deprotected sites.

We will use the spotting of oligonucleotides as a general example of the processes involved. Oligonucleotide arrays were originally introduced to accomplish sequencing by hybridization via a number of formats including both *in situ*- and *ex situ*-based tethering to a variety of substrates (see Chapter 3 and its references). Before we attempt to discuss specific protocols for the spotting of oligonucleotides, it is worthwhile to understand something about the fundamental processes involved. Four important factors should be considered in the spotting of *ex situ*-synthesized probes onto a surface:

1. Substrate
2. Probe (ink) composition
3. Printing environment
4. Printing mechanics

Substrates

The primary substrate for spotted arrays is the glass slide. The salient physical and chemical features of a microarray slide are optical clarity (including low fluorescence background), flatness, and coating uniformity. While

standard microscope slides were used originally, most commercial manufacturers replaced them with higher grade glasses to reduce the number of irregularities in surface features and improve optical quality.

Slides specifically selected for microarray applications should be used. They are available as ultracleaned (an important consideration) and untreated for those who wish to prepare their own surfaces or they can be purchased with a variety of precoated surface chemistries (e.g., lysine, aldehyde, or epoxide). The densities of reactive groups and surface coating uniformity are difficult to control. Thus, if lot-to-lot slide consistency is most important factor, consider using commercially available slides that are quality controlled.

Probe composition (print buffer)

Perhaps the least understood factor in the process of microarraying is the print buffer (probe ink) composition. This may not be too much of a surprise because manufacturers of computer printers offer consumers a multitude of different inks (whose formulas are closely guarded trade secrets) for use with a particular printer and kind of paper. In fact, it can be argued that the ink is perhaps the most important piece of the consumable product stream for this manufacturing sector.

Surely the printing of oligonucleotides onto the glass substrate is straightforward, is it not? No, not completely. A search of the very popular Gene Arrays site maintained by the University of California at San Francisco (gene-arrays@itssrv1.ucsf.edu) reveals repetition of many questions on this topic. A number of factors influence the ability to reliably print oligonucleotide probes onto a surface such as glass. The most important ones are represented in Figure 4.1.

First, consider oligonucleotide structure. Unmodified nucleic acid probes may be viewed as negatively charged polymers (they are polyanions)

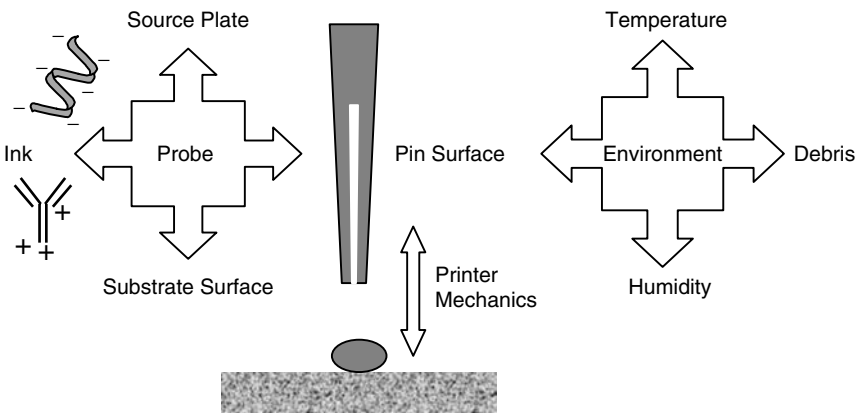


Figure 4.1 Factors influencing spotting.

under normal physiological conditions. In that state, they readily adsorb to positively charged surfaces such as glass slides coated with poly-L-lysine (PLL) because this polycation would carry a net positive charge. Oligonucleotides can also bind other counterions such as metal ions present on printing pins. The purine and pyrimidine bases participate in hydrogen bonding and van der Waals interactions, allowing for adsorption of nucleic acids onto a variety of materials. For example, oligonucleotides bind extremely well to polystyrene surfaces. Microtiter plates of molded polystyrene are often used for the storage of oligonucleotides or as source plates during the printing process.

The ink is composed of the probe, buffer, and most often a wetting agent that allows uniform deposition of the oligonucleotide to the substrate surface to control spot size and spot morphology. Other additives to the ink may be present to prevent or slow evaporation in an effort to control spot size. Fluorescent or other dye stuffs are sometimes included to monitor printing efficiency. The ink can therefore represent a complex matrix for the probe.

Selection of the ingredients must be undertaken with extreme care, keeping in mind that these components may interact with the various surfaces they contact and with the oligonucleotide probe itself. One of the most common effects associated with improper selection of printing ink is contribution to the background signal following hybridization, or ingredients may cause precipitation of the probe during the printing process. A detailed discussion of print buffer selection is provided in the protocols section later in this chapter.

Printing environment

Environmental conditions are also very important. For both quill pin printing and noncontact dispensers (piezo- or solenoid-based), we highly recommend filtering of all buffers and inks prior to use to reduce the levels of dirt and other debris that can clog a system. During printing, the temperature and relative humidity should be held within a specific range (20°C and relative humidity of 50 to 65%) to avoid print variation. Avoid wide swings in temperature and humidity during the print run. Do not locate a printer next to an air conditioning vent or where the room air flow is subject to frequent change.

Most microarray printers are now housed in boxes set up to control humidity and many are provided with high efficiency particulate air (HEPA) filtration to remove particulates. However, a word of caution: attempting to run a HEPA filter along with a humidifier may be counter-productive if the humid air is exhausted by the HEPA system. Check the design of the environmental chamber to verify whether both devices can be used during a print run.

Substrates should be kept clean, preferably sealed and opened just prior to use. Gloves should be worn to prevented unwanted fingerprints and other forms of contamination from hands. If possible, the printing operation

should take place in a clean room environment (class 10,000) in which the printer is housed within a class 100–1000 HEPA filtered box.

Finally, pins should be cleaned and inspected to ensure they are free of any obstructions of the tips prior to use. Manufacturers have specific recommendations on best methods of cleaning their pins. Some vendors such as TeleChem International (www.arrayit.com) offer cleaning solutions and equipment (sonic baths, holders, etc.) for their pins and print heads. In our laboratory, we routinely inspect pins under a microscope before and after cleaning and keep photographic images as records.

The manner in which the spotting pin makes contact with the substrate surface is dependent both on the design of the pin and the designs of the print head and printer mechanism. A variety of microarray printers, pin designs, and print head mechanisms are now commercially available. While printers may accept a variety of print heads with appropriate adapter hardware, there exist a very close match between pins and print heads. Such physical tolerances make it difficult to mismatch pins with print heads and achieve good spot quality. For example, in order for a quill pin to function properly in a free floating print head (Figure 4.3) the tolerance on the pin's shaft diameter might be O.D. +/– 0.002 inches. That represents about a 3% variance on pin shaft diameters of 1.5 mm to 2.0 mm that are commonly used.

Printing mechanics

The printing mechanics of dispense and pin printers are obviously different. Moreover, pin modes are of contact and noncontact types (Figure 4.2). We can refer to these as stamp and touch-off printing methods, depending upon whether the pin actually touches the surface. The most common microarray printing technique employs a split pin (quill) to deposit the probe. The fundamental process is simple enough: Fill the quill capillary with printing solution and deliver a very small droplet to the surface upon taping or striking the tip

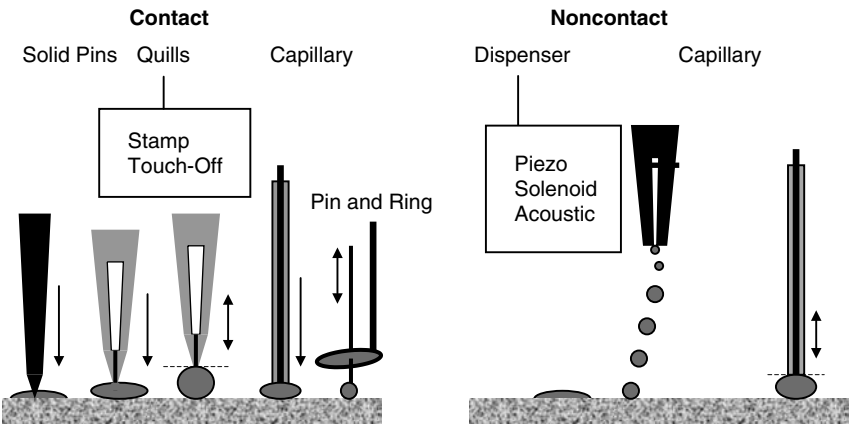


Figure 4.2 Microarray printing mechanisms.

at the surface. This is best accomplished using a computer-controlled x,y plotter with a Z axis motion control. However, the repetitive and precise printing of small droplets onto a glass slide can be difficult.

Commercialization of the microarrayer has been an important factor in the expansion of microarray-based applications. Today, most commercial arrayers meet the resolution and repeatability criteria necessary for producing medium to high density microarrays, although arrayers differentiate in areas such as cost, ease of use, deck capacity, and high-throughput robustness issues. Table 4.1 summarizes different types of arrayers and their throughput capabilities. It lists a selection of instruments from bench-top models for low capacity users to high-throughput printers enabling custom microarray manufacturing. Unfortunately, little standardization in design of pins and print heads has been achieved, making it difficult to directly compare performances of products (Figure 4.3).

In recent years, some consolidation in the industry has occurred and a few companies listed are no longer in business, for example, Genometrix, an early entry in the high-throughput custom microarray marketplace. Others including well-established companies like GeneMachines and Cartesian have survived as divisions of larger companies. Genomic Solutions (owned by Harvard Biosciences) recently acquired GeneMachines, Cartesian, and BioRobotics. However, their products are still supported and newer models are being introduced. More importantly, the table shows the many kinds of technologies that can be used to produce microarrays. Undoubtedly, new printing technologies that will further the expansion and advancement of the microarray field will continue to emerge.

The original quill pin developed in Pat Brown's laboratory at Stanford (see the "Print Tip Gallery" anthology of quill pins at http://cmgm.stanford.edu/pbrown/mguide/tips.html) was designed to strike the surface of the glass slide with enough force to eject a droplet from the capillary (see U.S. Patents 5,807,522 and 6,110,426). Doing so required overcoming the surface tension forces at the meniscus. The Stanford quill was commercialized by Majer Precision (Tempe, AZ) which was responsible for the fabrication of some of the more successful pin designs. Majer Precision also offers a specifically designed holder in which the quills are spring-loaded. This assures even striking of the substrate surface and reduces wear on the pin tips.

The holder is produced from a stress-resistance aluminum alloy that maintains the necessary structural tolerances. It is coated with nickel to guard against corrosion and permits a low coefficient of friction (embedded polytetrafluoroethylene [PTFE]) with the quill (Figure 4.4). Genetix Ltd. also designed a special holder to reduce friction. The quill shafts actually ride on opposing microball bearings, further reducing contact with the holder body (Figure 4.5).

In contrast to this stamping action is the touch-off strategy employed by TeleChem quills (see Martinsky U.S. Patent 6,101,946 for descriptions of design features). Unlike the pointed pin tips used for stamping, TeleChem's pins are flat at the tips (Figure 4.6). The force associated with the acceleration of the pin toward the surface is used to break the meniscus, allowing a

Table 4.1 Comparison of Arrayer Types

Manufacturer	Trade name	Print head type	Pin/dispenser type	Number of Pins/Tips	Number of slides on deck	Approximate print rate[a] (spots/sec)
Affymetrix	Model 417	Ring and pin	Solid–pin	4	20	1.7
Cartesian	ProSys	SynQuad	Solenoid tip, noncontact	8	9	2.4
Cartesian	PixSys 7500	TeleChem	Quill	4	6	3
GeneXP	BioGrid Arrayer	TeleChem	Quill	8	15	5.2
Genetix	QArray[2]	HPLF	Quill	12	15	8.4
GeneMachines	OmniGrid 300	TeleChem	Quill	48	300	083.6
Genospectra	N/A	Custom	Fiberoptic capillary bundle, noncontact	10,000	3,000	116
IMTEK	TopSpot/P	Nozzle	Noncontact	96	300	120
Genometrix	N/A	Custom	Capillary tube bundle, contact	256	15	427

[a] Print rates based on start-to-finish runs estimated from vendor specifications or supplied data with a certain number of slides.

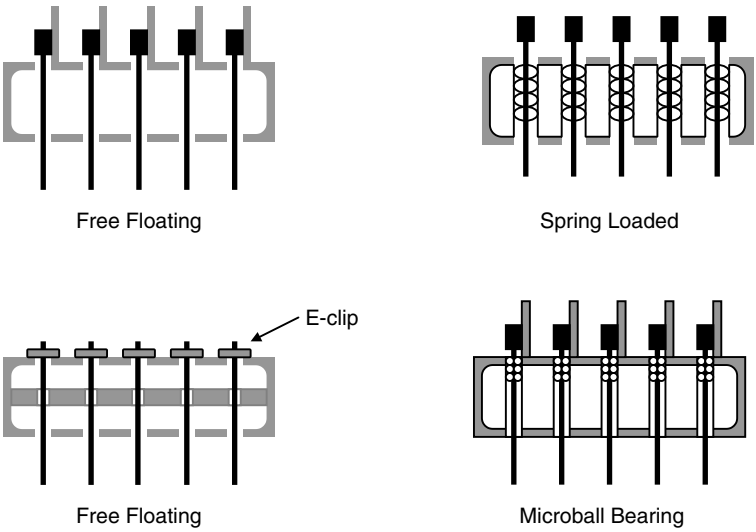


Figure 4.3 Print head designs.

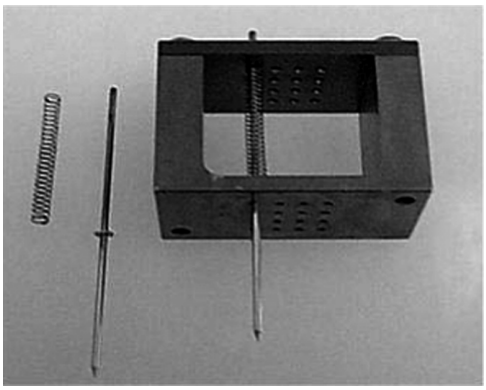
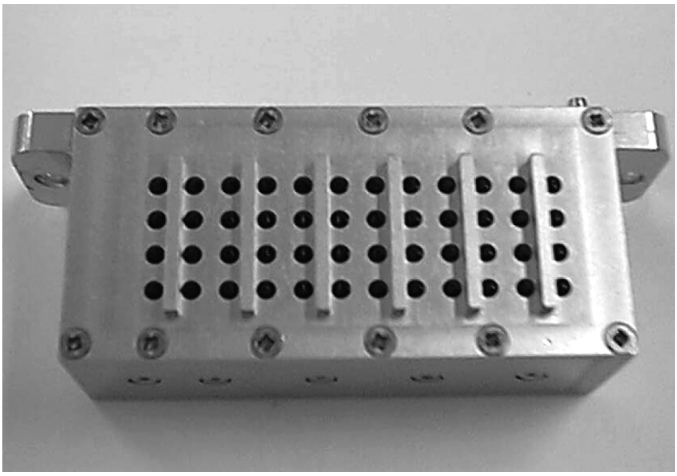


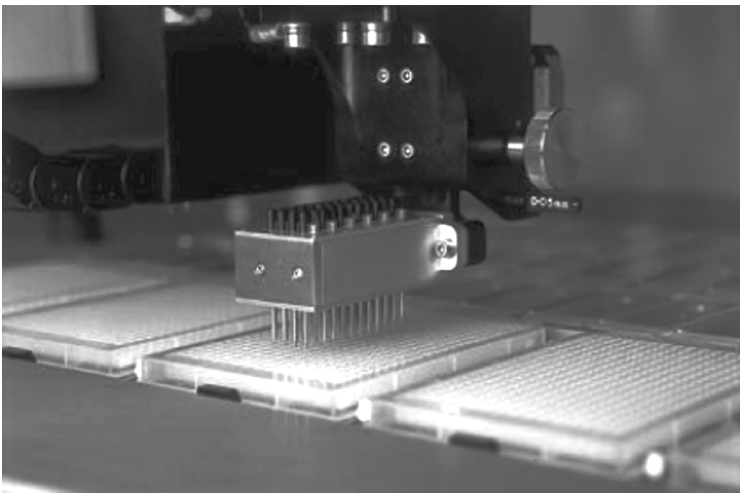
Figure 4.4 Majer Precision print head and MicroQuill® pins. (Majer Precision Engineering, Inc., Tempe, AZ.)

droplet to make contact only with the substrate. Relying upon the adhesive forces on the substrate to be greater than those on the pin allows the substrate to capture the droplet before the meniscus can reform. The pin does not strike the surface in order to propel the droplet. It touches the droplet onto the surface. However, this is a subtle difference between the two processes and it is not uncommon for a novice to overdrive a TeleChem pin into a slide. The most likely outcome is a reduced lifetime for the pin and a loss in spot quality during printing.

Solid pins are also used for microarray printing. These types of pins were adopted from gridding applications where they were primarily used



(a)



(b)

Figure 4.5 (a) Genetix high-precision low friction (HPLF) print head. (Genetix Limited, United Kingdom.) (b) HPLF with quills. (Photo courtesy of Genetix Limited, United Kingdom.)

to transfer colony plaques from agar or microtiter plates onto nylon membranes for cDNA-based gene expression analysis. For example, Beckman Coulter offers the high-density replicating tool (HDRT) for use with its Biomek series of liquid handling workstations for this purpose (Figure 4.7). Such blots were the forerunners of slide-based microarrays.

While automated grid blotting may be used to create microarrays, the linear accuracy requirements for these robots are not as strict as those found on microarray printers. This limits their use to low density arraying where

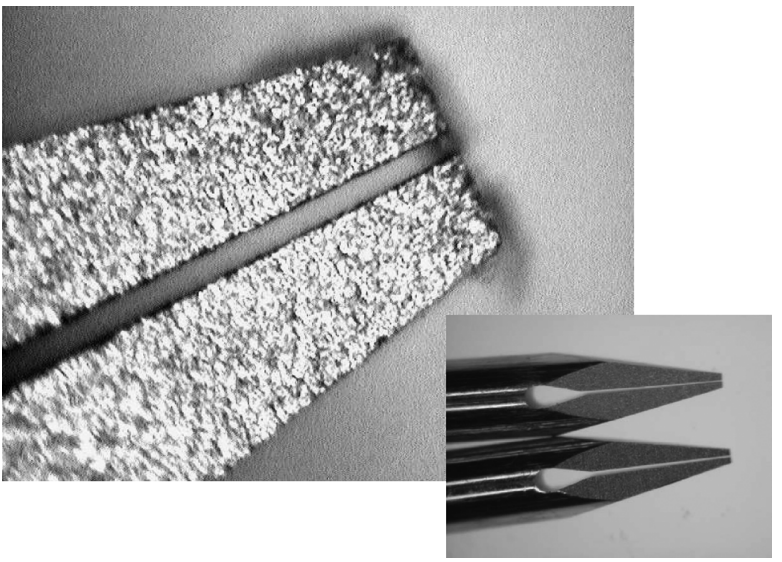


Figure 4.6 Close-up view of ChipMaker™ quill. (TeleChem International, Inc., Sunnyvale, CA.)

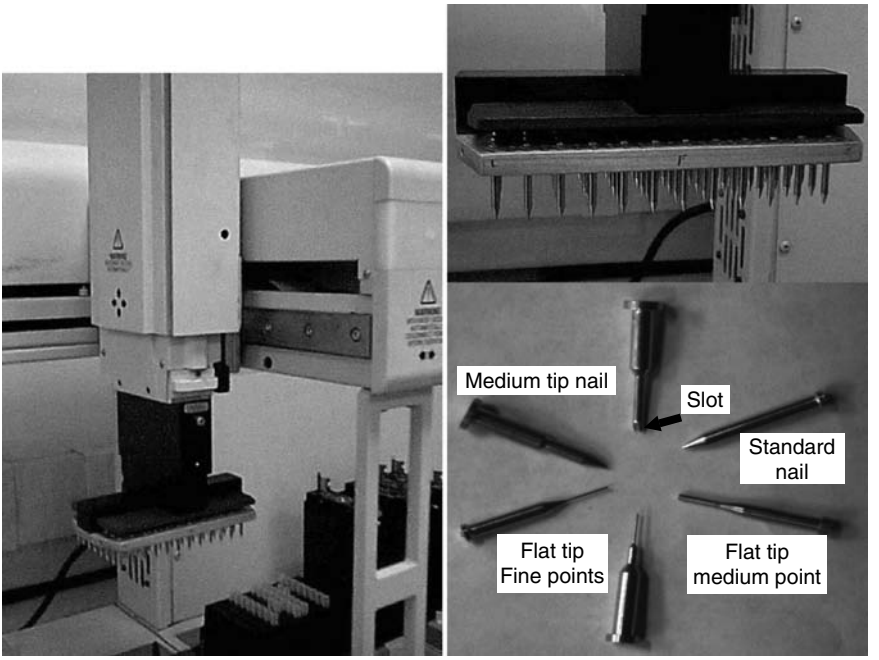


Figure 4.7 Biomek 2000 laboratory automation workstation equipped with HDRT and transfer nails.

the center-to-center spacing of the spots is on the order of 500 µ. However, advances in solid pin technology have progressed and delivered new surface features and pin shapes that provide precision printing comparable to quill pins. The main advantages of these new generations of solid pins are uniformity in spot formation and morphology, primarily due to more controlled droplet pick-up. Since they lack capillary slits, they are much easier to clean and maintain. The main disadvantage is that each print requires a re-inking of the pin from the source plate. This consumes more time than using a quill pin that can produce multiple prints before refill. However, unlike the quill pin, the solid pins do not require preprinting. Additional time can be saved by using an array of pins in the print head.

Genetic MicroSystems (GMS) originally offered a unique format that employed solid pins. The device is now known as the Affymetrix 417 arrayer since the acquisition of GMS by Affymetrix in February 2000. Commonly referred to as the pin-and-ring system, the microarrayer uses a ring to capture a film of print buffer (much the same way as a bubble wand functions). After loading with the probe solution, the print head is positioned over the substrate and a small amount of liquid is delivered from the tip of a solid pin that pierces through the film ring. The film ring reforms after the pin is drawn through it and the process is repeated. Arrays produced with the pin-and-ring system were of high quality. The main disadvantages were the need to use considerable print buffer in the ring and the inability to use a large number of such pins in the print head, thereby reducing throughput.

Another technique of spotting is capillary printing. The most notable process was developed by Genometrix, Inc. (no longer in business). Open capillary tubes were bundled to create a single head that would strike a surface to simultaneously deposit up to 256 spots. Each capillary terminated into a separate well of a 384-well source plate. Pressure was applied across the plate, forcing liquid into the capillary. Droplets were formed at the open end and released from the capillary following contact with the substrate. The printing process was very rapid and allowed the placement of tens of thousands of spots within a few minutes. A similar approach was recently introduced by GenoSpectra (Hayward, CA). The capillary of the device does not make direct contact with the surface, but allows the droplet to touch the substrate, then break away as the capillary retracts from the surface.

As for spotting precision, state-of-the-art microarrays such as the MicroGrid II (BioRobotics, now a division of Genomic Solutions–Harvard Biosciences) can achieve 10-µ spot (x,y) resolution at a repeatability of about 1 µ. This allows for relatively high density printing onto glass slides. However, control of the Z height travel and acceleration is also very important in producing good spot quality, especially when printing high density arrays. The BioRobotics system employs a "soft touch" that slows the pins (Z deceleration) upon approach to the substrate to ensure controlled spot size.

Most contact microarrayers utilize pin heads that permit the pins to float. As a pin strikes the surface, it can quickly rebound (float), thereby compensating for any surface irregularities encountered during printing. For

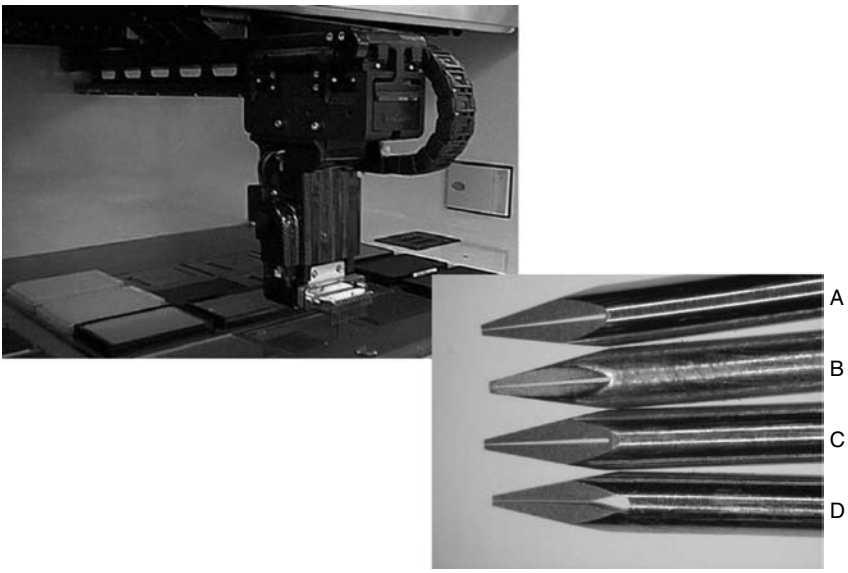


Figure 4.8 CMP™ quill pin types. (TeleChem International, Inc., Sunnyvale, CA.) A through C: CMP-10 standard quill pins. D: CMP-10B quill pin.

example, the TeleChem pin head is composed of a precision-bored brass block and holds stainless steel pins at a close tolerance. The brass permits low friction travel of the steel pins. The pin motion is controlled by the Z travel of the print head and gravity.

The Majer Precision print head employs spring action to control pin movement. Each pin is seated in a spring coil, and this allows a similar soft-touch landing and gentle but controlled rebound of each pin upon striking the surface. The Genetix pin head seats each pin between microball bearings to achieve low friction travel and precise rebound.

Microarray pins

A limited variety of quill spotting pins are available. They represent improvements of the original designs provided by Brown's group at Stanford. What changed were the materials used to construct the pins. Surface finish and alloy mix appear to exert some effect on printing performance. Thus, quills used to print DNA may not be optimal for the printing of proteins. Currently, TeleChem remains the major manufacturer of quill pins under the Chipmaker and Stealth trade names. The quills are constructed of stainless steel (Figure 4.8). Majer Precision offers quills under the MicroQuill trade name. They are produced from very hard 17-4 stainless steel that increases the lifetime of the tips and is corrosion-resistant (Figure 4.4). Point Technologies offers tungsten-coated quill pins under the PT Accelerator trade name (Figure 4.9).

Tungsten was selected because of its superior strength and hardness relative to stainless steel and titanium. The tungsten pins are electropolished

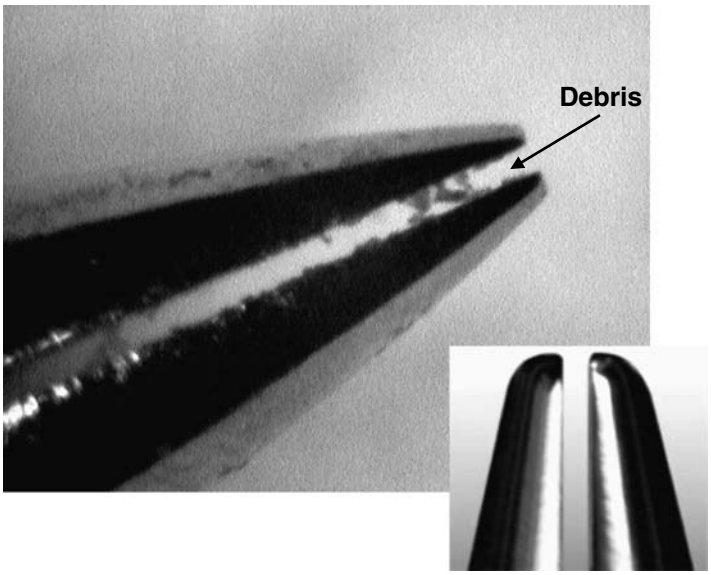


Figure 4.9 PT Accelerator® split pins. (Point Technologies, Inc., Boulder, CO.)

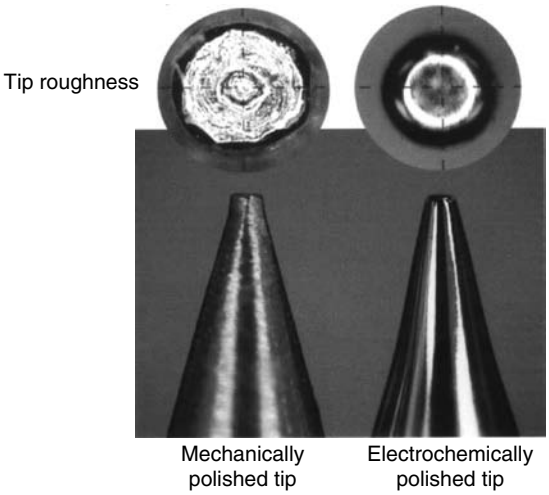


Figure 4.10 PT Accelerator® solid pins. (Point Technologies, Inc., Boulder, CO.)

for optimal fluid transfer. Point sharpness can be controlled down to about 25 µ. The electrochemical pointing (ECP) process leaves the surface extremely smooth and clean by chemically removing any rough features and surface impurities (Figure 4.10). No grinding or other mechanical polishing is involved because those processes can leave abrasions and metal fragments on the pins.

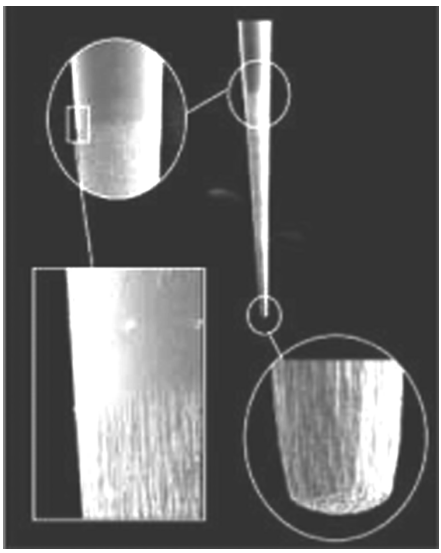


Figure 4.11 Zonal pin texturing. (Courtesy of Point Technologies, Inc., Boulder, CO.)

Highly reflective surfaces reduce fluid surface tension with the metal — making the pins essentially hydrophobic. For custom applications, Point Technologies can apply a proprietary "zonal" texturing process that creates different microsurfaces or zones on pins (Figure 4.11). For example, the tip can be rendered hydrophilic while the pin shaft is made hydrophobic to create a surface tension barrier. These features are accomplished by a combination of electrochemical etching and applied polymer coatings. The end result provides a means to control spot diameter and delivery without increasing pin size.

Other approaches

For the creation of lower density microarrays, several groups have successfully employed robotic gridding devices to print onto membrane substrates. The arraying of cDNAs (and proteins) onto membranes in this manner is well documented and still practiced. For example, Lane et al. (2001) used a standard Biomek 2000 equipped with a 384-solid pin HDRT used for gridding to create an array of 3456 cDNAs on a nylon membrane representing about 1000 *Candida albicans* genes in triplicate. A single run could produce up to 10 such blots using the available work surface.

Others such as Macas et al. (1998) successfully adapted the Biomek 2000 (Beckman Coulter), a commonly used liquid handling robot, to prepare microarray slides using a specially constructed print head and quill pins. Up to 28 microscope slides could be placed on a work surface for printing. Biomek's HDRT head was adapted to accept microarray quill pins held between two parallel plates with holes drilled on 9-mm centers to dip into 96-well source plates. The quill pins were spring-loaded similar to the design

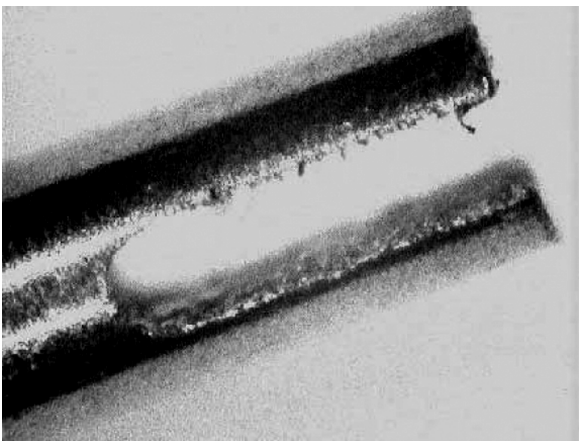


Figure 4.12 Close-up view of V&P Scientific's slot pin. (V&P Scientific, Inc., San Diego, CA.)

of Majer Precision. Using Tool Command Language (Tcl), the Biomek was reprogrammed for use as a microarrayer. Up to about 3000 elements can be spotted onto each microscope slide (125-µ spot diameter at a 500-µ spacing) using this system. The positional accuracy for n = 768 repeats on a 500-µ center was measured at a standard deviation of x ± 52.6 µ and y ± 63.7 µ.

V&P Scientific of San Diego, CA now offers a range of adapters, print heads, and slotted pins that can be used with the major commercially available liquid handling robots (Figure 4.12). These print heads can also be transformed into manual gridding devices if only a few microarrays are needed and where the cost of a robotic system is not warranted.

George and coworkers from the Lawrence Berkeley National Laboratory introduced the use of ceramic capillaries for printing with claimed improvements in the consistency of spot morphology (2001). The researchers compared the distribution of spot diameters created by the ceramic tips (K&S MicroSwiss) relative to distribution by a stainless steel quill pin (MicroQuill 2000, Majer Precision, Inc.). The ceramic capillaries (Figure 4.13) exhibited more uniform distribution of spot diameters than the quill (Figure 4.14). The mean spot diameters closely matched the size of the ceramic tip, e.g., a 50-µ tip produced a 56-µ spot, while a 132-µ tip deposited a 130-µ spot. The 50-µ tip was used to print a 12×12 array of labeled cDNA on 100-µ center-to-center spacing. While the stainless steel quills performed well, apparent tip wear following 100,000 stampings was noted. The ceramic tips are reported to perform with a lifetime beyond 225,000 deposits. However, ceramic tips require greater care in leveling the print head to the substrate surface in order to prevent damage, and drying times may need to be extended for ceramic materials.

A new quill pin design known as the "trench pen," has been created for high density microarray printing by Stephen Quake's group at the California

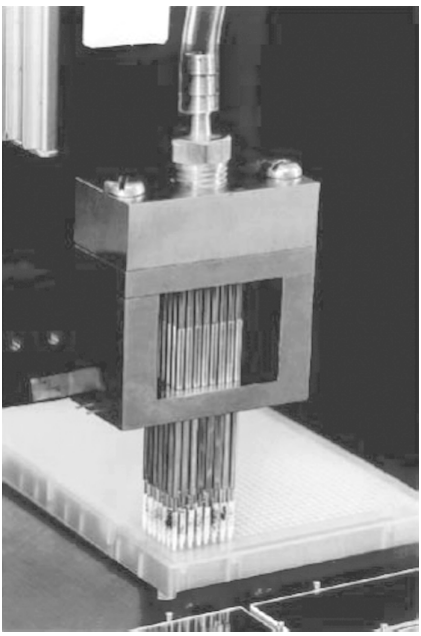


Figure 4.13 Ceramic capillary pin in holder. (From George, R.A. et al., *Genome Res.*, 11, 1780–1783, 2001. With permission.)

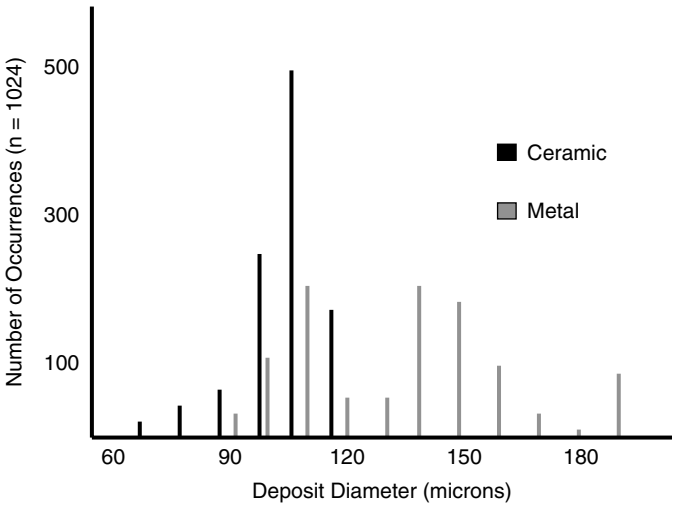


Figure 4.14 Comparison of spot diameter distributions for metal and ceramic capillary pins. (From George, R.A. et al., *Genome Res.*, 11, 1780–1783, 2001. With permission.)

Institute of Technology (Reese et al., 2003). Using optical lithography to etch away photolithographic resists on stainless steel foils, trench pens were designed with a rectangular geometry of 6 μ in depth, 30 μ in width with 30 μ sidewalls at the tips. To add structural support, the features anterior to the tip were expanded out so that the width of the sidewalls increased to 120 μ with a trench width set at 90 μ.

Because the trench pen was designed with considerable flexure, it serves as a miniature shock absorber and does not require any external fixturing (such as springs or ball bearings) in order to reduce shock from striking the substrate. In order to achieve capillary action the trench is coated with hydrophilic polyurethane. Rectangular spots ranging from 10 to 30 μ in width by 20 to 140 μ in length, depending upon the design, are produced. In one example, an array of >2500 dye spots was printed onto a 3.2-mm × 3.2-mm square corresponding to ~25,000 spots per cubic centimeter. Such densities are about twofold higher than other state-of-the-art microarray fabrication processes. According to the authors, an individual pen prints 5 to 20 spots per inking. Undoubtedly, higher throughputs will be required before the trench pen is considered a viable printing device.

In terms of performance relative to a conventional quill from Majer Precision, the trench pen showed spot size variance in the range of 10 to 20% CV, while the quill pin exhibited a spot CV ~14%. Hybridization signal variation from these trench pen printings (n = 3 arrays, 72 spots per array) ranged 1.6 to 4.3% CV, while the quill produced ~3% CV in signal. Thus, the trench pen performs at similar levels of print quality to more conventional quill pins.

There are, of course, noncontact printing devices useful for the construction of microarrays (see Figure 4.2). These are microdispensers that eject droplets by several different mechanisms (solenoid, piezoelectric, heated jet, acoustical wave). Perhaps the best-known commercial dispensers are the syringe driven-solenoid pump (e.g., Cartesian; BioDot) and piezo systems (e.g., Packard Biosciences).

IMTEK (Institute of Microsystem Technology, University of Freiburg, Germany) developed a microfluidic dispenser device for simultaneous printing from 24, 96, or 384 channels (Daub, 2002). The device consists of a reformat plate with microfluidic channels that terminate into nozzles (Figure 4.15). The plate reservoir is first filled with printing ink and then loaded automatically onto a print head. The print head piezo actuator forces the fluid from the nozzles. Droplets are formed and jetted down onto the surface (Figure 4.16). For example, 209-μ diameter spots can be printed down onto a glass slide reproducibly at CV <2% spot diameter or CV <5% signal intensity. The TopSpot E printer can produce 200 slides per hour at a 96-element density (5.3 spots/sec), while the TopSpot P unit can produce 300 slides per hour at a 1440 element density (120 spots/sec); see Figure 4.17.

Another approach is using a modified commercial ink-jet printer to dispense biological reagents. Over a decade ago, we took an HP DeskJet printer, emptied the ink from the print cartridge, and replaced it with a DNA

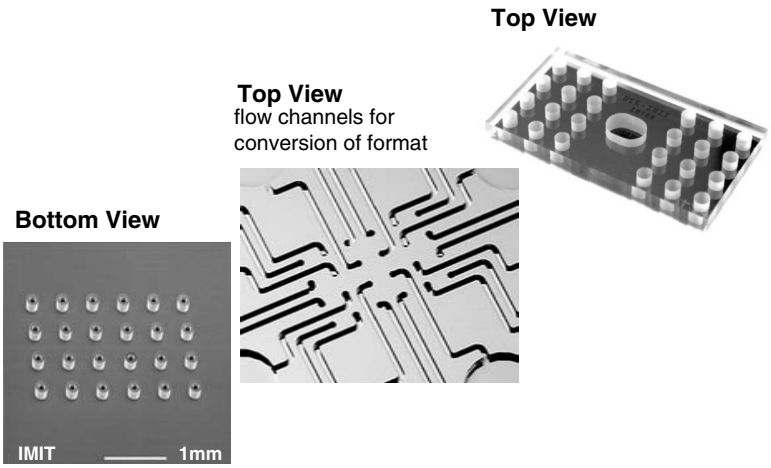


Figure 4.15 TopSpot reformatting done by capillary forces. (Courtesy of Martina Daub, IMTEK, Freiburg, Germany.)

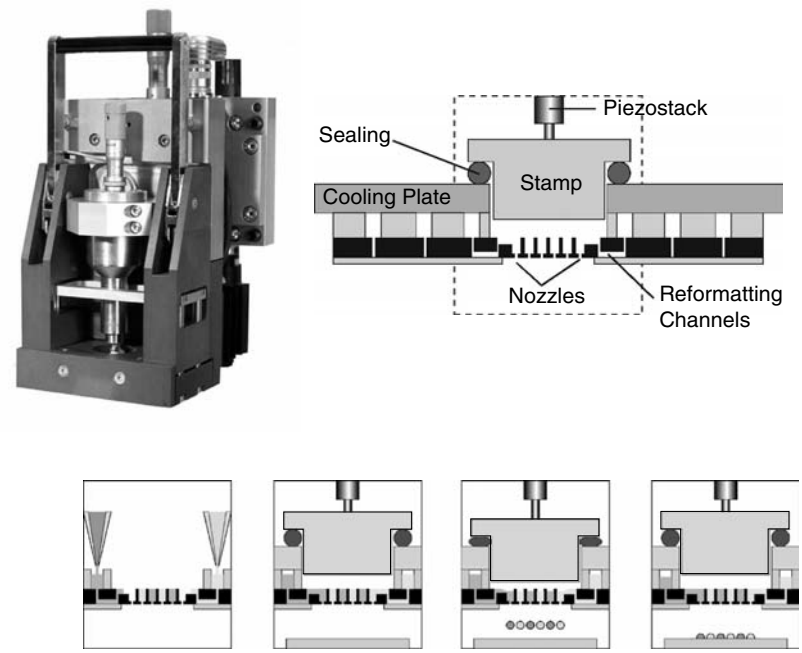


Figure 4.16 TopSpot actuator mechanism. (Courtesy of Martina Daub, IMTEK, Freiburg, Germany.)

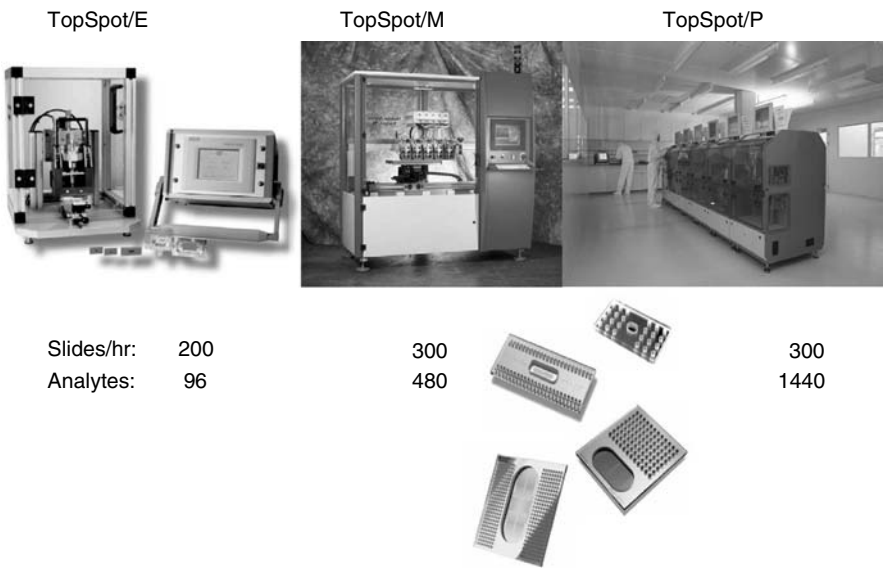


Figure 4.17 TopSpot platform scalability. (Courtesy of Martina Daub, IMTEK, Freiburg, Germany.)

or protein solution (R. Matson, unpublished data; Silzel et al., 1998). We were able to print arrays onto membranes or activated films without much trouble. The primary problem we encountered was a high number of missing spots. Moreover, in order to create larger array formats, a more versatile multihead printer and a low volume ink reservoir would have been required.

Stimpson et al. (1998) transformed an Apple StyleWriter II (Apple Computer) to handle thermal ink-jet printing of oligonucleotides. They too disassembled an ink cartridge and replaced the ink with oligonucleotides dissolved in PBS containing 7% isopropanol for printing onto membranes. Okamato et al. (2000) at the Canon Research Center, Kanagawa, Japan, reported on the fabrication of glass slide microarrays using a modified Canon Bubble Jet printer. They noted that the key to printing biologics is to avoid denaturation by heat or shearing stress forces. At least for DNA, an ink formulation was identified that allowed the printing of 10-nt to 300-bp nucleic acids at suitable concentration for optimal hybridization performance on the microarray. The ink was a combination of glycerin, urea, thiodiglycerol for wetting, and acetylenol to control viscosity. In one instance, a 5′-thiolated 18-mer oligonucleotide (the substrate was activated with maleimide groups) was prepared in this ink at 8 μM concentration and transferred onto the print head (BC-62, Canon). The printer (BJC-700J, Canon) permitted the simultaneous dispensing of six inks (24 pL/droplet) onto the glass substrate. Approximately 70-μ diameter spots were printed in an 8×8 pattern at a resolution of 100×100 dots per inch (dpi). That corresponds to roughly 4 spots per millimeter or 70-μ spots spaced at ~180 μ.

In order to print down 64 different probes on the substrate, two BC-62 print heads were aligned. This allowed the simultaneous firing of up to 12 jets and permitted the construction of a p53 gene mutation array. Both point mutations and single nucleotide polymorphism (SNP) analysis were demonstrated using this array. However, one reported drawback to Bubble Jet printing was the need to change print solution after each dispense to create the 64-element microarray. Nevertheless, this is not a difficult task with disposable cartridges, provided the operation for replacing cartridges is relatively simple and no alignment issues must be resolved.

Printer performance

Microarray printing requires careful attention and vigilance. A good understanding of the performance features of the arrayer is highly recommended. This is best accomplished by designing experiments that will measure the robustness of the arrayer, i.e., determine at what point the print quality begins to deteriorate. These experiments should be able to differentiate mechanical or software issues and those related to pin performance.

Arrayer performance is commonly measured in terms of its positional accuracy, or resolution, and repeatability. For example, a typical manufacturer's specification may state a positional resolution (x,y) at 10 µ with a repeatability of 1 µ (x,y). What this means is that the arrayer can print down a spot within 10 µ of the desired position and return to within 1 µ of that x,y coordinate most often. For printing down 100-µ spots, such levels of resolution are easily met. However, how does this arrayer perform after printing 10,000 spots on 100 slides?

To answer such a question, we must obtain data across the print run and measure variance, typically in the form of intra- and inter- spot CV. Was the last slide printed as well as the first? How many spots are missing? Do spots begin to merge or do the diameters increase or decrease across the print run? Are there systematic variations? These issues are not easily resolved and one must be very careful not to confuse the robustness of the arrayer with the performance of the pin. For example, a loss in spot quality or the disappearance of spots may be attributed to a clogged quill pin, while randomly missing spots could be due to failure of the pin or arrayer mechanism. If a pin failed to float, it might not touch down on the slide surface or pick up enough print buffer from the source plate. If the Z travel of the arrayer became erratic, the pin may not have touched down completely and may have failed to deposit the droplet. However, it would be most difficult to imagine adjacent spots suddenly appearing within 10 µ of each other from pins set in a print head on 4.5 mm centers. Most likely this problem would be a mechanical or software failure. It is always a good idea to understand the limitations of the arrayer and the performance of the pin before proceeding with the production of microarrays to be used in studies.

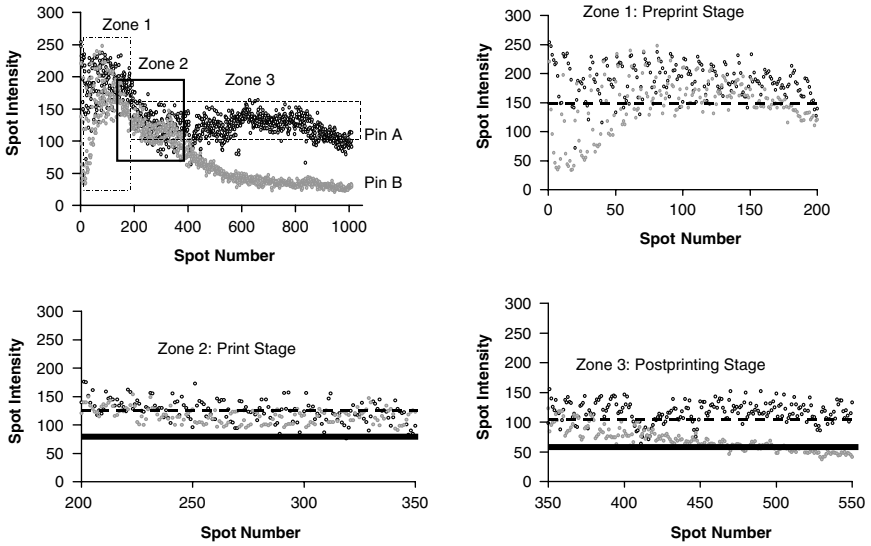


Figure 4.18 Quill pin printing profiles.

Pin performance

While the arrayer's controller mechanism and the geometry of the print head determine the spacing between spots (often expressed as center-to-center distance or pitch), the physical characteristics of the pin generally determine the spot size. Thus, we can rate pins based upon the delivery of a particular volume (e.g., 1 nL) corresponding to a spot size range (e.g., 120 to 130-μ diameter). Printing with quill pins has certain limitations. Aside from the common problems with clogging of the capillary or damage to the tip, these instruments are individually machined to a specific tolerance, i.e., no two pins are perfect matches in performance.

For example, consider the printing profiles for two pins from the same vendor as shown in Figure 4.18. The two (A and B) pins were loaded into the print head and dipped once into print buffer (containing a fluorescent dye). The number of printed spots each pin could deliver from a single inking were counted. Spot intensity per unit diameter was used as a measure of pin delivery. The first thing to notice is that the initial spotting pattern (Zone 1) is rather erratic. Zone 1 is usually reserved for preprinting — a certain number of spottings are performed on a substrate prior to initiating the construction of the actual array. This process provides a means to remove any excess ink from pins (Pin A) and permit clearing of the capillary or tip (Pin B) of any obstruction such as salt build-up or air bubbles or simply to fully wet out the tip for optimal delivery. Preprinting is generally performed elsewhere, such as on a sacrificial slide, after which printing is resumed onto a new slide at a different location on the arrayer's deck.

Each pin should be assessed for the number of preprints required. In this particular example, the pins required 150 to 200 preprints before achieving

uniform delivery (Zone 2). Notice that both pins within Zone 2 perform at about the same level that would be ideal for producing several microarrays. Depending upon the number of replicates on each microarray, Zone 2 spotting (150 dispenses) could be used for construction of 50 slides (n = 3 replicate spots per slide) per inking. Beyond Zone 2 is the postprint stage (Zone 3), where the pin runs out of ink. This is more evident for Pin B; Pin A appears to deliver far beyond that limit. However, since more than a single pin would most likely be used for printing, the number of spots per inking would be necessarily dictated by the lowest performing pin.

Obviously, the above example involved the use of higher loading capacity pins. In addition, we printed onto a hydrophobic plastic substrate. While the profile shown in Figure 4.18 applies to more hydrophilic glass slides, we expect that the number of preprints could be greatly reduced as explained in the following discussion.

The number of preprints that a particular pin requires will depend upon several parameters. First, a substrate that is hydrophilic may allow for increased wetting of the surface, thereby drawing out more liquid from the capillary, and more importantly, from the outside of the pin. The net effect would be a reduction in the number of preprints. Depending upon the buffer, this might also reduce the effective capacity of the quill as well.

In contrast, if the substrate were rather hydrophobic, the tendency would be to deposit less fluid by wicking action. This could also reduce the number of preprints because the deposited spots would smaller and more uniform at an earlier stage in the printing process. Less fluid deposited would also increase the number of spottings possible under these conditions. Printing is an empirical process, one in which the pin, buffer, and substrate properties must be matched to produce the desired microarray (see Figure 4.1).

We cannot overemphasize the importance of the buffer composition. A case in point is illustrated in Figure 4.19. Two pins (A and B) previously filled with Buffer A are now filled with Buffer B. While the initial printing zone characteristics of these pins remain essentially the same, the number of spots produced by both pins is dramatically curtailed. How could this happen? Buffer B was comprised of a rather volatile component while Buffer

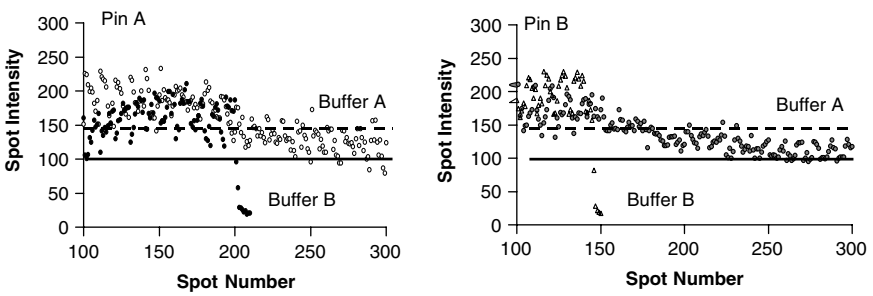


Figure 4.19 Buffer effects on pin performance.

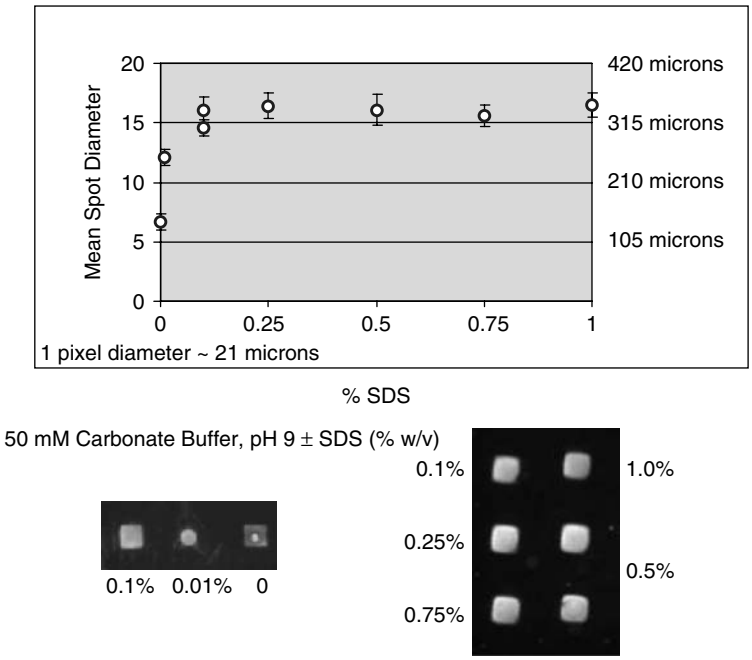


Figure 4.20 Detergent effects on spotting.

A was nonvolatile. The pins dipped in Buffer B simply dried out and left salt deposits to block the tip of the pin, thus preventing the pin from printing.

In another example shown in Figure 4.20, a detergent sodium dodecyl sulfate (SDS) is used to control spot diameters of oligonucleotide probes immobilized onto a plastic (hydrophobic) surface. At zero SDS content, a very small spot was produced (~100-μ diameter) even though the same quill pin on a glass surface was reported by the manufacturer to produce a spot diameter approaching 300 to 330 μ (TeleChem Inc., ChipMaker CMP-10). SDS wetting action overcomes the surface tension on the substrate, allowing the probe to spread out from the tip of the quill, eventually filling the area around the flat tip, leaving a characteristic rounded square spot pattern. As expected, printing of these oligonucleotide probes in SDS buffer on a glass surface results in the deposits of very large spots by the CMP-10 quill.

Microarray design

Designing an array is not a trivial task. In addition to the probes of interest, an array should include appropriate numbers of positive and negative control elements, such as housekeeping genes and controls that can be used to monitor the efficiency of important steps within the process. For example, you may wish to spike in internal standards that track recovery or labeling efficiency among different samples. It is also important to consider how you will print. How many replicates do you want? Should these replicates be

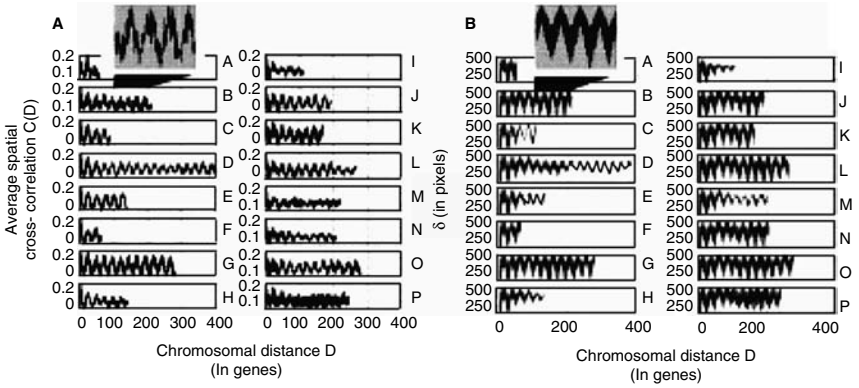


Figure 4.21 Gene expression periodicity effects due to pin-tip bias. (From Balázsi, G. et al., *Nucleic Acid Res*., 31(15), 4425–4433, 2003. With permission.)

placed next to each other or randomly distributed throughout the array? If there is concern about print uniformity, replicates should be distributed across the array.

The pin configuration and print order may also produce impacts. Balázsi et al. (2003) have shown that the printing process can lead to significant biasing of gene expression data, i.e., how you arrange the pins in the print head and the probes in the source plate can affect the outcome of your experiment! Balázsi and co-workers examined microarray gene expression data collected on the growth cycle for *Saccharomyces cerevisiae* from the work of Spellman et al. (1998) at Stanford University. cDNA probes were arranged in 64 (96-well) microtiter plates according to chromosomal order (centromer to left telomere, then right telomere). The array had been printed using four pins in a commonly employed 2×2 pattern yielding blocks of 44×44 spots or 1936 probes printed per probe.

Each pin can be tracked through the printing process and the genes grouped according to print position. When the corresponding gene expression values were arranged according to chromosomal distance, a spatial periodicity was observed: a 2-gene period superimposed on a 24-gene periodicity (Figure 4.21). This periodicity was traced back to tip-specific biases introduced by the 2×2 print head configuration (Balázsi et al., 2003; Yang et al., 2002). A new 176-gene periodicity appeared as a result of applying algorithms to filter out the above periodicity from the gene expression data. The 176 periodicity was attributed to print location bias, i.e., 176 spots (or 4 pins printing 44 spots per print cycle) were printed before the printer returned to a given position on the slide (Figure 4.22). Corrections for such systematic error due to printing bias were shown to improve the gene clustering (Figure 4.23) for the printed arrays with the average minimum distances between functional classes decreased from 13.3369 to 12.4067.

The important point is to be aware of the performance characteristics of individual printing pins (or ink-jets) and how multiple pins match up.

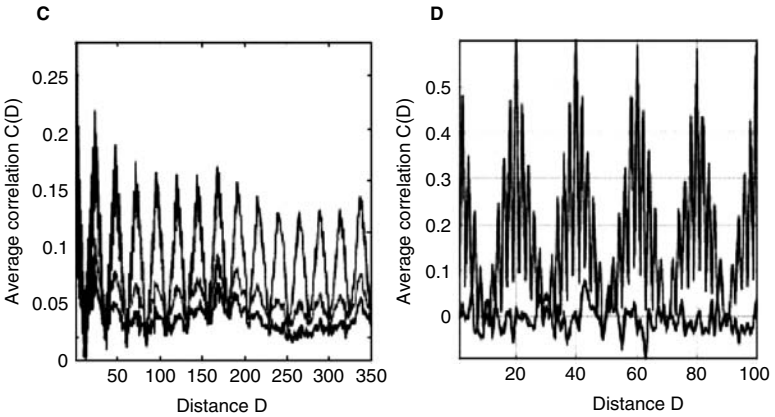


Figure 4.22 Higher order periodicity caused by print location. (From Balázsi, G. et al., *Nucleic Acid Res.*, 31(15), 4425–4433, 2003. With permission.)

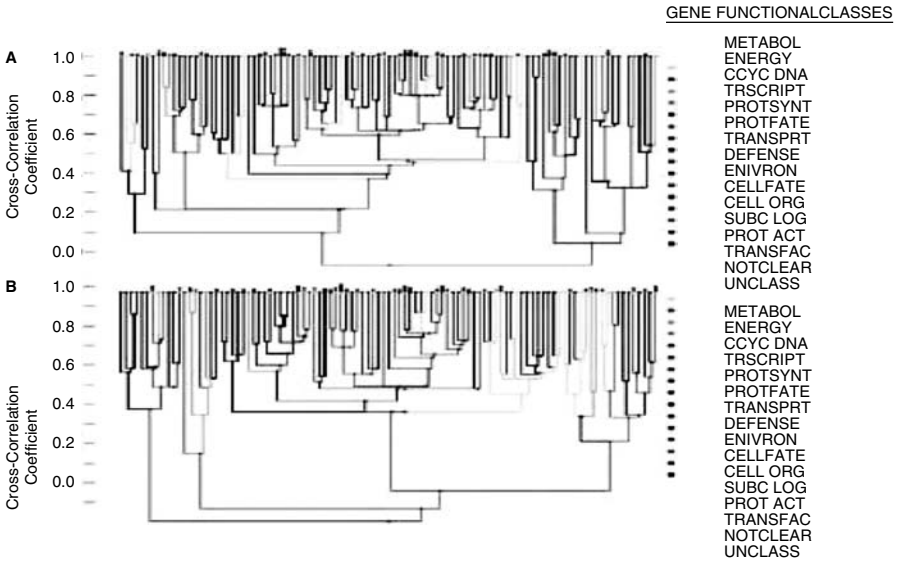


Figure 4.23 Corrections to gene clustering for pin printing bias. (From Balázsi, G. et al., *Nucleic Acid Res.*, 31(15), 4425–4433, 2003. With permission.)

Reinking of pins can lead to systematic differences in spot diameters or spot intensity, dividing the array into regions. This can lead to periodicity effects (Figure 4.24). Individual pins may also exhibit different printing patterns and deliver slightly different volumes or spot diameters. If significant differences (e.g., based upon signal mean ± standard deviation) appear in results obtained from the same probe pairs printed by different pins, then the data must be adjusted to normalize these differences.

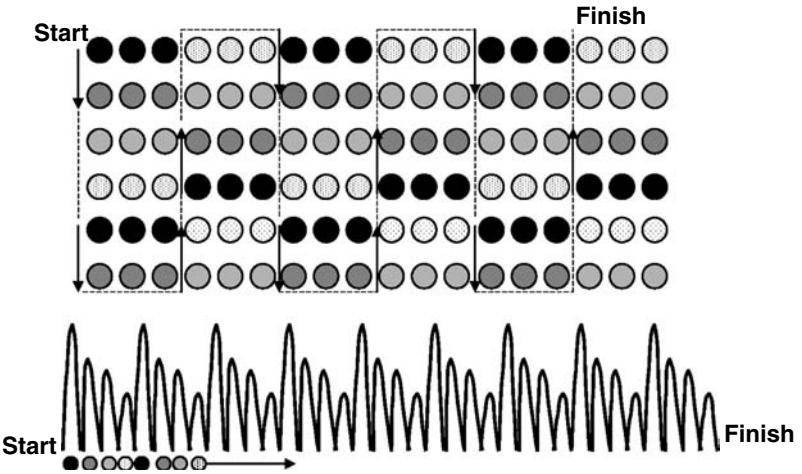


Figure 4.24 Periodicity effects due to reinking.

Setting up a print run

Assume that you are now ready to create your first array on a previously selected substrate. How many elements (spots) do you wish to print? This number will determine what kind of pin you will need; and while it appears to be a rather fundamental question to ask, it may not be simple to answer. As noted, spot density is directly related to spot size and pitch (Figure 4.25). The pitch will determine how many spots can actually be printed on a slide (Figure 4.26). The pin will deliver a specific droplet volume that will spread to a certain diameter largely based upon the tip's diameter and the print buffer used (Figure 4.27). The larger the spot diameter, the fewer the spots that can be printed (Figure 4.28).

The selected pin will only fit into that manufacturer's print head; pins and print heads of different manufacturers cannot be interchanged without hardware modifications. Obviously, the number of pins and their configuration in the print head are determined by the print head design and this can also vary among manufacturers. Pins and print heads are expensive so choose wisely!

After you obtain the correct pin and print head, you may wish to consider the number of pins and their arrangement (configuration) in the print head in terms of throughput. If time is limited, the more pins are configured, the faster the printing will be accomplished. However, also consider that printing with a larger number of pins will require more duplicate probe source wells in the source plate for that purpose. This may also complicate the preparation of the source plate. At that point, it is desirable to enlist the aid of a robotic liquid handler to fill the source plate from the probe stocks. This will reduce errors in filling and save considerable time by allowing the preparation of many source plates at once.

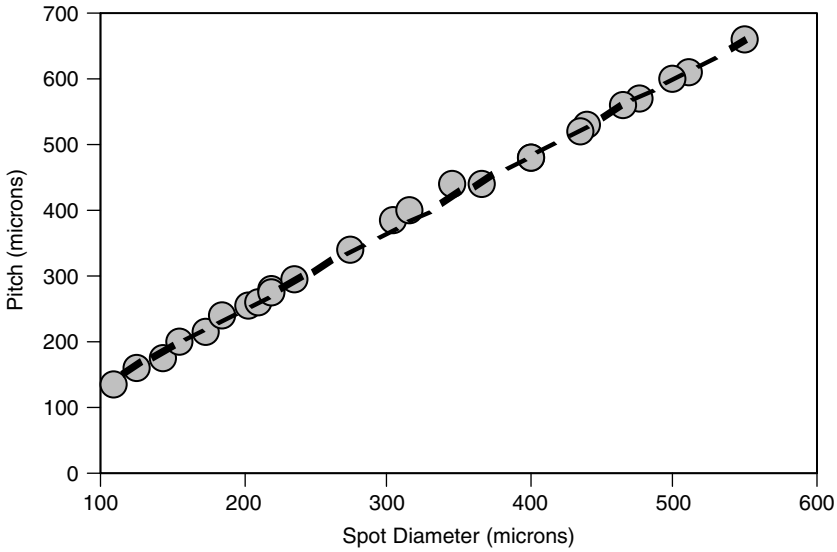


Figure 4.25 Relationship of spot size and pitch. (From Table 1, ChipMaker Micro Spotting Pin Matrix, http://arrayit.com/Products/Printing/Chipmaker/chipmaker.html)

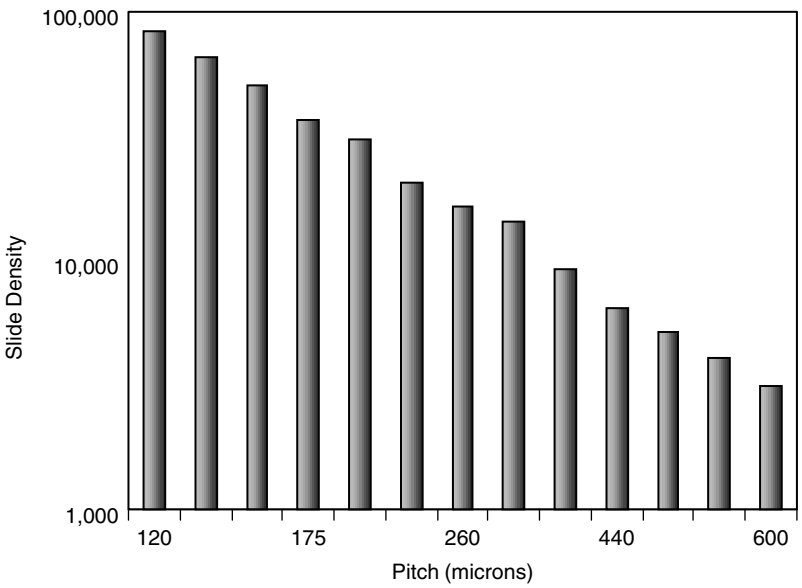


Figure 4.26 Spot density vs. pitch. (From Table 1, ChipMaker Micro Spotting Pin Matrix, http://arrayit.com/Products/Printing/Chipmaker/chipmaker.html)

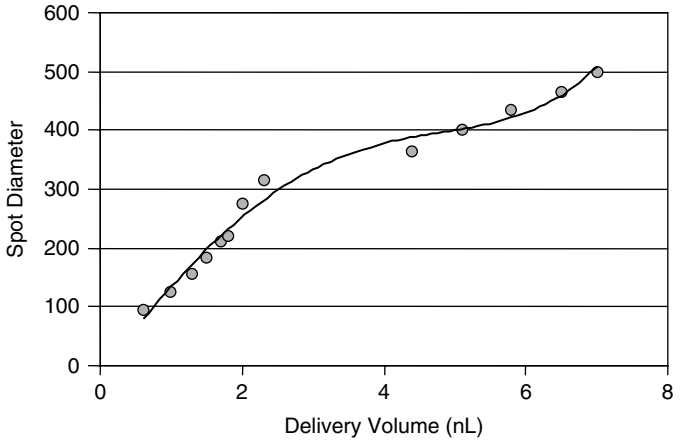


Figure 4.27 Pin delivery volume vs. spot diameter. (From Table 1, ChipMaker Micro Spotting Pin Matrix, http://arrayit.com/Products/Printing/Chipmaker/chipmaker.html)

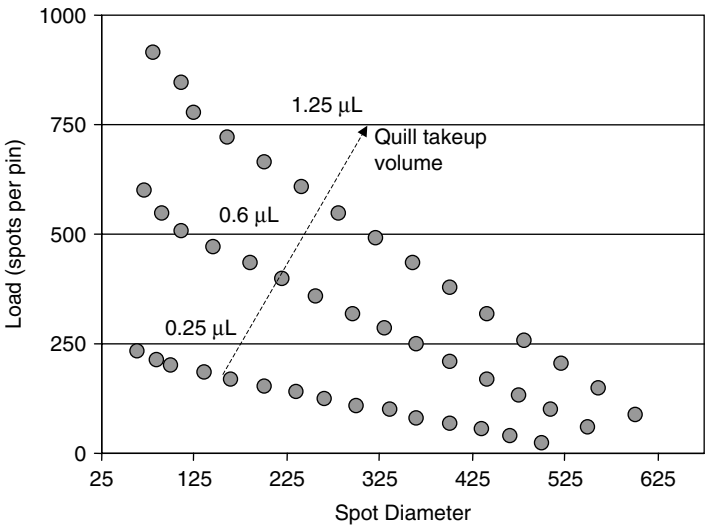


Figure 4.28 Relationship of pin load, delivery, and spot diameter. (From Table 1, ChipMaker Micro Spotting Pin Matrix, http://arrayit.com/Products/Printing/Chipmaker/chipmaker.html)

Printing parameters

A number of printer-defined parameters must be considered in setting up a print run (Figure 4.29). They are briefly described in the following discussion.

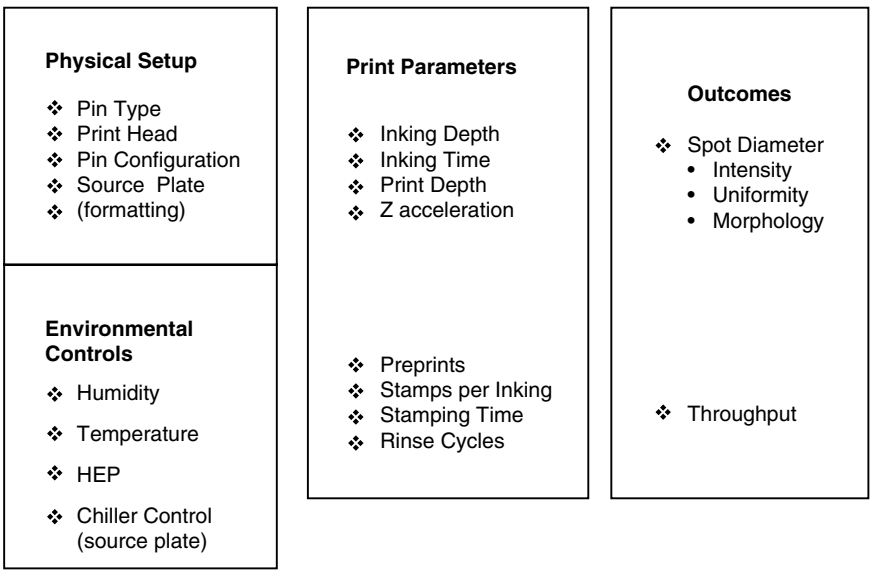


Figure 4.29 Printer setup parameters.

Inking Depth — Depending upon the quill (or solid pin) type selected, certain precautions in filling must be observed. Generally, quill pins should not be dipped too far into the source plate well because the printing ink will deposit on the outside of the quill. This can produce two effects. First, surface wetting means that the quill will eventually drain off the excess upon first striking the substrate, leading to very large spots of varying spot diameters. Thus, more preprints will be required to achieve uniform spot diameters. The other issue with improper depth filling is that capillary action can be slowed or prevented by fluid coming in from the side. Removal of the quill from the well may allow formation of an air gap that essentially blocks delivery of the fluid from the tip. Quill manufacturers can provide the recommended filling depth data for their pins.

Inking Time — Pins may have different capillary fill and wetting rates depending upon surface characteristics and geometries. For example, inking times on quills are in the range of a few seconds, while it may not be necessary to keep solid pins in contact with a source plate for more than a second to allow uptake.

Print Depth — This is a critical parameter affecting pin performance and lifetime. A quill pin striking a hard surface (such as a glass slide) has a limited lifetime. Manufacturers rate pin lifetimes according to the number of spottings. However, misuse can dramatically reduce this number. One particular risk is overdriving the pin. Such excessive force at the tip of a pin will cause damage and reduce performance, leading to poor quality array production. Avoid overdriving by adjusting print depth. This can be accomplished by adjusting the Z height of the print head assembly and in some

cases the deck height of the substrate. Essentially, the pin should be seen to barely float up upon striking the surface. For example, TeleChem quills are generally held to a print depth of about 50 μ relative to the slide surface. If your arrayer also has Z acceleration control, the striking process can be slowed down by reducing the final acceleration of the print head just prior to striking the surface.

Preprints — Quill pins for the reasons described above require a number of preprints. This is largely an empirical exercise. For that reason, the best approach is to conduct a print test and measure the number of printings necessary to achieve consistent spot size. For example, 10 spottings may be required before consistent spot size is reached. Therefore, the number of preprints required to assure high quality array production would be about 15.

Stamps per Inking — This measurement takes into account how many printings can be obtained before print quality deteriorates. It also determines how many slides per inking can be printed. The number of acceptable spots printed upon a single inking is limited. For example, spot quality or diameter may begin to fall around 80 prints. Thus the useful range for printing with a single dip into a source plate would be about 75 prints. If 15 preprints were required, then a pin with a single inking could effectively deliver a single spot to each of 60 slides. The more time involved in reinking, the longer the print run will take because most of the time for printing is spent on rinsing the pin and replacing the ink in the quill.

Stamping Time — This is the amount of time that the pin resides on the substrate surface. The more time spent on the surface, the greater ink volume deposited on the substrate. The greater volume has a tendency to cause spreading of the fluid, thereby increasing spot diameter. However, other factors such as the contact angle of the substrate and the capillary hydrostatic head also influence the size and spread of droplets.

Z Acceleration — Certain printing mechanisms involve striking the surface with the quill pin to dispense a droplet onto a substrate. Control of the acceleration rate can be useful in ejecting the droplet without crashing the pin into the substrate (sometimes called overdriving; Figure 4.30). Not all arrayers have this feature.

Finally, when setting up a print routine for the first time, it is advisable to use dummy or break-away pins (Figure 4.31). That way, in the event of a programming error, you have reduced your liability in replacing damaged pins. Certain arrayers also offer features such as running in slow-motion mode, allowing the user to more easily follow in real time the printing steps such as pin positioning within the reservoir well or stamping depth.

Preparing probe ink

There are many approaches by which to prepare a probe for immobilization to a substrate by printing. The exact nature of the buffer (ink) composition will depend upon the surface characteristics of the substrate and the surface

Figure 4.30 Pin damage due to excessive overdriving into substrate.

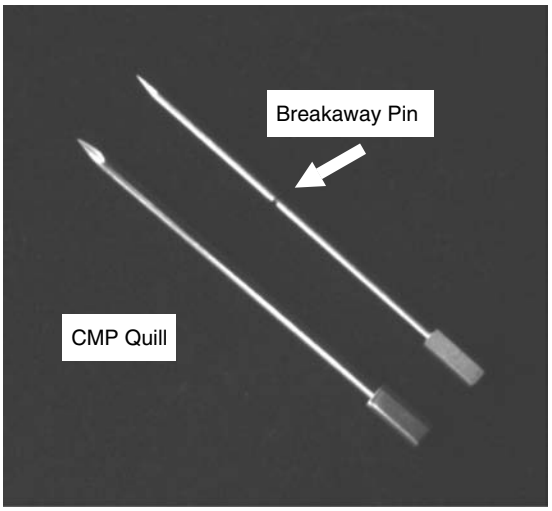


Figure 4.31 Breakaway "dummy" pins.

chemistry employed for immobilization. However, some fundamental precautions should be followed. First, the print buffer must be compatible with the probe. Certain buffers and salts will cause precipitation or probe aggregation. It is highly recommended that the physical and chemical stability of the probe be well understood prior to printing. All printing solutions should be filtered to remove aggregates or debris that might occlude the capillary

tip. Stock solutions should be prepared and aliquots frozen and stored until needed.

Source plates may be prepared in advance and stored frozen, depending upon the probe ink stability. In any case, source plates should be brought to ambient temperature to assure that all components have redissolved. We recommend that source plates be centrifuged to remove any entrapped air bubbles from the bottoms of the wells. While manual preparation of source plates is possible, the use of a robotic dispensing system is highly recommended in order to avoid mistakes in placement of the probe inks in the wells and to avoid cross-contamination. The use of such a device will also allow more uniform preparation of subsequent source plates.

Optimization of probe concentration

What concentration of probe should be used for printing the array? While one can theoretically estimate the monolayer surface coverage for a particular biomolecule, this question is best answered by empirical determination. Practical reasons exist for performing probe loading vs. hybridization efficiency. First, probe stocks may exhibit differences due to variation in their production. It is well known that the synthesis and purification of oligonucleotide probes can vary considerably among vendors or even within lots from the same vendor. It is not uncommon to obtain probes with varying amounts of salts or other materials remaining after high performance liquid chromatography (HPLC) purification. As a result, differences in the chemical compositions of various probe stocks can lead to significant differences in immobilization efficiency.

Second, certain secondary structures within oligonucleotides can lead to concatenation and aggregation, thereby reducing the coupling efficiency to the surface. Proteins can also undergo aggregation or become denatured by adsorption to the surface, thereby reducing overall binding efficiency. Here are some general starting ranges for optimized loading. Obviously, these figures may vary based on molecular weights and size of biomolecules.

Oligonucleotides	5 μM to 40 μM
cDNA	1 nM to 1 μM
Proteins	0.1 to 1 mg/mL

Protocols for printing nucleic acids

cDNA microarray

Slide-based microarray technology was first introduced by Schena et al. (1995). The processes and equipment for preparing (arrayer) and analyzing (laser scanner) microarray slides comprised a portion of Dari Shalon's thesis work at Stanford University. Polymerase chain reaction (PCR) products (cDNA probes) were attached to PLL-coated glass microscope slides. The

mechanism of attachment most likely involves electrostatic interaction through adsorption of the negatively charged nucleic acid to the positively charged lysine residues coating the glass substrate. The probes (48 cDNAs) were laid out in duplicate in an array format using a custom-built arrayer equipped with a single quill pin. The following is a summary of that early printing protocol:

1. Load quill tip with 1 µL PCR product (cDNA probe) from a 96-well microtiter plate.
2. Print probe onto 40 slides depositing ~5 nL per slide at 500-µ center-to-center spacing.
3. Rehydrate slides in a humid chamber for 2 hr.
4. Snap dry for 1 min at 100°C.
5. Rinse slides with 0.1% SDS.
6. Block (cap) residual lysine residues with a carboxylic acid by acylation using succinic anhydride prepared in 1-methyl-2-pyrrolidinone, boric acid buffer.
7. Just prior to hybridization, heat denature cDNA probes by a 2-min soak in distilled water held at 90°C.

The fundamental processes described above have not changed significantly. In fact, many researchers continue to follow this simple protocol, and a few introduced slight modifications for improvement in spotting consistency and/or reduced background (Hedge et al., 2000; Diehl et al., 2001; Hessner et al., 2003b). Aminosilane-coated slides (such as Corning's CMT-GAPS™ slides) are reported to offer more uniform surface coatings and reductions in fluorescent background over PLL slides. Problematic spot morphologies such as the occurrence of donuts (also known as ring spots) were reported to be substantially reduced using aminosilane surfaces (Hedge et al., 2000). While the earlier printing inks employed high salts such as 3X SSC (saline sodium citrate) for depositing spots of cDNA, issues regarding spot homogeneity even on the aminosilane surface remain.

Oligonucleotides

The printing of single-stranded oligonucleotide probes is similar to the process for the printing of cDNA. The input concentration to achieve optimal surface loading is usually 10-fold higher. In our laboratory, we typically print cDNA (>300 bp to ~1000 bp) at 1 nM and 5′-amino-oligonucleotides (15 to 30 mer) at 20 µM. However, we have found that certain oligonucleotides are optimal at higher or lower concentrations and recommend performing a loading study at 5 to 40 µM in most instances. For synthetic oligonucleotides, we highly recommend using HPLC purified and desalted stocks, especially for amino-oligonucleotides.

Print buffer composition should remain simple. For most printing applications, a sodium phosphate buffer at pH ~8 to 9 works well. Our laboratory

employs a carbonate buffer system (50 to 150 mM sodium carbonate-bicarbonate, pH 9) that works very well for the immobilization of oligonucleotides or proteins to activated supports. We can easily change spot size by the addition of SDS to this buffer (0 to 0.25%).

For aldehyde slides, Genetix recommends printing oligonucleotides at 10 to 50 pmole/μL (10 to 50 μM) and cDNAs at 0.2 to 1 μg/μL (~0.3 to 1.5 μM based upon 1 μg of 1,000 bp cDNA = 1.52 pmoles). For aminosilane slides, AstroGen Biosciences recommends resuspending DNA to a maximum of 0.25 mg/mL (~ 0.4 μM) in phosphate buffer (150 mM, pH 8.5). For the printing of oligonucleotides onto polymer-coated PowerMatrix™ slides, the manufacturer recommends a final concentration of 0.5 to 1.0 μg/mL in 150 mM sodium phosphate buffer, pH 8.5 to 9.5 (see *User Guide* published by Full Moon BioSystems, Sunnyvale, CA).

Accelr8 Technology offers a polymer-coated glass slide product under the trade name of OptiPlate™-DNA. The company provides a printing protocol for amino-oligonucleotides that differentiates long and short print runs. For short runs, the company recommends 50% relative humidity for contact printing in a buffer composed of 20 μM oligonucleotide in 300 mM sodium phosphate, pH 8.5, 0.005% Tween-20, and 0.001% sarcosyl. For long runs, the suggestion is 30% relative humidity in a print buffer comprising 150 mM sodium phosphate, pH 8.5, and 0.001% Tween-20. Sarcosyl is not included in this buffer. Also, this manufacturer does not recommend the use of dimethyl sulfoxide (DMSO) with its slides.

U-Vision Biotech (www.u-vision-biotech.com) markets an epoxy-activated slide product under the trade name EasySpot® Oligo for the immobilization of oligonucleotides (20-mer to 70-mer). Unmodified oligos, PCR products, and RNA are preferred over amine-modified forms. The company states, "Our experiments showed that the amine modification slightly lowers the attachment efficiency of oligonucleotides." Its protocol suggests resuspending DNA from to 2 μM approximately 16 μM in 50% DMSO–distilled H_20 only and caution against using salts such as SSC and sodium bicarbonate. The presence of salts on this slide chemistry affects the spot morphology and the efficiency of immobilization.

Longer oligonucleotides (50-mer to 70-mer) have been employed recently. Kane et al. (2000) described the covalent attachment of 5′-amino 50-mer oligonucleotides to 3D-Link slides (Surmodics, Eden Prairie, MN). The 50 mers were prepared at 20 μM in 150 mM sodium phosphate buffer, pH 8.5. The Gene Arrays URL (gene-arrays@itssrv1.ucsf.edu) contains considerable discussion about the printing of longer oligonucleotides onto different substrates. For example:

70-mer (operon), 40 μM, 50% DMSO, Corning UltraGAPS: "You can go down to 20 μM but the spots look a lot better with 40." Gregory Khitrov, July 31, 2003.

50-mer (operon), 10 μM, 150 mM sodium phosphate, pH 8.5, 0.01% SDS, epoxy slides: "We also found 40 μM to be excessive. I did a dilution series

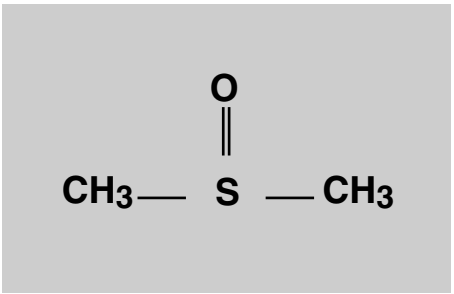


Figure 4.32 Structure of dimethyl sulfoxide.

and could see no change in signal until I reduced the oligo concentration below 5 µM. So we print oligos at 10 µM." Patty Holman, July 30, 2003.

Most likely, both observations are valid, depending upon substrate, probe, and printing conditions.

Dimethyl sulfoxide

DMSO (Figure 4.32) was introduced as an additive with demonstrated improvements in spot morphology and an increase in hybridization efficiency (Hedge et al., 2000). This solvent serves as a denaturant for nucleic acids and presumably permits more efficient tethering of single stranded probes. Its hygroscopic property is responsible for slowing of evaporation of printed spots as well as probes remaining in the source plate. Hedge et al. (2000) examined spot morphology from cDNA spotted down on aminosilane slides in 3X SSC containing 50% DMSO as a function of relative humidity (RH) and temperature and concluded that optimal conditions were obtained at 22.2°C and 45% RH.

Betaine

While Diehl et al. (2001) agree that the addition of DMSO to print buffer improves spot uniformity, they argue that DMSO is also toxic and a good solvent for other materials. As a result, they explored alternative chemistries to replace DMSO and also to improve upon postprint blocking conditions in an effort to find a replacement for borate-NMP (1-methyl-2-pyrrolidinone) buffer used for preparing solutions of succinic anhydride for capping of residual amine groups.

The Heidelberg group (Diehl et al., 2001) chose (carboxymethyl) trimethylammonium hydroxide, commonly known as betaine (Figure 4.33) as a substitute for DMSO. Why these researchers selected betaine and did not consider other additives was not discussed. However, we do know from other works (see cited references) that this compound is effective in reducing stability differences between A:T and G:C base pairing during hybridization, much like the action of tetramethylammonium chloride or formamide (Rees et al., 1993). Thus, betaine also provides a simple means to denature DNA and maintain such probes in a single-stranded state.

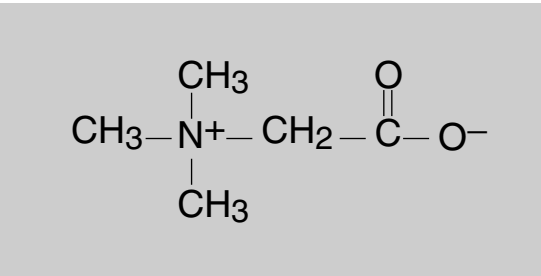


Figure 4.33 Structure of betaine.

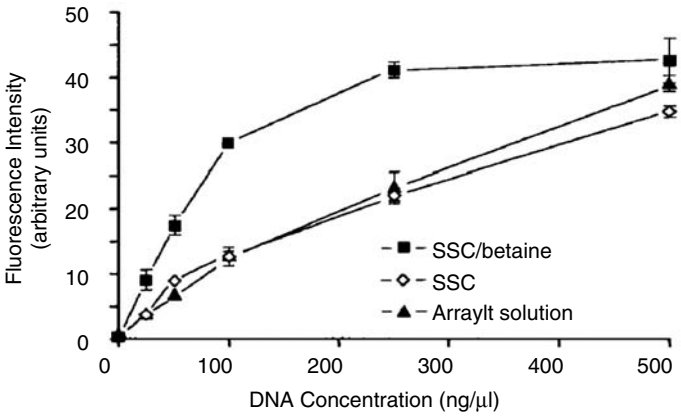


Figure 4.34 Efficiency in the delivery of DNA onto a substrate using various additives. (From Diehl, F. et al., *Nucleic Acid Res.*, 29(7), 1–5, 2001. With permission.)

Single-stranded probes are preferred for attaching the probes to a support. In addition, higher concentrations of betaine are rather viscous and slow evaporation. Thus, betaine would most likely improve spot size and uniformity by slowing evaporation and thereby impeding the spread of the DNA spot out into the familiar donut shape. It also would slow evaporation in the wells of the source plate, preventing the unwanted and nonuniform plate edge evaporation effect on probe concentration.

Thus, 3X SSC ± 1.5 M betaine was used to print down a 500-bp cDNA onto both PLL and aminosilane glass slides and a commercial print buffer, ArrayIt™ microspotting solution or MSS (TeleChem International, Inc., Sunnyvale, CA) was compared. Curiously, 3X SSC ± 50% DMSO was not included in this study. While MSS may contain DMSO, the experiments would have been better designed by including 3X SSC + 50% DMSO as a control.

Nevertheless, it is clear from the work that the inclusion of betaine into the print buffer was an improvement over SSC or the MSS on several fronts. First, the SSC–betaine spotting was found to increase hybridization efficiency as measured by a 2.5-fold higher hybridization signal intensity for probe–target hybrids relative to those probes spotted in SSC or MSS alone (Figure 4.34).

Spot morphology was thought to have improved based upon the observed levels of signal intensity variation within spots. The addition of betaine reduced the spot pixel intensity CVs from 14% down to 5% when DNA was spotted at 100 ng/mL. Finally, betaine contributed the lowest on-spot (buffer residual) background relative to SSC or MSS. In summary, the inclusion of betaine in the print buffer improved printing.

The other issue addressed by Diehl et al. (2001) is more refined control of the capping process to prevent the occurrence of an elevated interspot background. The dissolution of various print buffer components back into solution can lead to comet-tailing effects as well as readsorption back onto the slide, resulting in a build-up of the background. Slides may appear blotchy and exhibit uneven signal distribution. In other instances, a decrease in spot uniformity or a significant reduction in signal intensity may become evident. The most severe result occurs when the interspot background intensities are much greater than the spot intensities, leading to the appearance of "black" spots surrounded by bright interspot backgrounds.

The authors reasoned that the aqueous succinic anhydride capping buffer (comprised of 96% NMP and 4% sodium borate) may have led to the redissolving of probe DNA that was subsequently randomly redeposited over the entire slide, leading to elevated background. As a result, a reformulation of succinic anhydride into a nonaqueous medium of dichloroethane (DCE) solvent containing N-methylimidazol (acylation catalyst) was undertaken. Significant improvements in interspot backgrounds were evident.

Evaporation

McQuain et al. (2003) undertook a detailed study on the effects of relative humidity and a direct comparison of the impacts DMSO vs. betaine in print buffer on the overall performance of quill pin printing. A video microscope was employed to visualize and track the drying behaviors of the various printing inks. A Cy5-labeled 466-bp dsDNA probe was used to monitor the printing process. Drop-drying behavior, bulk evaporation from the quill reservoir, surface tension changes, and spotting characteristics (spot diameter, spread, and number deposited) were examined at different RH levels.

Print buffers 3X SSC, 3X SSC + 50% DMSO, and 3X SSC + 1.5 M betaine were evaluated at 40, 60, and 80% RH for spot intensity, spot diameter, intraspot variation, and CV (Figure 4.35). The reductions in quill drop volumes and droplet drying times were measured by video microscope and the quill reservoir volume changes determined by weight. In summary, "Solvent evaporation from the print buffer reservoir is the major factor responsible for the variations in the transfer of fluid to the slide surface."

In order to control spot deposition and ultimately spot diameter and morphology, one must first control the rate of evaporation from the quill reservoir. The change in surface tension causes variations in spot characteristics. Evaporation of water from the bulk solution held in the quill reservoir increases the salt and probe concentrations, which in turn increases the

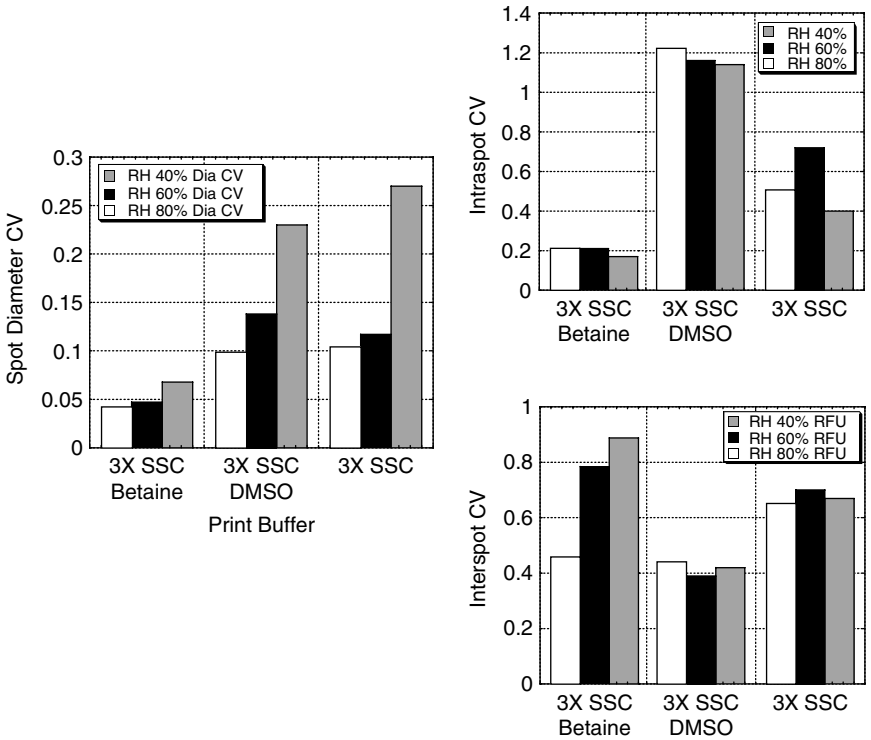


Figure 4.35 Spot variation upon printing in various buffers at different relative humidities. (From McQuain, M.K. et al., *Anal. Biochem.*, 320, 281–291, 2003. With permission.)

surface tension of the solution. As a result, less fluid volume is transferred from the pin, and the smaller droplets transferred to the substrate surface are at a higher surface tension. This results in a higher surface contact angle.

Opposing surface tension force is the pinning force. Essentially, pinning forces promoted by surface features tend to fix the contact line of the droplet and drive the DNA toward the contact line (solvent perimeter). Solutes such as salts and probes spread to the perimeter by convection as the droplet evaporates (Figure 4.36). Uneven evaporation leads to differences in spot uniformity.

How effective were DMSO and betaine at improving print performance? It really depends upon what is most important. Based on the work of McQuain et al. (2003), Table 4.2 summarizes the rankings based upon variations in inter- and intraspot signal intensity and spot diameter. The addition of DMSO provided the best spot-to-spot signal reproducibility, while betaine appeared to produce the most consistent spot diameters and homogeneity. Interestingly, betaine produced spots with greater signal intensity variations than either 3X SSC + DMSO or 3X SSC alone while intraspot CVs were the highest when DMSO was included in the buffer. If you are most concerned about spot-to-spot

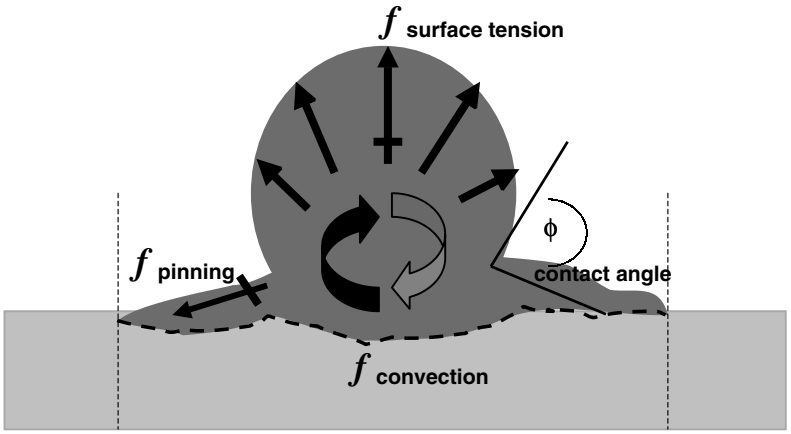


Figure 4.36 Droplet drying mechanisms.

Table 4.2 Ranking Print Buffer Additives

Buffer composition	Variation Rankings		
	Interspot (spot to spot)	Intraspot (within spot)	Spot diameter
3X SSC	2	2	3
3X SSC + DMSO	1	3	2
3X SSC + Betaine	3	1	1

Note: Rankings-based calculated CVs from reported mean intensity, standard deviation 1 = lowest percent CV (least variation). SSC = saline sodium citrate. DMSO = Dimethyl sulfoxide.

Source: From McQuain, M.K. et al., *Anal. Biochem.*, 320, 281–291, 2003. With permission.

variations, use 3X SSC + 50% DMSO; if spot homogeneity is more important, betaine may be a better choice as a print buffer additive.

Of course, one must ask whether a combination of DMSO and betaine would be of even greater benefit for spot quality. Hessner et al. (2003a), in efforts to develop a tracking system to measure microarray performance, found that 1.5 M betaine in 3% DMSO provided the greatest probe retention on PLL slides. Their major concern was that when cDNA probes were printed at low concentration (<100 ng/μL), one could not accurately distinguish expression differences.

With a reduced probe population on the surface, it would be possible to saturate the probe–target binding sites (i.e., target excess). As Hessner et al. note, under such a condition, the transcript ratios would be compressed. To achieve quantitative information from microarrays (determining more than the presence of a target), it is desirable that probes remain in "binding" excess over targets. The problem was that when probes were printed down on glass slides, only a portion of the cDNA remained attached to the surface (or

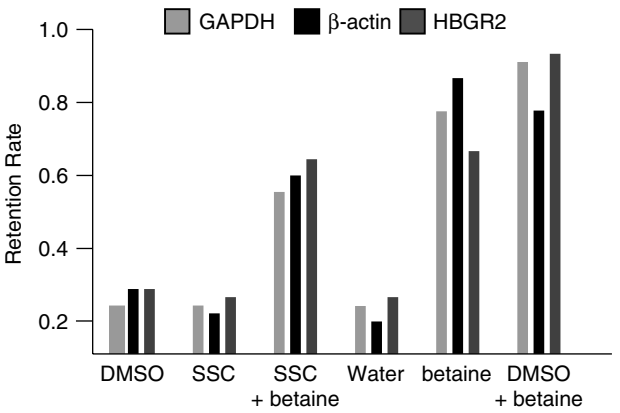


Figure 4.37 cDNA retention on glass subtrates after printing in various buffers. (From Hessner, M.J. et al., *Nucleic Acid Res.*, 31(4), 1–6, 2003a. With permission.)

remained available for hybridization) following processing. This was found to be print buffer-dependent (Figure 4.37).

Both 50% DMSO and 3X SSC were about as effective as water as a print buffer in terms of cDNA probe retention (~20 to 30%), while 3X SSC + 1.5 M betaine retained ~60 to 70% of hybridizable probes on the surface. Interestingly, betaine alone appeared to provide better retention than in combination with 3X SSC although the data scatter permits us only to suggest a trend. However, the combination of DMSO and betaine also provided a level of retention comparable to that of betaine (~70 to +100%) with the added benefit of being able to titrate with DMSO to control spot diameter. Thus, Hessner et al. (2003a) determined that 1.5 M betaine in 3% DMSO provided the optimal print buffer for their studies involving immobilization of cDNA probes onto PLL slides.

Print quality assessment

One important outcome of the Hessner et al. studies was development of a method useful for determining microarray spot quality (2003a and 2003b). One of the drawbacks to using microarrays has been how to address quality assurance and quality control (QA/QC) issues regarding the printed product prior to use. After an array is printed, how do you best determine spotting consistency from slide to slide or batch to batch?

Several approaches such as Syber Green II staining (Battaglia et al., 2000) and the hybridization of fluorescently labeled sets of randomly synthesized short oligonucleotides (Panomer™ 9 Molecular Probes) have been used. While these methods do work, it is necessary to melt off or dissociate the short oligonucleotides from the support by chemical means. Battaglia's group claimed a reversible staining for Syber Green, but we have found the need for extensive destaining to reduce the fluorescence background to an acceptable level. We

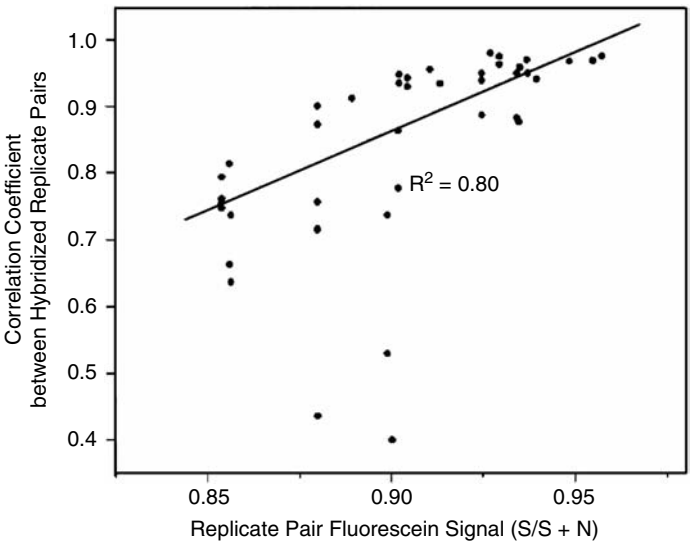


Figure 4.38 Spot quality vs. probe quality. (From Hessner, M.J. et al., *Nucleic Acid Res.*, 31(4), 1–6, 2003a. With permission.)

also noted certain sequence bias in staining with Syber Green, for example, polyA strings (R. Matson, unpublished data). In both cases, the lingering question is the effects such reagents have on microarray performance.

The question is similar to asking how many times one can strip and reuse a microarray before performance deteriorates. An alternative approach is provided by Hessner et al. (2003a) in which the cDNA probes are permanently labeled using fluorescein-labeled primers to the clone's vector insert region. Fluorescein is excited at 488 nm and emits at 508 nm, while Cy3 may be excited at 543 nm to reduce any spectral overlap with fluorescein. Thus, fluorescein-labeled cDNA probes may be printed down and the slide scanned for QC/QA purposes prior to hybridization. Since the same region is primer-labeled in each cDNA, a direct comparison between the relative fluorescence units (RFUs) and the amount of cDNA probe can be determined.

We can now evaluate the microarray slide for a variety of parameters: spot diameter, intensity, morphology, and retention upon processing. Hessner et al. (2003a) examined 50 pairs of slides varying in spot quality and found a significant ($P < 0.001$) difference in hybridization performance based upon the fluorescein probe quality (Figure 4.38). Probes with low signal-to-noise ratios are most likely to produce hybridized arrays having low signal-to-noise values; and such sets of microarrays do not show good inter-slide correlations.

In a further examination of printing behavior on slide performance, Hessner et al. (2003b) looked at the effect of print order on spot quality (Figure 4.39). They noted an obvious trend in that the first arrays printed showed higher overall spot intensity.

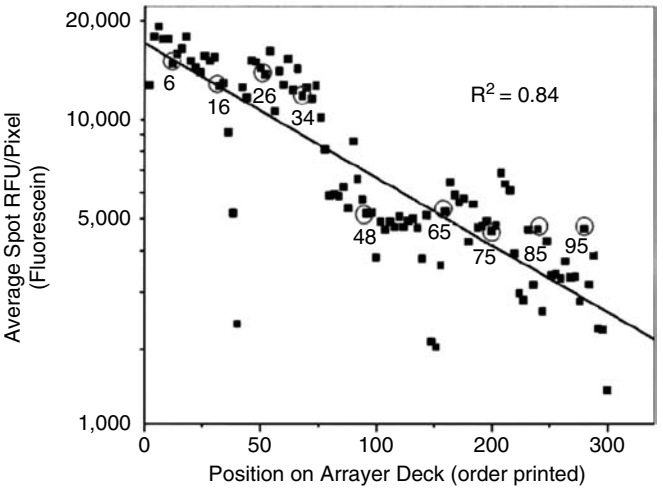


Figure 4.39 Spot quality vs. print order. (From Hessner, M.J. et al., *Nucleic Acid Res.*, 31(11), 1–9, 2003b. With permission.)

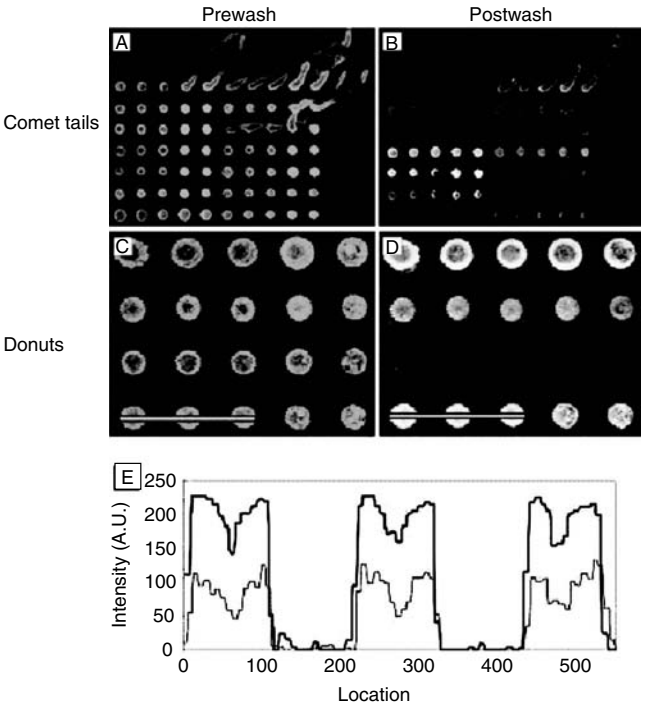


Figure 4.40 Assessment of spot quality using dye-labeled deoxynucleotide triphosphates (dNTPs). (From Shearstone, J.R. et al., *Biotechniques*, 32, 1051–1057, 2002. With permission.)

A spot quality metric called a quality composite score (q_{comp}) derived from the weighting of five spot-related criteria (size, signal-to-noise ratio, background level, background uniformity, and saturation) was developed based upon Matarray software (Wang et al., 2001 and 2003). Using q_{comp} to screen for spot quality on microarrays, it was possible to remove spots with low scores from the analysis. In doing so, Hessner et al. (2003b) concluded that limiting probe concentration leads to greater hybridization signal variability while probes with high fluorescence intensity (higher probe concentration) are less variable. They observed that a fluorescein intensity threshold at 5000 RFU per pixel was necessary for good hybridization performance. Slide acceptance criteria were suggested to achieve reproducible results allowing differential expression change to 1.5-fold:

1. Array mean element intensity >5,000 RFU/pixel
2. Signal intensity CV <10%
3. Mean S/S + N score >0.85 (S = signal intensity; N = noise level)
4. Spot size CV <20%

The works of Hessner et al. (2003b) and Wang et al. (2003) (Max McGee National Research Center for Juvenile Diabetes, Milwaukee, WI) provide a quantitative and systematic approach toward cDNA probe microarray quality assessment. Their utility for the assessment of oligonucleotide arrays is less certain. Oligonucleotide probes labeled with fluorescein may not be as sensitive due to weaker signal strength at lower probe concentration with the added potential for quenching of fluorescence signal at high probe concentration.

A simpler method has been described by Shearstone and coworkers from Biogen (2002). Instead of employing dye-labeled oligonucleotides, they spiked Cy3-deoxycytidine 5′-triphosphate (dCPT) or Cy5-dCTP into the print buffer (containing unlabeled oligonucleotide probes). Spiked Cy3-dCTP was more sensitive than Cy5-dCTP. The Cy3-dCTP could be reliably detected from 10 µM down to at least 20 nM. The latter concentration was chosen in order to prevent carryover. The potential for cross-contamination was recognized due to difficulty in removing the dye from the quills at higher concentrations. It was noted at higher concentrations that the spiked Cy3-dCTP (e.g., 500 nM) may have interfered slightly with oligonucleotide attachment.

The effect was minor at about 87% of signal strength compared to control hybridization to probes without spikes. This may be within the range of assay variation. However, what is particularly attractive about the dye-labeled dNTP route is that it is relatively inexpensive compared to primer labeling. Moreover, spot characterization (e.g., spot morphology, signal intensity, spot diameter) could be easily determined from the scans (Figure 4.40).

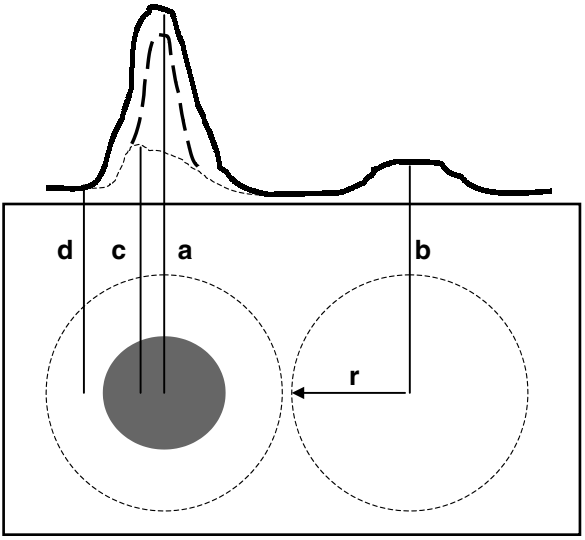


Figure 4.41 Background profiles. a = Array pixel intensity profile. b = Nonspecific background on-spot (false positive). c = Hidden nonspecific background on-spot. d = Off-spot background.

Backgrounds

An understanding of what contributes to background and how best to correct microarray data for background remains an important issue for both DNA and protein microarray analysis. Fundamentally, we can divide background into two basic categories: off-spot contributions (nonspecific signals surrounding spots) and on-spot contributions (nonspecific signals within specific signal regions).

For example, consider a fluorescence pixel intensity scan for a portion of a microarray as described in Figure 4.41. The positive signal (a) is easily distinguished while the much weaker signal at (b) could be identified as either a nonspecific signal such as fluorescence arising from the print buffer or as a true signal, in which case by our example it would be a false-positive signal. Nevertheless, by means of a background subtraction (b – d) and/or establishment of a threshold ratio (a/b > n value), it would be possible to score the signal at (b) as desired in order to improve the data.

Many image analysis software programs examine the pixel intensity histograms for the on-spot and off-spot signals and then define local background (at a radius r from the spot) from which to subtract background. This is an acceptable practice provided that the backgrounds are similar within and outside of the spot. However, this background assumption may not always be valid. As depicted in Figure 4.41, the nonspecific signal (c) could contribute to the overall intensity within the spot. An example of this type of background would be fluorescent residue from components in the print buffer (at spot b) or rinsing solutions. Another approach would be then

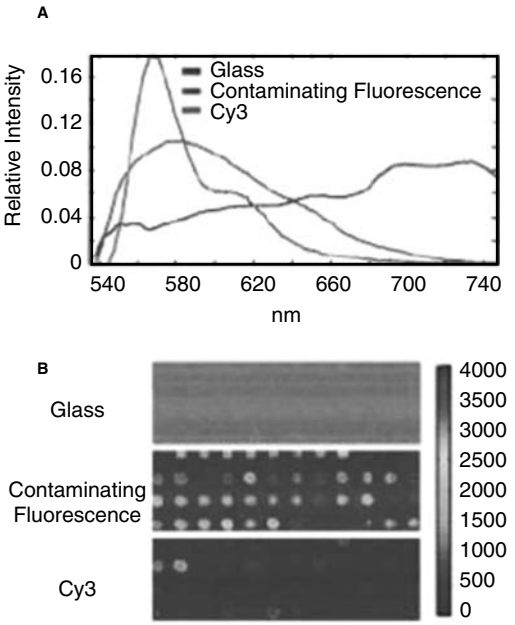


Figure 4.42 Examining microarray backgrounds using hyperspectral scanning. (From Martinez, M.J. et al., *Nucleic Acid Res.*, 31(4), 1–8, 2003. With permission.)

to background subtract (a – b) assuming an equivalent background contribution to the on-spot intensity.

What if the background were primarily associated with the spot and the background varied from spot to spot? Martinez et al. (2003) examined on-spot contaminating fluorescence backgrounds for a number of commercial and in-house printed slides. Using the Axon 4000B scanner and a hyperspectral imaging scanner (developed at Sandia National Laboratories, Albuquerque, NM) capable of looking very closely at intra-spot intensity profiles, significant spot-localized backgrounds were detectable in the green channel in the absence of Cy3 (Figure 4.42).

Such contaminations discovered in mock hybridizations were found to be highly variable with average spot intensities ranging from 840 ± 689 for Corning preprinted CMT and 682 ± 382 for Operon's OpArray preprinted yeast array slides. The printing of 70-mer oligonucleotide probes onto a series of commercial slides also exhibited on-spot backgrounds, especially when the print buffer included TeleChem's MSS where backgrounds were, for example, 3116 ± 1405 on Corning's GAPS I aminosilane slides.

A post-printing treatment involving the following sequential rinsing protocol reduced these backgrounds to ~200 relative intensity units:

1. 0.1% SDS, 10 min
2. 2X SSC, 2 min
3. Boiling water, 3 min
4. Ice cold 100% ethanol, 5 min

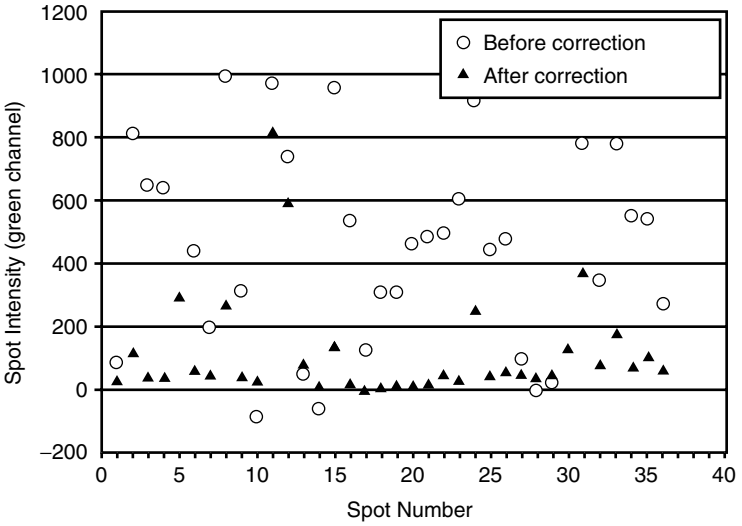


Figure 4.43 Background correction based upon multivariate curve resolution following hyperspectral scans. (From Martinez, M.J. et al., *Nucleic Acid Res.*, 31(4), 1–8, 2003. With permission.)

Aging of certain slides (e.g., Corning GAPS) by exposure to humidified air for several hours prior to printing also reduced the background. It is not known whether this applies to all substrates.

Utilizing a hyperspectral imaging scanner together with MCR (multivariate curve resolution) algorithm analysis made it possible to evaluate the contribution of on-spot backgrounds to errors in the Cy5 and Cy3 gene expression ratios. Martinez et al. reported that for the green channel (Cy3) intensities, ~75% of spots were off by a factor of 2, while 50% were off by a factor of 3; at least 25% of all spots exhibited errors greater than a factor of 4.5 (Figure 4.43). Such variation has an obvious impact on Cy5 and Cy3 expression ratios and upon the interpretation of the data set in terms of the biology studied.

Protocols for printing proteins

The printing of proteins is not difficult as long as the proteins are antibodies. While this is obviously not a completely accurate statement, there is much more danger in generalizing about printing proteins than printing oligonucleotides. While the physical and chemical properties of nucleic acids are very well understood and we have had over a decade to learn how to print them, our knowledge base for creating protein microarrays is essentially derived from work on antibody arrays. While some may disagree, I contend that the antibody array represents the easiest example to follow. The complexity of the proteome will provide ample challenge to those interested in using microarrays.

Antibody arrays

Many protocols are now available for printing antibodies onto various substrates. However, most protocols involve the following fundamental processing steps: (1) exchange of protein from storage buffer into print buffer, (2) adjustment of protein concentration, (3) arraying, (4) postprinting rinse to remove unbound excess protein, and (5) a blocking step. In some cases where a protein is covalently immobilized, a capping step may be required in order to inactivate residual reactive groups. One of the best examples was provided by Haab et al. (2001) in which 94 antibody–antigen pairs were printed onto a PLL slide. This work comes from Pat Brown's lab at Stanford University where the DNA microarray based upon the same slide chemistry was invented.

Protocol of Haab et al. (2001)

1. Transfer antibodies and antigens from glycerol buffer into glycerol-free PBS solution using spin columns (BioSpin P6, Bio-Rad Laboratories, Hercules, CA).
2. Prepare proteins at 0.1 to 0.3 mg/mL.
3. Transfer 4 μL into 384-well source plate.
4. Array solutions from source plate onto poly-L-lysine coated glass slides.
5. Rinse microarrays briefly in 3% nonfat milk in PBS, pH 7.4, containing 0.1% Tween-20. This step is intended to remove unbound protein.
6. Soak slides overnight at 4°C in 3% nonfat milk in PBS, pH 7.4, containing 0.02% sodium azide as a preservative. This step is used to block the slide surface to reduce nonspecific adsorption of analytes or other interfering substances in the sample in subsequent analysis steps.
7. Just prior to use, rinse slides with PBS three times at room temperature (1 min for each rinse). Maintain slides in PBS buffer until incubation with sample.

At the forefront of arraying proteins is a demonstration by MacBeath and Schreiber (2000) in which proteins were immobilized to a glass slide coated with bovine serum albumin (BSA) and subsequently derivatized with NHS (N-hydroxysuccinimide) groups. The BSA-NHS slide served two roles. First, the BSA as a common blocking protein masked out any regions on the slide that might have otherwise contributed to nonspecific binding. Second, the NHS-modified BSA served as a convenient scaffold for covalent immobilization of the protein probe. The adsorbed BSA layer most likely prevented the probe from any surface interactions with the substrate, while keeping it oriented into the surrounding media for efficient capture of analyte.

Protocol of MacBeath and Schreiber (2000)

1. Prepare proteins for spotting in PBS, pH 7.5, containing glycerol (40% v/v) at 100 µg/mL.
2. Array proteins onto NHS-BSA-coated slide.
3. Incubate protein microarrays in a humidified chamber at room temperature.
4. After 3 hr, remove slides and drop them face down into PBS, pH 8.0, containing 500 mM glycine. (Glycine is commonly used as capping reagent, in this case, to inactivate residual NHS esters.)
5. Soak slides for 1 min, then turn them upright and immerse in PBS-glycine. Incubate with gentle agitation for 1 hr at room temperature, followed by a final rinse in PBS prior to use.

In the previous examples, the substrate of choice was the glass microscope slide — a format that allowed these researchers to rapidly move ahead with experiments, largely due to the accessibility of equipment and technologies used in DNA microarray analysis such as arrayers, scanners, and image analysis software. Around the same time, a higher throughput microarray format was under development. This new format was based upon the microtiter plate, a standard format for automated liquid handling and target screening assays.

Genometrix (Mendoza et al., 1999) introduced the use of optically flat glass plates masked with Teflon to create a pattern of 96 wells. The surface was first coated with aminosilane and subsequently reacted with bis-sulfo-succinimidyl suberate to create an NHS surface for covalent attachment of proteins. Various immunoglobulins (creation of an antigen array) were printed down into individual wells in 50 mM carbonate buffer, pH 8.3, at 100 µg/mL using a capillary contact printer to deposit droplets (200 pL) of protein, ~275-µ diameter spots on 300–µ centers. Arrays were first rinsed with Tris-buffered saline (TBS) containing 0.1% Tween-20 to remove unbound protein. Blocking was accomplished using Blocker™ casein in PBS (Pierce Endogen, Rockford, IL) for 1 hr at room temperature. A semiautomated microarray ELISA was performed in a well volume of 25 to 35 µL, depending upon the incubation step, using a HYDRA 96® Liquid Pipettor (Robbins Scientific, Sunnyvale, CA).

This author and coworkers at Beckman Coulter first described the use of a low form 96-well plastic microplate for automated micro-ELISA immunoassays (Matson et al., 2001). The polypropylene plate was first modified by a radiofrequency plasma amination process (Matson et al., 1995) followed by conversion to an acyl fluoride surface chemistry for rapid covalent attachment of biomolecules. Proteins (1 to 2 mg/mL) were prepared in 50 mM carbonate buffer, pH 9, containing 4% sodium sulfate (to improve spot uniformity) and printed using a conventional arrayer system. Approximately 200-pL droplets of monoclonal antibodies (anti-cytokine) were deposited into the bottom of the microwells using a Cartesian PS7200 system equipped

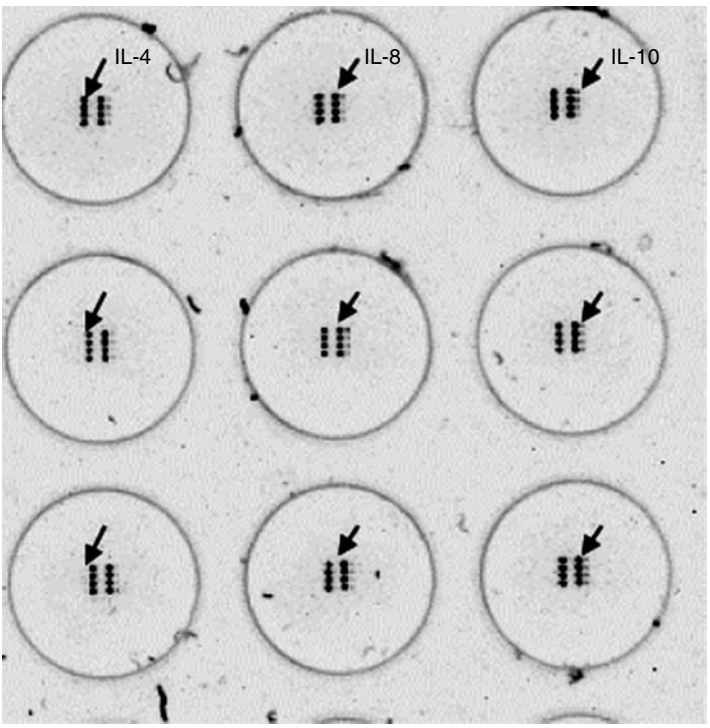


Figure 4.44 Cytokine antibody array.

with Majer Precision quill pins (Figure 4.44). Following a drying step, the wells were soaked in a carbonate–casein buffer to quench residual reactive groups and block the wells in order to reduce nonspecific protein adsorption.

Moody et al. (2001) immobilized anti-cytokine monoclonal antibodies (MAbs) in a 3×3 pattern to the bottom of a Maxisorp™ polystyrene 96-well plate (Nalge Nunc, Rochester, NY) using a Biochip Arrayer (Packard Bioscience, now PerkinElmer, Boston, MA). The MAbs were prepared at 50 µg/mL in Dulbecco's PBS without calcium or magnesium (Life Technologies, now Invitrogen, Carlsbad, CA) and deposited at 20 nL per spot with a resulting spot diameter of about 400 µ. Well microarrays were then blocked with SuperBlock™ buffer (Pierce Endogen) for use in a micro-ELISA. Pierce's SearchLight microplate antibody microarray products are based upon this technology.

Angenendt et al. (2003) evaluated several slide surface chemistries for use as protein and antibody microarrays. They reasoned that because proteins vary greatly in their surface charge and relative hydrophobicity, a careful selection of surface chemistry may be important for obtaining optimal performance for a particular protein. Thus, several commercially available slide surface chemistries were evaluated for performance in model arrays

comprising a protein dilution series composed of anti-fibrinogen antibody (antibody array) and human serum albumin (antigen array).

The surfaces included PLL, polystyrene, epoxy-terminated polyethylene glycol (PEG) or dendrimer slides, various amine-derivatized surfaces, and nitrocellulose-coated slides. All proteins were printed in PBS, rinsed in TBS, and then blocked in 3% nonfat dry milk powder dissolved in TBS–0.1% Tween-20. A final rinse in TBS was performed prior to incubation. While no attempt was made to optimize print buffer or blocking conditions for each of the selected surfaces, it was apparent that with the exception of activated polystyrene, most chemistries performed at about the same levels, i.e., within two- to threefold at saturation.

Detection limit ranges for the antigen array were ~60 to ~90 amol and for the antibody array were ~100 to <400 amol. Dendrimer-coated slides appeared to offer some advantage for immobilization of antibodies, providing the lowest detection limits and highest mean signal intensities. However, with CVs ranging from 16 to 43%, it is difficult to predict an optimal substrate for proteins. The study does show that diverse surface chemistries can be used to create protein microarrays at similar levels of performance by applying a simple printing protocol based upon PBS, a buffer commonly employed in working with proteins. What was not addressed is protein stability on these supports.

In summary, we examined a number of studies concerned with the printing of proteins onto various substrates possessing different coating chemistries. Most employed classic methodologies whose origins rested on decades of research related to the development of the ELISA or in the coupling of proteins to affinity matrices. Table 4.3 provides an overview.

Where the protein microarray differed from the classic ELISA was in the much smaller quantities of proteins deposited on the substrate. While micrograms (~10^{-6} g) of protein are employed in the coating of an ELISA well, microarray spots may contain only picograms (~10^{-12} g) or less of protein. The accurate delivery of small volumes [picoliters (pL) or nanoliters (nL)] containing small amounts of protein can be problematic. A certain amount of protein is likely to adsorb onto the quills or capillaries used for printing. The degree at which adsorption takes place is dependent upon a number of factors such as buffer composition, quill surface features and, more importantly, the physical and chemical nature of the protein itself.

This point is illustrated in the work of Delehanty and Ligler (2003) of the Naval Research Laboratory (Washington, DC). They described glass capillary piezoelectric dispensing of antibodies and the effects of buffer composition on printing efficiency. The printing ink was aspirated into a glass capillary surrounded by a piezoelectric collar. The application of voltage to this element caused the aspirate solution to compress and expel a small droplet from the capillary orifice. Typically, picoliter volumes (0.1 <X <1 nL) were dispensed to the surface. The observed problem was nonspecific adsorption of protein within the borosilicate capillary tube.

Table 4.3 Conditions for Printing Protein Microarrays

Study	Print buffer	Additives	Protein concentration	Substrate	Rinse	Blocking buffer
1999 Mendoza	50 mM carbonate, pH 8	None	100 µg/mL	NHS, glass	TBST	PBS, casein
2000 MacBeath	PBS	40% glycerol	100 µg/mL	NHS–BSA, glass	PBS, glycine, pH 8; PBS, pH 7.5	None; PBS, 1% BSA
2000 Joos	PBS	10% glycerol, 0.1% SDS, 5 µg/mL BSA	0.2 to 1.2 mg/mL	Aldehyde, glass	PBST	PBS, 1.5% BSA; 5% milk
2001 Haab	PBS	None	0.1 to 0.3 mg/mL	Poly-L-lysine, glass	PBST, 3% milk	PBS, 3% milk, 0.02% NaN_3
2001 Matson	50 mM carbonate, pH 9	4% sodium sulfate	0.2 to 1.0 mg/mL	ACF, polypropylene plastic	TBST	TBST, 0.1% casein
2001 Moody	PBS	None	50 µg/mL	Nunc Maxisorb, polystyrene plastic	None	SuperBlock
2003 Angenendt	PBS	0.1% NaN_3 preservative	0.3 µg to 1.2 mg/mL	Various: glass coated; membrane coated; plastic with eposy, poly-L-lysine, amine, NC	TBS	TBST, 3% milk
2003 Delehanty	PBS	0.1% BSA	Biotinylated IgG, 2.5 to 20 µg/mL	Avidin, glass	PBST	PBS, 1% BSA, 0.1% NaN_3 (prespotted slide storage)

In this study, the effects of ionic strength and carrier protein (BSA) were examined in terms of the outcome for microarray printing. As noted previously, many factors can unduly influence printing performance. Proteins are notoriously bad when it comes to nonspecific adsorption. It should not be much of a surprise that some portion of a protein probe will adsorb to the printing device, whether it is a stainless steel quill or glass capillary.

In the case of the piezo system, biotinylated antibodies were dispensed at 1 nL to produce spot diameters of 230 μ on avidin-coated glass slides (NeutrAvidin™, Pierce). The concentration of biotinylated antibody (Cy5-labeled mouse IgG) was varied from 2.5 to 20 μg/mL, while the ionic strength varied by dilution of PBS (10 mM to 150 mM, where 150 mM = 137 mM NaCl, 3 mM KCl, and 10 mM phosphate, pH 7.4) in distilled water. BSA was added as a carrier protein at 0.1% (w/v). After printing, the slides were rinsed in 150 mM PBS containing 0.05% Tween-20 (PBST), followed by distilled water and dried under a nitrogen stream.

Two effects were observed. First, in the absence of carrier BSA, the amount of IgG deposited onto the slide was reduced at high ionic strength. Presumably more IgG remained adsorbed to the borosilicate glass capillary under high salt conditions. As the input concentration of IgG increased, less of an ionic strength buffer effect was noted. Most likely, at the higher IgG (20 μg/mL protein) concentration, sufficient coating of the capillary was achieved. However, the addition to the print buffer of 0.1% BSA (1 mg/mL) was more effective in coating the capillary, thereby significantly reducing the nonspecific adsorption of IgG. Thus, more IgG was deposited onto the slide. An added benefit with the addition of BSA was the finding that spot morphology improved. A reduction in the so-called donut spot morphology was most likely due to less peripheral drying occurring with the higher protein content within the spot.

References

Angenendt, P., Glokler, J., Sobek, J., Lehrach, H., and Cahill, D.J., Next generation of protein microarray support materials: evaluation for protein and antibody microarray applications, *J. Chromatogr. A*, 1009, 97–104, 2003.

Balázsi, G., Kay, K.A., Barabási, A.L., and Oltvai, Z.N., Spurious spatial periodicity of co-expression in microarray data due to printing design, *Nucleic Acid Res.*, 31(15), 4425–4433, 2003.

Battaglia, C., Salani, G., Consolandi, C., Bernardi, L.R., and DeBellis, G., Analysis of DNA microarrays by non-destructive fluorescent staining using SYBR green II, *Biotechniques*, 29(1), 78–81, 2000.

Brown, P.O. and Shalon, T.D., Methods for fabricating microarrays for biological samples, US Patent 5,807,522, issued 1998.

Daub, M., TopSpot Technology: highly parallel dispensing for production of microarrays, Presentation 2002, Euro Biochips Conference, Berlin, Germany.

Delehanty, J.B. and Ligler, F.S., Method for printing functional protein microarrays, *Biotechniques* 34, 380–385, 2003.

Diehl, F., Grahlmann, S., Beier, M., and Hoheisel, J.D., Manufacturing DNA microarrays of high spot homogeneity and reduced background signal, *Nucleic Acid Res*., 29(7), 1–5, 2001.

George, R.A., Woolley, J.P., and Spellman, P.T., Ceramic capillaries for use in microarray fabrication, *Genome Res*., 11, 1780–1783, 2001.

Haab, B.B., Dunham, M.J., and Brown, P.O., Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions, *Genome Biol*., 2, 0004.1–0004.13, 2001.

Hedge, P., Qi, R., Abernathy, C., Gay, C., Dharap, S., Gaspard, R., Hughes, J.E., Snesrud, E., Lee, N., and Quackenbush, J., A concise guide to cDNA microarray analysis, *Biotechniques*, 29, 548–562, 2000.

Hessner, M.J., Wang, A., Hulse, K., Meyer, L., Wy, Y., Nye, S., Guo, S.W., and Ghosh, S., Three color cDNA microarrays: quantitative assessment through the use of fluorescein-labeled probes, *Nucleic Acid Res*., 31(4), 1–6, 2003a.

Hessner, M.J., Wang, X., Khan, S., Meyer, L., Schlicht, M., Tackjes, J., Datta, M.W., Jacob, H.J., and Ghosh, S., Use of a three-color cDNA microarray platform to measure and control support-bound probe for improved data quality and reproducibility, *Nucleic Acid Res*., 31(11), 1–9, 2003b.

Kane, M.D., Jatkoe, T.A., Stumpf, C.R., Lu, J., Thomas, J.D., and Madore, S.J., Assessment of the sensitivity and specificity of oligonucleotide (50-mer) microarrays, *Nucleic Acid Res*., 28(22), 4552–4557, 2000.

Lane, S., Birse, C., Zhou, S., Matson, R., and Liu, H., DNA array studies demonstrate convergent regulation of virulence factors by Cph1, Cph2, and Efg1 in *Candida albicans*, *J. Biol. Chem*., 276, 48988–48996, 2001.

Macas, J., Nouzova, M., and Galbraith, D.W., Adapting the Biomek® 2000 laboratory automated workstation for printing DNA microarrays, *Biotechiques*, 25, 106–109, 1998.

MacBeath, G. and Schreiber, S.L., Printing proteins as microarrays for high-throughput function determination, *Science*, 289, 1760–1763, 2000.

Martinez, M.J., Aragon, A.D., Rodriguez, A.L., Weber, J.M., Timlin, J.A., Sinclair, M.B., Haaland, D.M., and Werner-Washburne, M., Identification and removal of contaminating fluorescence from commercial and in-house printed DNA microarrays, *Nucleic Acid Res*., 31(4), 1–8, 2003.

Martinsky, R.S., Microarray printing device including printing pins with flat tips and exterior channel and method of manufacture, US Patent 6,101,946, issued 2000.

Matson, R.S., Rampal, J., Pentoney, S.L., Anderson, P.D., and Coassin, P., Biopolymer synthesis on polypropylene supports: oligonucleotide arrays, *Anal. Biochem.*, 224, 110–116, 1995.

Matson, R.S., Milton, R.C., Cress, M.C., and Rampal, J.B., Microarray-based cytokine immunosorbent assay, Poster 20, Oak Ridge Conference, Seattle, WA, 2001.

McQuain, M.K., Seale, K., Peek, J., Levy, S., and Haselton, F.R., Effects of relative humidity and buffer additives on the contact printing of microarrays with quill pins, *Anal. Biochem*., 320, 281–291, 2003.

Mendoza, L.G., McQuary, P., Mongan, A., Gangadharan, R., Brignac, S., and Eggers, M., High-throughput microarray-based enzyme-linked immunosorbent assay (ELISA), *Biotechiques*, 24, 778–788, 1999.

Moody, M.D., Van Arsdell, S.W., Murphy, K.P., Orencole, S.F., and Burns, C., Array-based ELISAs for high-throughput analysis, *Biotechniques*, 31, 1–7, 2001.

Okamoto, T., Suzuki, T., and Yamamoto, N., Microarray fabrication with covalent attachment of DNA using Bubble Jet technology, *Nat. Biotechnol.*, 18, 438–441, 2000.

Rees, W.A., Yager, T.D., Korte, J., and Von Hippel, P.H., Betaine can eliminate the base pair composition dependence of DNA melting, *Biochemistry*, 32, 137–144, 1993.

Reese, M.O., van Dam, R.M., Scherer, A., and Quake, S.R., Microfabricated fountain pens for high-density DNA arrays, *Genome Res.*, 13, 2348–2352, 2003.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O., Quantitative monitoring of gene expression patterns with a complementary DNA microarray, *Science*, 270, 467–470, 1995.

Shalon, T.D., and Brown, P.O., Methods for fabricating microarrays for biological samples, US Patent 6,110,426, issued 2000.

Shearstone, J.R., Allaire, N.E., Getman, M.E., and Perrin, S., Nondestructive quality control for microarray production, *Biotechniques*, 32, 1051–1057, 2002.

Silzel, J.W., Cerecek, B., Dodson, C., Tsong, T., and Obremski, R.J., Mass-sensing, multianalyte microarray immunoassay with imaging detection, *Clin. Chem.*, 44, 2036–2043, 1998.

Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B., Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization, *Mol. Biol. Cell*, 9, 3273–3297, 1998.

Stimpson, D.I., Cooley, P.W., Knepper, S.M., and Wallace, D.B., Parallel production of oligonucleotide arrays using membranes and reagent jet printing, *Biotechniques*, 25, 886–890, 1998.

Wang, X., Ghosh, S., and Guo, S.W., Quantitative quality control in microarray image processing and data acquistion, *Nucleic Acid Res.*, 29(15), 1–8, 2001.

Wang, X., Hessner, M.J., Wu, Y., Pati, N., and Ghosh, S., Quantitative quality control in microarray experiments and the application in data filtering, normalization and false positive rate prediction, *Bioinformatics*, 19, 1341–1347, 2003.

Yang, H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., and Speed, T.P., Normalization of cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation, *Nucleic Acid Res.*, 30(4), 1–10, 2002.

chapter five

Gene expression: Microarray-based applications

Introduction

In the past several years, we have witnessed exceptional growth in both the depth and breadth of microarray-based applications. The focus of this chapter will be on representative examples of how DNA microarrays are able to address a number of problems in the biomedical fields. The most popular microarray format is a glass slide to which are immobilized by some means a library of nucleic acids. The microarray slide originated from the work of the Brown and Davis groups at Stanford University (Schena et al., 1995). It was an adaptation from filter grid arraying of cDNA clones to examine gene expression on a global scale.

Miniaturization allowed several benefits to be realized. First, higher density arrays could be produced, allowing the monitoring of many more genes. Second, the sample volume, reagent consumption, and processing steps were greatly reduced. Third, fluorescence labeling and detection could be used in place of radiolabeling and autoradiography. With the development of a slide scanner based upon fluorescence detection, the throughput from sample to result was greatly increased. Through dual-label competitive hybridization analysis (e.g., Cy3-labeled control gene population compared to Cy5-labeled sample population), the slide microarray became an enabling technology for genomic studies.

Applications demonstrating DNA microarray utility

Gene expression

"The temporal, developmental, topographical, histological, and physiological patterns in which a gene is expressed provide clues to its biological role."

With this introductory statement by Schena et al. (1995), the Stanford groups of Davis and Brown heralded what would soon become a new paradigm for biological investigation. At the center of this global examination of gene expression was the glass slide cDNA microarray. The Stanford paper in *Science* not only showed the power of the microarray in determining gene expression patterns, but also provided much of the detail that enabled others to construct their own micorarrays. This is perhaps the most important contribution made by this landmark paper.

Based upon Dari Shalon's Ph.D. thesis and methods later published on Ron Brown's website, biologists were quickly introduced to the building of robotic pin printers, preparation of slide arrays, the utility of two-color labeling, and the construction of laser scanners (Shalon et al., 1996). While others "fiddled to the tune" of sequencing by hybridization (SBH) or mutation detection in hopes of moving DNA arrays rapidly into diagnostics, the field of genomics was born. SBH has essentially been abandoned and while the potential for DNA arrays in diagnostics remains, the slide microarray continues to advance as an important genomics tool. In this section we will review what has been accomplished with gene expression microarrays.

Armed with this new tool, Schena et al. (1996) created a microarray of 1,046 human cDNAs of unknown sequence. They were derived from human peripheral blood lymphocytes transformed with Epstein–Barr virus. Suitably sized inserts [>600 base pairs (bp)] were cloned into a lambda vector, subsequently infected into an *Escherichia coli* strain, and finally amplified by polymerase chain reaction (PCR) using 5′-amino-modified primers. The resulting 5′-amino-modified cDNA amplicons were then arrayed onto silylated microscope slides. Next, the expression levels in human Jurkat cells undergoing heat shock or phorbol ester induction were examined.

Total mRNAs from control and induced cells were labeled using reverse transcriptase with the incorporation of fluorescene-dCTP (control, green label) or Cy5-dCTP (induced, red label). The two populations were hybridized to separate arrays. However, the labels were also swapped to verify that any differences in labeling efficiency did not affect the result.

While more than 95% of the arrayed cDNA probes showed hybridization at signal intensities ranging over 3 logs, only a few "genes" displayed significant differences in expression. In fact, only 17 of 1046 (1.6%) underwent changes between two- and ~six-fold. While the absolute expression levels of these genes varied considerably between microarray hybridization and RNA blots, the relative fold changes showed a good correlation (R^2 ~0.8) between the microarray and blot results (Figure 5.1).

As we shall see, most genes are not expressed *a priori* in response to a particular metabolic or environmental change; rather a smaller number are significantly induced or repressed (Figure 5.2). Moreover, it is the response of these particular genes that is important in the drug discovery process. As a case in point, 4 of 17 (23.5%) of the twofold expressed genes were discovered to be novel upon sequencing of the cDNA clones. This also points to the power of the cDNA microarray in that the monitoring of gene expression

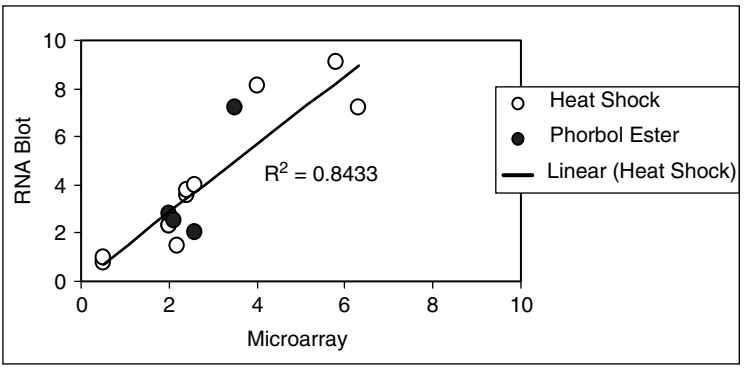


Figure 5.1 Comparison of gene expression levels: Northern (mRNA) vs. microarray. (From Schena, M. et al., *Proc. Natl. Acad. Sci. USA*, 93, 10614–10619, 1996. With permission.)

Fold-Discrimination ≥ 2.0

Study	Diff. Expressed/Total Genes	% Diff. Expressed
Lock et al.	1080/7026	15.4
Moustafa et al.	213/12530	1.7
Lapteva et al.	9/1081	0.8
Nishizuka et al.	29/7000	0.4
Reilly et al.	332/11000	3.0
Maeda et al.	8/2304	0.3
Katsuma et al.	82/4224	1.3

Figure 5.2 The comfort zone. (From Schena, M. et al., *Proc. Natl. Acad. Sci. USA*, 93, 10614–10619, 1996. With permission.)

patterns and the discovery of novel genes interactions can be achieved without knowledge of gene sequence.

While the Schena papers (1995 and 1996) served as first demonstrations of cDNA microarray technology, it was clear that further refinements were necessary in order to realize the full potential of the microarray. Arraying technology was in its infancy and suffered from inconsistency in uniform spotting, making it difficult to compare slides. Refinements in labeling and detection were also needed.

Despite these shortcomings, relative expression levels could be monitored with some degree of confidence in the data by employing comparative hybridization (DeRisi et al., 1996). In this approach, the labeled control and test mRNA populations are mixed and then applied to a single slide array, thus avoiding problems associated with differences in slide-to-slide uniformity. Also, well-characterized synthetic targets could be doped into the samples to serve as internal standards and permit quantitative estimates of relative abundance levels between the two expressed gene populations. The

two-color, comparative hybridization approach also allowed for a visual interpretation of the microarray results.

DeRisi et al. (1996) labeled tumorigenic cell mRNA with a green dye and nontumorigenic control cell mRNA with a red dye. Thus genes overexpressed in tumor cells would appear as green spots and those preferentially expressed in control cells would have red spots. Equivalent levels of gene-specific expression between the cell lines would show an equal mix of red and green, resulting in the appearance of yellow spots. Of course, if no genes were expressed, then the spots would remain dark.

In the DeRisi study, an additional level of control was provided by examining 90 so-called "housekeeping" genes that would remain invariant between control and test mRNA populations. In fact, some differential expression will occur, and therefore user-defined cutoffs in relative gene expression are necessary. For example, in this study, the red-to-green (R:G) ratio for housekeeping gene expression was 1.13 while for internal standards it was 0.97. Setting minimal cutoffs at 3 standard deviations, ratios <0.52 and >2.4 are required to measure statistically significant levels of differential expression, i.e., at least a twofold change or "comfort zone" is necessary to compensate for variance in constitutive biological levels and assay performance. As a result, DeRisi et al. found that only about 9% of the 870 genes (15 of 870 down-regulated; 63 of 870 up-regulated) on the microarray were observed to be differentially expressed. Attempts to utilize data from lower levels of differential expression (below twofold) have to date been controversial although with careful experimental control this should be possible.

One of the most elegant early experiments with cDNA microarrays was the assessment of gene expression during the classic metabolic (diauxic) shift in *Saccharomyces cerevisiae* in going from anaerobic to aerobic states (DeRisi et al., 1997). Thus, superimposed upon the metabolic pathways characterizing the flow of metabolites during glycolysis and gluconeogenesis was now the temporal relationship of the pathway genes as the shift from fermentation to respiration takes place (Figure 5.3).

What was most impressive about this study was that the microarray contained essentially the entire yeast genome of 6400 ORFs (open reading frames), providing a global genomic overview. Gene expression "snapshots" could be taken during the diauxic shift using micorarrays and the entire metabolic process reviewed frame by frame. The mRNA from cells grown in glucose-rich media (fermentation state) were labeled with Cy3 (green) and served as the reference, while mRNA from cells transitioning into glucose-depleted media (shifting to aerobic respiration) were labeled with Cy5 (red). [*Note:* The introduction of the Cy5:Cy3 ratio has been largely adopted as standard practice for microarray gene expression analysis although newer dyes such as the ALEXA series (Molecular Probes, Inc., Eugene, OR) are now also in use.]

The diauxic shift experiment turned out to be an extremely important demonstration. While *S. cerevisiae* exhibited very little in the way of differential expression activity (19 of 6400 genes showing twofold expression)

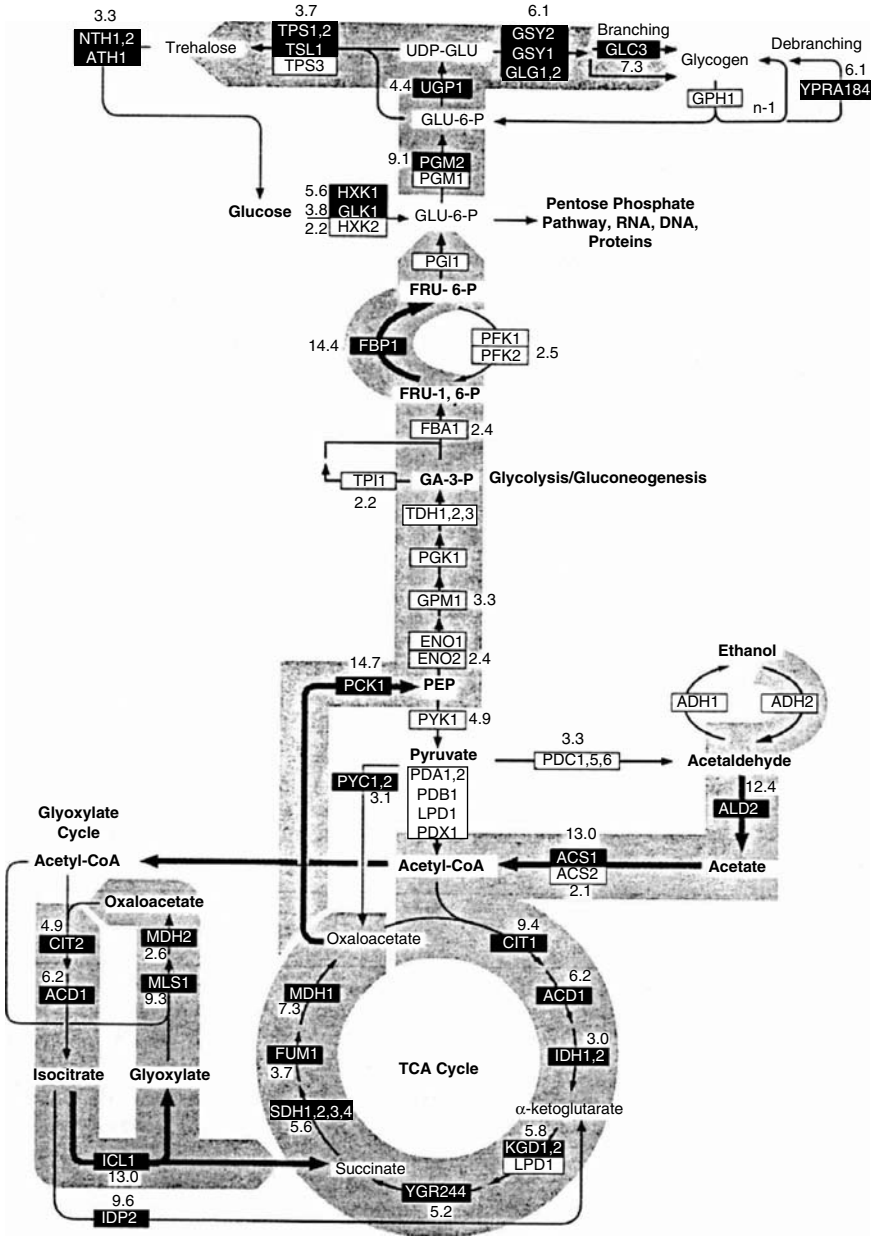


Figure 5.3 Diauxic shift in *Saccharomyces cerevisiae*. Monitoring metabolic pathways using a gene expression microarray. (From DeRisi, J. et al., *Science*, 278, 680–686, 1997. With permission.)

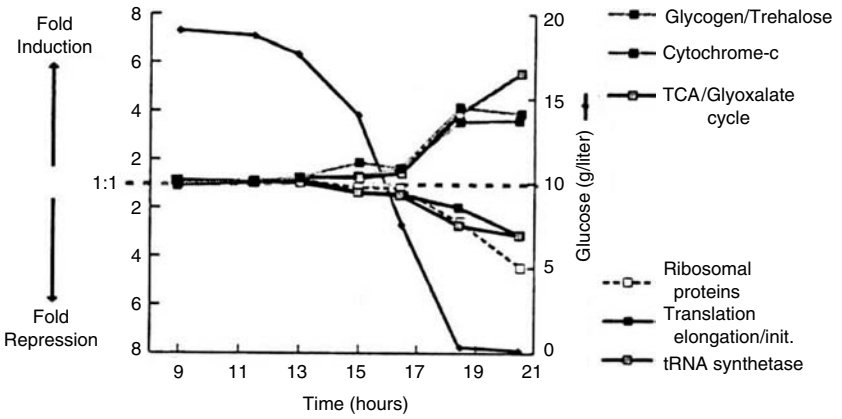


Figure 5.4 Global assessment of gene expression among various metabolic states. (From DeRisi, J. et al., *Science*, 278, 680–686, 1997. With permission.)

during exponential growth in glucose-rich media, this was certainly not the case during glucose depletion. At expression levels measured at twofold or greater, more than 25% of the yeast genome had undergone induction (710 of 6400) or repression (1030 of 6400) in response to the anaerobic-to-aerobic shift.

The experiments also resulted in the discovery of at least 400 genes that at the time were found to have no known function and did not appear on public databases. Those genes coding for enzymes associated with the metabolic pathways permitted a dynamic view of shifts in metabolite flow. For example, the ALD2 gene encoding for aldehyde dehydrogenase and the ACS1 gene for acetyl-CoA synthase underwent 12.4- and 13-fold inductions, respectively. These enzymes are responsible for moving acetyl-CoA into the tricarboxylic acid (TCA) and glyoxalate cycles. Similar inductions of the PCK1 gene (encoding phosphoenolpyruvate [PEP] carboxykinase) and the FBP1gene (encoding fructose 1,6-biphosphatase) were responsible for reversing the flow of oxaloacetate ultimately into glucose-6-phosphate for glycogen storage.

Obviously, with changes in 25% of the yeast genome occurring at some point during the diauxic shift, a large number of regulator events remain to be mapped in this manner. One way to address this is to group gene families rather than monitor individual genes (see Eisen et al., 1998, regarding cluster analysis). For example, genes associated with a particular pathway such as glycolysis or protein synthesis can be monitored collectively for average fold induction or repression during the time course of glucose starvation (Figure 5.4). For additional details regarding the strategy for constructing specific yeast ORF arrays for gene expression analysis and yeast strain comparisons, see Lashkari et al. (1997).

In a related application of this approach, Ferea et al. (1999) examined variations in gene expression of progeny during adaptive evolution. The yeast genomes were monitored in evolving strains subjected to growth under

a glucose-limited chemostat for 250 generations. All the evolved strains studied appeared to experience similar changes in their gene expression profiles relative to the parental strain. Essentially, those genes involved in respiration were up-regulated while genes encoding enzymes of the glycolytic pathway (fermentation of excess glucose) were repressed relative to the parental strain (Figure 5.5). This is consistent with the physiological response of yeast in shifting from fermentation to respiration. Each new generation adapts to become more efficient in the utilization of glucose (via oxidation) under limiting concentrations.

Iyer et al. (1999) extended microarray pathway expression profiling to examine the effects of serum on fibroblast growth. It is well known that certain growth factors are required in order to propagate mammalian cells in culture. Serum sources such as fetal calf serum contain these growth factors. Using a cDNA microarray comprising 8613 human genes, the changes in mRNA levels were monitored from 15 min to 24 hr following the introduction of fibroblasts into serum. Cluster analysis was conducted on 517 (6% of the cDNA microarray) genes showing significant change (≥2.20-fold expression change). The genes were grouped into ten cluster families on the basis of the similarities of their expression profiles (Figure 5.6).

Genes involved in the encoding of transcription factors and signal transduction proteins were induced within 15 min following fibroblast transfer from serum-poor into serum-rich media. The genes involved in cell cycle progression began to appear (were induced) about 16 hr following passage into serum. For example, induction of mRNA encoding various subunits of the RNA/DNA polymerases and cyclins that regulate growth phases appeared in this cluster. Of particular interest was the discovery that genes involved in wound healing were induced within hours following serum stimulation. For example, COX2 (chemotaxis, neutrophil activation), MCP1 (macrophage recruitment), IL-8 (T lymphocyte activation), ICAM-1 (B lymphocyte activation), and VEGF (angiogenesis) among others were all induced within a few hours (Figure 5.7). In addition, at least 200 previously unknown genes were also identified as participating in the fibroblast growth and cell cycle progression.

The cDNA microarray format provides the ability to monitor gene expression without prior knowledge of the probe cDNA sequence. However, because cDNAs are typically 200 bp to 600 bp in size, considerable potential for the occurrence of cross-hybridization is present. When the gene sequence is known, oligonucleotide probes can be carefully designed to avoid this problem. However, the selection of a unique gene-specific probe sequence of an appropriate size is also problematic because partial but nevertheless hybridizable sequence copies may be present at random within the genome. Under low stringency conditions, it is possible to hybridize 6- to 8-mer oligonucleotides (Drmanac et al., 1990).

In order to reduce the effects of cross-hybridization that lead to false positives, researchers at Affymetrix created a series gene-specific but closely

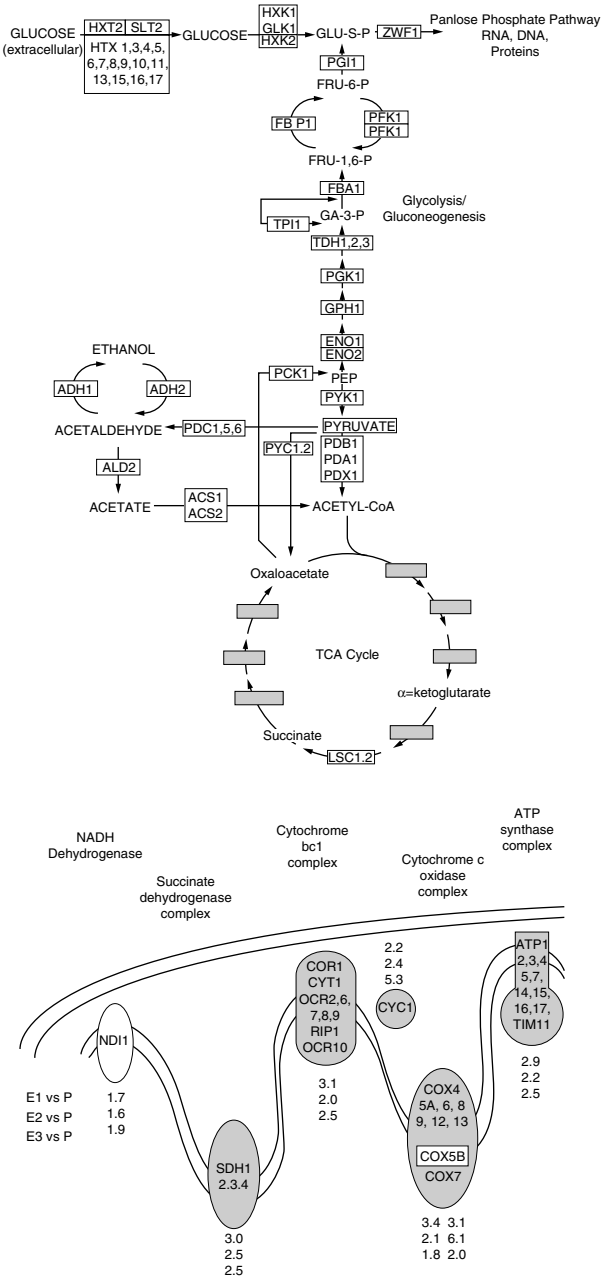


Figure 5.5 Examining regulation of yeast genes from evolving strains. (From Ferea, T.L. et al., *Proc. Natl. Acad. Sci. USA*, 96, 9721–9726, 1999. With permission.)

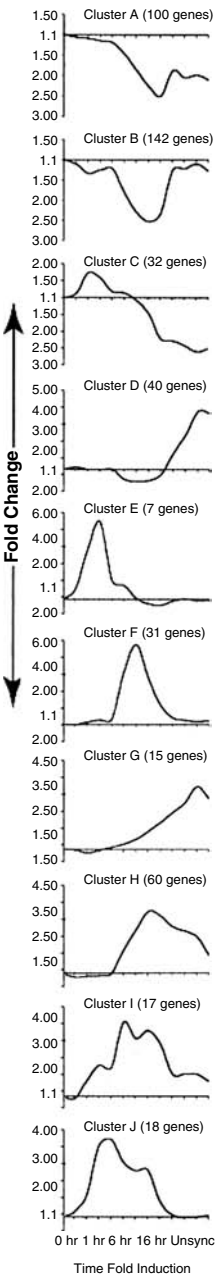


Figure 5.6 Gene cluster profiling. (From Iyer, V.R. et al., *Science*, 283, 83–87, 1999. With permission.)

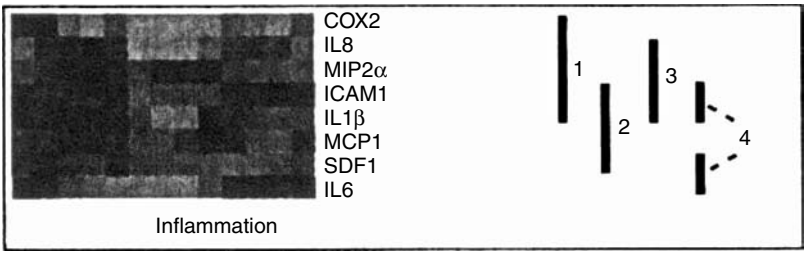


Figure 5.7 Human fibroblasts responsive to serum monitoring genes associated with inflammation. (From Iyer, V.R. et al., *Science*, 283, 83–87, 1999. With permission.)

related complementary (PM or perfect match) probes and then introduced corresponding mismatched (MM) probes in which the single-base mismatch is placed at a central location in the sequence. For short oligonucleotide probes (e.g., 25-mer in the case of the GeneChip array), this central location provides the optimal instability relative to the perfect match. As a result, the ratio of PM:MM provides a convenient means to improve both specificity and sensitivity on the array.

Wodicka et al. (1997) created such an array for measuring yeast gene expression based upon 25-mer oligonucleotides covering 6200 ORFs. Each ORF was represented by 20 PM and 20 MM probes. Why so many probes? Simply put, not all probes hybridize in a predictable manner. Averaging across a number of probes improves the outcome. Thus, the yeast expression chip comprised over 65,000 probe features and required a set of four chip subarrays.

The Wodicka et al. (1997) paper also defined the performance of the Affymetrix chip. Semiquantitative measurement of the absolute abundance of mRNA species was possible. Hybridization of total yeast-genomic DNA to the chips revealed the mean hybridization signal across 6049 probe sets to vary by 25% coefficient of variance (CV). The use of gDNA serves to normalize because most genes are represented only once in the population. In fact, the majority (98%) of the intensities were found to cluster well within two standard deviations. Thus, the concentration of a given mRNA could be estimated at >95% probability to reside within ± twofold of its actual concentration. Measurement at widely different total gDNA concentrations did not appreciably affect this outcome.

Finally, the Affymetrix chip further corroborated the fact that only a relatively small number of genes show significant levels of differential expression in response to particular stimuli. A comparison of rich and minimal media revealed that 36 mRNAs (genes) were more abundant (5- to 10-fold higher) in rich media, while 140 genes were more abundant in minimal media. This collectively represents less than 3% of the yeast genome. In terms of "absolute" concentration or copy number per cell, the mRNA distribution was very similar in cells grown in rich and minimal media (Figure 5.8). An estimated 50% of the mRNA population (~15,000 total copies

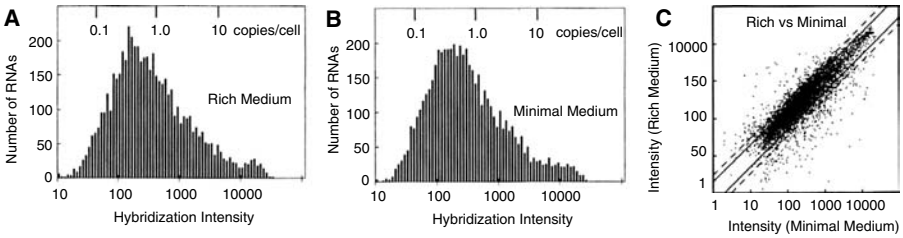


Figure 5.8 Distribution of expression levels for cells grown in rich or minimal media. (From Wodicka, L. et al., *Nature Biotechnol.*, 15, 1359–1367, 1997. With permission.)

per yeast cell) were present at 0.1 to 1 copy per cell; 26% (1 to 10 copies per cell); 5% (>10 copies per cell); and 19% (<0.1 copies per cell).

An interesting application of the Affymetrix chip yeast genome array is direct allelic variation scanning (Winzeler et al., 1998). The yeast expression chip provided 21.8% coverage of nonrepetitive regions of the yeast genome. Winzeler et al. reasoned that this was enough to capture a small but significant portion of the genetic variation found between strains and hoped to find markers that would map out phenotypic differences. Two strains of *S. cerevisiae* that were phenotypically distinguishable were tested.

When labeled genomic DNA was hybridized to the high density array, a total of 3714 (contributing ~4.7% of the estimated strain-to-strain variation) marker candidates exhibited a greater than 99% probability of differentiating the two strains. These biallelic markers were spaced at about 3500 bp. Cycloheximide sensitivity (phenotype) was mapped to the PDR5 gene (a multidrug resistance pump) in the one strain exhibiting hypersensitivity to cyclohemimide. Thus, phenotypic variation between the yeast strains was accurately mapped directly based upon differential hybridization of genomic DNA.

De Saizieu et al. (1998) examined the transcriptional activity of bacterial genomes (influenza, pneumonia) using an Affymetrix chip of 64,000 probes that were complementary to 100+ genes for each genome or about 150 probe pairs per gene. This represents about 5% coverage for the *S. pneumoniae* genome. Because bacterial mRNA lacks the 3′ poly(A) tailing, enrichment of mRNA from total RNA by affinity purification is not possible. Instead, labeling of total RNA must be undertaken. In this case, the most efficient labeling was accomplished using psoralen–biotin, providing incorporation of about one biotin per 120 nucleotides.

Fragmented, biotinylated RNA prepared in this manner was hybridized to the array and the signal developed using streptavidin-R-phycoerytherin. A confocal laser scanner was used for detection. The researchers estimated that they could detect two transcripts per cell based upon labeling efficiency and an estimated 4% mRNA content in total bacterial RNA. Thus, chip detection of labeled transcripts was found to be more sensitive than detection by Northern blot. Specific genes (e.g., basal levels of cinA) undetectable on Northern blots were quantifiable on the microarray. In addition, it was

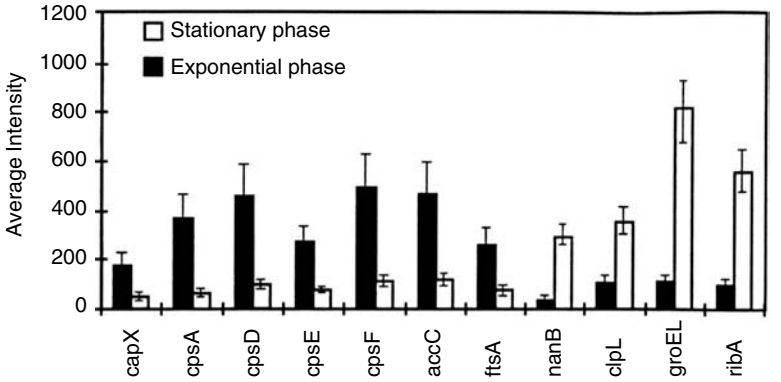


Figure 5.9 Gene expression microarray monitoring of bacterial growth. (From de Saizieu, A. et al., *Nature Biotechnol*., 16, 45–48, 1998. With permission.)

possible to monitor gene expression in moving from exponential to stationary growth phases (Figure 5.9).

The issue of whether to use the enriched poly(A) RNA or total RNA was addressed in a later paper by Mahadevappa and Warrington (1999). Using human adenocarcinoma cells, they examined the recovery of detectable transcripts from varying numbers of cells for both protocols. Of the ~1800 genes represented on the chip, about 35% were observed to be detectable using either of the two preparations.

Of these "detectable" transcripts (number of copies per cell basis), about 86% of the same transcripts were reported across several levels of relative abundance. For example, at levels >10 copies per cell, poly(A) RNA-derived transcripts ranged from 102 to 134, and total RNA transcripts ranged from 119 to 141. At levels <2 copies per cell, poly(A) RNA ranged from 141 to 165 transcripts, and total RNA ranged from 152 to 162 transcripts. Intermediate abundance levels were also similar. Added benefits from total RNA labeling included the need for less starting material and improved yield of higher quality material (less degradation by elimination of the extraction process).

In this section, we described some of the important early work on DNA microarrays. What we have discovered from these examples is that the microarray format can be used quite effectively to view differences (or similarities) in gene expression between control and test populations of cells. It is also evident that such temporal changes do not represent genome-wide levels of response but rather involve relatively small numbers of genes (1 to 10%) that are up- or down-regulated. Because of the large number of genes represented on a microarray and the level of biological variability, it is difficult to draw conclusions from examining single-gene events. A more meaningful approach for large gene expression data sets is to monitor these events by grouping genes of similar expression patterns together in clusters. As we have seen, clusters tend to bring together genes that participate in related functions or those involved in the junctures of metabolic pathways.

In the following sections, we will look at representative applications of DNA microarrays in the biomedical research field.

Biomedical research applications

Drug discovery

As we examine various applications for microarrays, it is important to understand the distinction between *pharmacogenomics* and *pharmacogenetics*. Microarrays are used for studies in both fields but for slightly different purposes. Pharmacogenomics is the application of genomics to the drug discovery and development process. Microarrays reveal which targeted genes are turned off or turned on in response to candidate drugs.

In primary screenings, thousands of candidate drugs (compound libraries) are directed against a few targets, e.g., receptors. Screening generally involves a biological assay using the putative target and measuring direct drug binding and response. Microarrays are usually not used in primary screening. They are most suitable for analyses involving the secondary screening of a few lead drug candidates produced from the primary screen.

After an intended target (gene product) is identified, secondary screening is an attempt to determine which drugs interact. As noted earlier, microarrays can be constructed of gene-specific probes, i.e., cDNA clones, without prior knowledge of the gene's function. The ability to examine alterations in gene expression patterns in response to a candidate drug may lead to the identification of new (gene) targets. For example, a gene of unknown function may associate or cluster with known genes in response to a specific drug. This may provide important clues to the unknown gene's function and suitability as a target (Ivanov et al., 2000).

Pharmacogenetics involves understanding an individual's genetic make-up relative to drug action. For example, based on a particular genotype we may ask which drug is most effective for treatment without adverse (off-target) side effects. Microarrays are used here to monitor the up- or down-regulation of genes involved in various off-target pathways in order to assess potential drug toxicity — often referred to as the field of *toxicogenomics*. In this case, a single drug whose target has been identified is examined for off-target responses such as activation or shut-down of important metabolic pathways. For additional review on the role of the DNA micorarray in the drug discovery process, see the collections of reviews by Jain (2000), Zanders (2000), and Ivanov et al. (2000). This section will examine studies aimed at adapting DNA microarrays for drug discovery.

The fundamental approach in using microarrays for gene expression analysis is really to seek out differences between a control cell gene population and the gene population from test cells. We then examine those genes whose expression levels changed and attempt to discover what caused the change. In reality, we observe many changes and, as we have seen, this leads to a rather interesting but complicated chain on events. Drug discovery

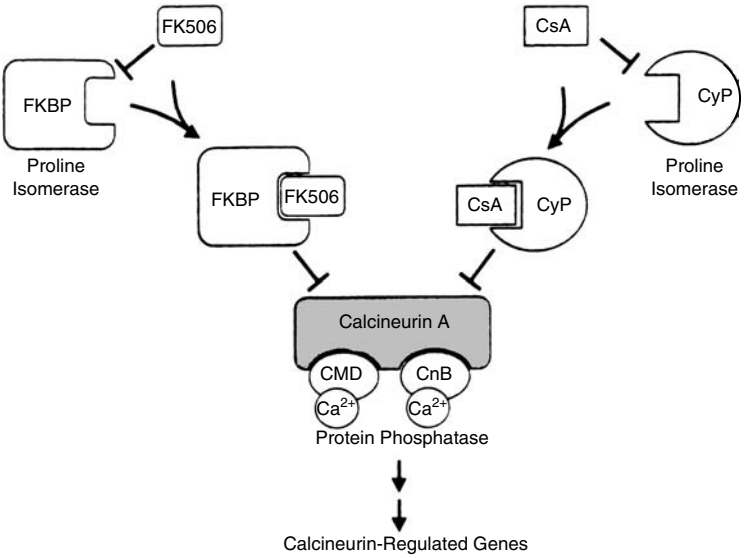


Figure 5.10 Calcineurin signaling pathway in yeast. (From Marton, M.J. et al., *Nature Med.*, 4, 1293–1301, 1998. With permission.)

involves attempts to find compounds that interact with specific targets of interest. The ideal candidate drug should be very specific for a target and free of unwanted side effects. However, traditional approaches to determining so-called off-target drug responses are time consuming, expensive, and not always precise.

Because of the ability of the DNA microarray to examine global changes in gene expression, it has become an important new tool for measuring off-target drug interactions. However, with so many changes, how do we determine whether a drug's mechanism of action is through interaction with the presumed target? One clever way is to "knock out" a particular gene (the putative target) and assess whether the drug is still effective.

Marton et al. (1998) used mutant yeast strains and examined gene expression patterns in the presence or absence of a specific drug. Comparison of these "signatures" provided clues to the mechanisms of action of specific drugs. They chose to study the calcineurin signaling pathway in yeasts (Figure 5.10); calcineurin was involved in the regulation of a number of key cellular functions such as the onset of mitosis in yeasts. In mammals this calcium-activated protein phosphatase has been implicated in a wide range of biological functions from T cell activation to serving as an effector in short- to long-term memory transition.

Calcineurin activity is inhibited by two immunosuppressant drugs: tachrolimus (FK506) and cyclosporin A (CsA). Do these drugs interact in the same manner? In order to answer this question, the gene expression microarray signatures for both drugs in wild type yeasts were obtained. Next, signatures were obtained in mutant strains in which the putative target was

deleted. The hypothesis was that if the mutated gene encoded a protein in the pathway affected by the drug, the mutant signature of the drug would be different from the wild-type signature or absent.

Drug toxicity

Cleary et al. (2001) studied amphotericin B toxicity to human mononuclear cells using a commercially available cDNA nylon membrane-based microarray (Clontech, Palo Alto, CA). Amphotericin B, a fungicide, elicits immune responses such as activation of interleukin (IL)-1β and tumor necrosis factor (TNF)-α. The IL-1β gene expression is related to the accumulation of intracellular calcium known to be mediated by amphotericin B (see Cleary et al., 2001, References 4 and 5). Because calcium plays a role as a second messenger, it was recognized that additional genes may undergo regulation in response to the fungicide.

The Clontech microarray used in this study contained 588 cDNA fragments arranged in several relevant functional categories: oncogenes, tumor suppressor genes; ion channels, transducers; apoptosis; transcription factors; cell receptors such as for interleukins, hormones, chemokines; and extracellular cell signaling-related genes. Of the 588 genes represented on the microarray, 16 were up-regulated and 4 were down-regulated in the THP-1 (human acute monocytic leukemia cell line) 6 hr following administration of the drug. Most genes (75%) exhibited greater than 10-fold differences relative to control cells. These genes encoding the listed proteins were identified as having important regulatory functions implicating potential mechanisms of action for amphotericin B:

1. Transcription factor (activator protein) AP-1 is up-regulated. AP-1 is involved in the induction of genes encoding inflammatory responses, e.g., IL-1.
2. MAL (myelin and lymphocyte) protein is up-regulated. This protein is involved in cell signaling and protein trafficking.
3. Caspase-4 is down-regulated and represents a potential block to activation of the apoptosis pathway.
4. Cell adhesion protein (intercellular adhesion molecule) ICAM-1 is up-regulated. This protein plays a role in cell–cell adhesion and leukocyte migration.
5. IL-8 is up-regulated. This chemokine is responsible for neutrophil activation.

The application of the cDNA microarray revealed that the fungicide appeared to affect a number of cellular processes. While the content of the Clontech microarray was rather limited, it provided a substantial amount of new information regarding amphotericin B-mediated cellular toxicity.

Now that we have identified most of the human genome, a higher density array could be easily applied to refine the study. By applying this

global look at gene expression, we can greatly enhance our understanding of drug toxicity. The next studies cited expand upon such applications using microarrays with expanded gene contents.

Reilly et al. (2001) used a mouse model to study the global gene expression profile associated with drug-induced liver toxicity brought about by the analgesic drug, acetaminophen (APAP). The mechanism of APAP toxicity is not well understood because it is implicated in a variety of biochemical events leading to cellular damage such as oxidative stress, disruption of calcium and mitochondrial hemostasis, alterations to transcription, inflammation, and programmed cell death pathways. For this reason, a global examination of gene expression events using the microarray was undertaken in the hope that additional information regarding the mechanisms of toxicity would be found.

APAP was administered to mice and the progression of hepatotoxicity monitored by histochemical means and by microarray analysis. The Affymetrix oligonucleotide array (Mul1K sub A, sub B) was used to access gene expression activity across 11,000+ genes and expressed sequence tags (ESTs). Monitoring began 6 hr after administration in order to establish the early onset of hepatotoxicity. Fold change above 2.0 was used as the threshold level. Of the 11,000 genes on the array, a total of 332 or ~3% were scored as up- or down-regulated. Of these, >50% exhibited fold changes between 2.0 and 2.9; >90% demonstrated changes between 2- and 10-fold.

APAP toxicity as revealed by gene expression analysis was even more extensive than toxicity described in a similar proteomic study conducted by Fountoulakis et al. (2000). The results of the oligonucleotide microarray were reported to be consistent with reverse transcriptase–polymerase chain reaction (RT-PCR) estimates of fold changes for selected genes within several of the assigned functional clusters. These clusters included, among others, stress-responsive genes (n = 22); cell cycling and growth inhibition (n = 14); inflammation (n = 14); cell signaling (n = 18) and cell metabolism (n = 13).

As depicted in Figure 5.11, the drug-induced liver toxicity resulted in the induction of genes involved in stress and inflammation, while genes involved in cellular metabolism were down-regulated. What is missing from the report was information regarding the timing of expression. It would be of interest to see which genes responded first to APAP and how that temporal expression varied during the course of the hepatotoxicity. Such time-course snapshots may have provided useful information regarding the off-target mechanisms of action of APAP.

Katsuma et al. (2001) undertook the time-course approach to help elucidate factors involved in drug-induced lung fibrosis. They followed histopathological changes associated with bleomycin-induced pulmonary fibrosis in a mouse model and correlated these with the gene expression profiles obtained using a cDNA microarray. A so-called lung chip was prepared with spotting of 4224 cDNAs obtained from a normalized lung cDNA library. Mice were subjected to bleomycin and sacrificed 2, 5, 7, and 14 days following intratracheal instillation.

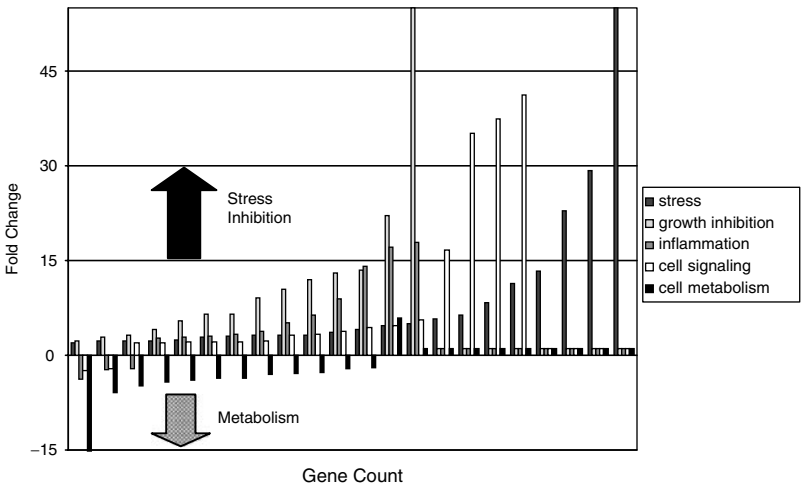


Figure 5.11 Gene expression changes associated with acetaminophen toxicity in liver. (From Reilly, T.P. et al., *Biochem. Biophys. Res. Commun.*, 282, 321–328, 2001. With permission.)

Lungs were fixed in formalin and embedded in paraffin and thin sections examined for pathology. While left lungs were reserved for pathology, right lungs were used for RNA preparation. Two micrograms of poly(A)$^+$ RNA was used to prepare labeled cDNA by incorporation of Cy3- or Cy5-deoxyuridine triphosphates (dUTPs) into control and test populations, respectively. Approximately 80% of the array was found to hybridize to the labeled cDNA targets from lungs. A twofold change in expression level was scored as significant.

Of the 4224 cDNA clones, a total of 159 (82 nonredundant genes) were observed to be differentially expressed over the 14-day period. Most genes were found to be up-regulated by day 5 and very few genes down-regulated. The differentially expressed genes were divided into four clusters for the purpose of analysis:

Cluster 1: genes up-regulated at 5 days, then returning to basal level
Cluster 2: genes maximally up-regulated at day 5, then remaining up-regulated
Cluster 3: genes continuously induced over 2 to 14 days
Cluster 4: genes down-regulated after bleomycin administration

How do these groupings correlate with the phenotypic changes encountered? Pulmonary fibrosis involves the accumulation of collagen, the growth of fibroblast cells, and thickening of the lung septa. Cluster analysis revealed that genes involved in inflammatory response (e.g., complement C3, osteopontin) were induced in the early stages of the disease, followed

by the appearance of genes related to fibrosis activity (collagen, fibronectin) occurring at a later stage in the process. Therefore, the authors postulate that bleomycin-induced fibrosis may first involve an inflammatory response in which the damaged lung is repaired by laying down of additional extracellular matrix proteins such as collagen, thereby leading to the progression of fibrosis pathogenesis.

Cancer

One of the most extensive undertakings employing DNA microarrays for research on cancer models was achieved by Ross et al. (2000). Sixty tumor-derived cell lines (NCI 60) were profiled for alterations of gene expression when subjected to anticancer drugs. These cell lines maintained under the National Cancer Institute's Developmental Therapeutics Program had been assessed for drug sensitivity against over 70,000 compounds. The cDNA slide microarray used in this study comprised over 8000 different cDNAs of which 3700 were well identified to previously characterized human protein (gene) products. The other elements on the array were ESTs (2400) and homologs (1900) from other organisms. Essentially, about 80% of the genes were correctly identified in the cDNA clones.

The experimental design is rather straightforward. The mRNAs from all test cell lines were converted to Cy5-labeled cDNAs by RT. They were compared against a reference of pooled mRNAs from 12 separate cell lines that had been reverse transcribed to produce a Cy3-labeled cDNA reference. This pool represented the maximum diversity across the 60 cell lines and was used as the reference throughout the study.

Thus the hybridization signal intensity ratio Cy5/Cy3 served to normalize for each cell line permitting comparison across all 60 cell lines. To minimize differences due to culturing conditions and cell density, cells were grown out to 80% confluency and mRNA isolated 24 hours after transfer of the cell line into fresh media. Results were clustered in hierarchical fashion to allow groupings of similarly expressed genes in relation to their tissue of origin. However, approximately 1200 genes exhibited wide variation in expression across the 60 cell lines, and these were of the most interest (Figure 5.12).

What did this study accomplish? First, the clustering approach was found to be generally valid in that cell lines derived from the same tissue were found to group together while those from different tissues appeared in separate branches. For example, colon- and ovarian-derived cell lines were observed in separate branches. On the other hand, breast tumor cell lines were found to distribute across several branches, suggesting a higher degree of heterogeneity in gene expression.

Other cell lines such as those from melanoma tissue exhibited clusters containing as many as 90 genes with high levels of expression, many of which are involved in melanin metabolism. When cell line expression patterns were found among different tissues of origin, it was also observed that the patterns

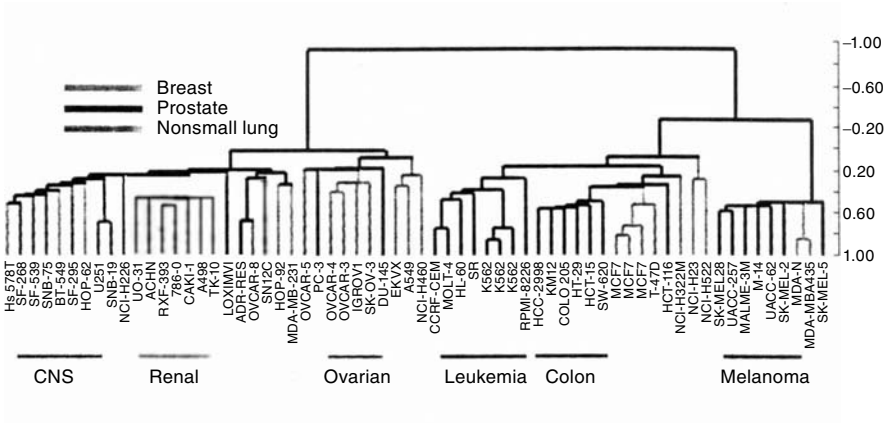


Figure 5.12 Hierarchical clustering of expressed genes from the NCI 60 tumor cell line. (From Ross, D.T. et al., *Nature Genetics*, 24, 227–235, 2000. With permission.)

carried with them genes from the different tissues, that is, one cell line may have shown a relationship with epithelial-derived cells while another cell line may have been more stroma cell-like in its gene expression pattern.

A gene cluster comprising cells derived from colon carcinoma, ovarian, and breast cancer shared genes involved in formation of the basolateral membranes of epithelial cells. In another branch were found all glioblastoma-derived cell lines in which many clustered genes were associated with stroma cellular function. Yet, no one gene was distributed across all of the clustered cell lines. Each cell line held a characteristic expression pattern related to the regulation of the extracellular matrix proteins. This is a good case in point to support the rationale for performing hierarchical clustering. A single gene expression pattern is generally not sufficient to track down relationships.

Unger et al. (2001) examined the gene expression profiles of adjacent breast tumors using microarray analysis. The occurrence of multiple tumors in tissue is problematic for diagnosis and appropriate treatment because of the potential for metastasis. The tumors may have originated from different primary tumors from different tissues. In those cases, treatments of the individual tumors may have to be different or compromised. Current methodology based upon X chromosome polymorphisms and used to check for tumor clonality apparently are limited in their scope to assess the genomic characters of individual tumors. For this reason the value of examining the gene expression profiles was assessed using clinical samples.

The Affymetrix U95A GeneChip containing 12,500 known human genes was used for this purpose. Two adjacent breast tumors from an 87-year old woman were removed and used in the study along with an additional five breast tumors obtained from different individuals. Thus, seemingly related tumors from one individual could be compared with tumors from different sources to validate the microarray-based prognosis. In addition to validation of the microarray for this purpose, the researchers examined a number of

important aspects of microarray performance. We will first discuss the issue of system performance that ultimately impacts the clinical interpretation presented by this study.

First, the total cRNA representation on the chip was found to approximately 50% of the available gene content. This, of course, means that the chip could not read half of the sample's content. While arguably assessment of 50% of the gene population is significant, it is also difficult to accept the fact that we are, from an analytical perception, starting at a disadvantage by not being able to see the complete profile. Simply put, significant information that could support or invalidate the author's conclusions may as likely lie within the missing genes.

The chip-to-chip variations between duplicate hybridizations of the same sample were low, resulting in a pair-wise correlation of 0.995 for both adjacent tumors. The small difference (0.5%) between duplicates represented 50 genes from tumor 5A and 36 genes from tumor 5B. None of the 36 genes found at a twofold expression for 5B replicates was found among the 50 genes expressed by 5A replicates.

Examining the two adjacent tumor samples (5A and 5B), 149 genes were differentially expressed in common (r = 0.987). However, based upon the level of variation in duplicates, some of the observed expression could be attributed to experimental noise. Couple this with a sample-to-sample variance (r = 0.915) reported larger than the observed chip variance (for tumor 5A, 50 genes from hybridization replicates and 100 genes differentially expressed between sample replicates), it would be difficult to believe that the tumor expression profiles were the same or different.

The authors suggest that because the number of differentially expressed genes between the two tumors was less than the number of genes expressed from replicates of the same sample, the tumors were virtually indistinguishable. I would argue that the results were inconclusive. As the authors note, "a simple pair-wise correlation comparison may not fully represent the relationship between gene expression profiles." In fact, when the researchers focused on those genes implicated in breast cancer, significant differences among various tumors that were not evident based upon pathology were revealed via microarray analysis (Table 5.1).

Wang et al. (2000) wished to identify new cancer markers for potential use as antigens in tumor-specific immunotherapy, especially for treatment of lung squamous cell carcinoma (LSCC). Nonsmall cell lung carcinoma comprises about 80% of all lung cancers; the 5-year survival rate is less than 10%. LSCC is a member of this disease pathology that currently lacks sufficient markers for early diagnosis and treatment. Thus, the object of Wang's study was to use microarrays to identify overexpressed genes in lung tumors relative to normal lung tissue in the hope of finding candidate markers. Wang's group recognized that the presence of high abundance genes on the microarray would limit the representation of lower copy number genes. As a result, they sought to combine the process of subtractive hybridization that

Table 5.1 Genes of Known Interest in Breast Cancer

Gene	Tumor		
	5A/5B	5A/7	5A/9
ER	NC	D (12.7)	NC
PR	NC	NC	I (2.7)
Androgen receptor	NC	NC	D (2.5)
Epidermal growth factor receptor	NC	NC	NC
ERBB2	NC	NC	NC
VEGF	NC	I (2.0)	D (2.2)
TP53	NC	NC	NC
Ataxia telangiectasia (*ATM*)	NC	NC	NC
FHIT	NC	NC	NC
BRCA2	NC	NC	NC
RAD50	NC	NC	NC
BARD-1	NC	D (2.9)	D (3.5)
Retinoblastoma-1 (*RB1*)	NC	NC	NC
Amplified in breast cancer (*AIB1*)	NC	NC	NC
Breast cancer transcription factor (ZaBC1)	NC	NC	NC
Thymidylate synthetase	NC	D (3.0)	NC
Multidrug resistance (*MDR-1*)	NC	NC	NC
Thrombospondin-1	NC	NC	D (2.2)
KI-67 antigen	NC	NC	NC
Breast epithelial antigen (*BA46*)	NC	I (4.3)	NC
Human mammaglobin	NC	I (3.6)	I (36.9)
Human mammaglobin β precursor	I (40.1)	NC	I (29.1)

Note: NC = No Change. D = Fold Decrease. I = Fold Increase.

Source: From Unger, M.A. et al., *Breast Cancer Res*., 3, 336–341, 2001. With permission.

would eliminate highly abundant transcripts while enriching for lower abundant genes of interest for use in the generation of microarray probes.

The subtractive hybridization process is outlined in Figure 5.13 and will be briefly described here. First, total RNA pools were isolated from normal and diseased tissues and the corresponding poly A^+ RNA (mRNA) purified. The mRNA was converted by RT to cDNA to create a *tester* cDNA library from LSCC tissue and a *driver* library from normal lung cells. The driver library was used to remove common and highly abundant transcripts from the tester population by first biotinylating the driver cDNA to be later bound to streptavidin and precipitated out. Tester and driver cDNAs were mixed with the driver (which was in excess of the tester cDNA) in order to efficiently capture the more abundant and common transcripts. Following mixing to hybridize the tester and driver, streptavidin was added and the mixture precipitated.

cDNA not binding to the streptavidin–biotin driver was isolated and thus enriched after several rounds. The enriched tester cDNA [lung squamous tumor-specific subtracted cDNA libraries (LST-S1…)] were then ligated

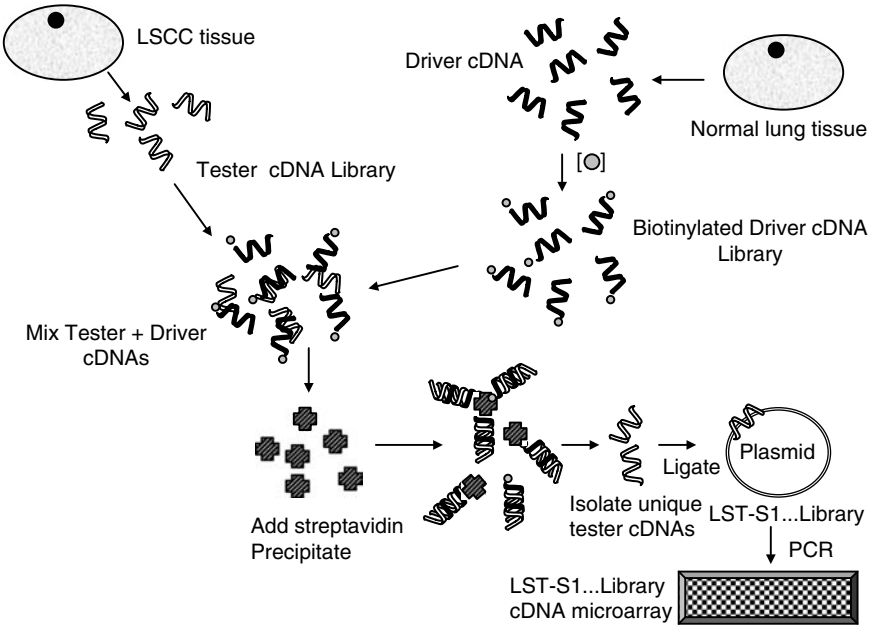


Figure 5.13 Subtractive hybridization. (From Wang, T. et al., *Oncogene*, 19, 1519–1528, 2000. With permission.)

into plasmids and amplified by PCR. The LST-S1, -S2, and -S3 donated the various levels of step-wise subtractive hybridization and enrichment of tumor-specific clones. The resulting cDNA amplicon-subtracted library (LST-S1 …) could then be spotted down on a glass slide microarray to create tumor-specific gene probes. The relative gene expression levels of genes present in tumors could then be assessed and overexpressed genes identified as potential cancer markers.

Wang et al. were able to isolate approximately 2000 cDNAs as LST-S1, -S2, and -S3 libraries and prepare probes. Using the competitive hybridization dual-labeling (Cy3/Cy5) method, a total of 17 genes were differentially overexpressed in LSCC. This included 13 known genes and 4 unknown genes. Figure 5.14 is a graph of normal vs. squamous cell tumor results; Table 5.2 lists the 17 genes. These expression results were confirmed by Northern analysis and real-time RT-PCR. The genes could also be classified as tissue or tumor-specific markers. Tumor-specific genes comprised cell signaling, enzymes, and antigens, while tissue-specific genes included cytoskeletal and squamous cell-specific markers.

Chen et al. (2001) validated the use of cDNA microarrays for gene expression studies involving an established cell line model for lung metastasis. Recognized as the leading cause of mortality in cancer patients, metastasis is a complex process in which cancer cells from the primary tumor move into (invade) other tissues and organs. Cell-to-cell interactions and influences

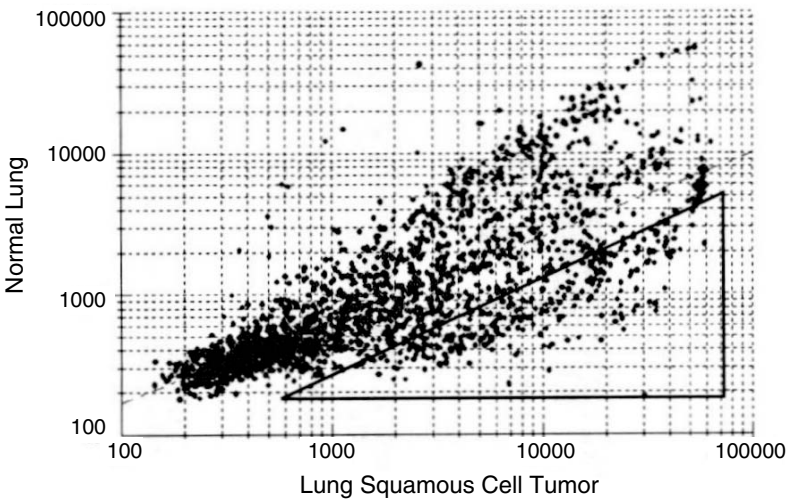


Figure 5.14 Lung squamous tumor-specific subtracted cDNA libraries. (From Wang, T. et al., *Oncogene*, 19, 1519–1528, 2000, Macmillan Publishers Ltd. With permission.)

from the surrounding microenvironment are important factors governing tumor invasiveness. Certain molecules such as laminin receptor and CD44 are known to promote metastasis, while others such as cadherin are inhibitory. However, genetic instability of cancer cells has proven to be a challenge in the further identification of genes involved in metastasis. Chen and coworkers (2001) used human lung adenocarcinoma cell lines (CL1-0 and sublines, CL 1-1 and CL 1-5) varying in metastatic potential and assessed their gene activities under defined growth conditions using microarrays.

The microarray comprised 9600 nonredundant ESTs from the integrated molecular analysis of genomes and their expression (IMAGE) collection of human cDNA clones arrayed onto nylon membranes. Of these, 1875 clones (19.5%) were verified by resequencing; 110 of the 589 genes expressed (18.7%) that correlated with the metastatic potential were among those verified by sequencing.

Invasiveness of the chemiluminescence (CL) lines was measured by *in vitro* and *in vivo* methods. The *in vitro* monitoring process comprised the movement of cells across a membrane of defined pore size within a specially designed growth chamber or MICS (membrane invasion culture system). A 10-μ diameter Nucleopore membrane was coated with a mixture of laminin (to promote invasion), collagen, and gelatin. Cells were added to the top side of the chamber in media and the extent of cell movement into the bottom of the chamber (invasion) through the membrane determined.

The *in vitro* invasive assay consisted of grafting rat tracheas by implanting into severe combined immunodeficiency disease (SCID) mice. The tracheas were first injected with the adenocarcinoma cells to promote tumor growth. Invasiveness was determined by histochemical staining of the iso-

Table 5.2 Lung Squamous Cell Carcinoma Differential Gene Expression

cDNA Clone	Gene	Functionality	Hits	Northern	RT-PCR	Specificity
520	SPRC	Cell marker	1	++		Tissue
513	PVA	Cell marker	2	+++	+++	Tissue
521	SPR1	Cell marker	2			Tissue
525	Plakophilin	Cytoskeleton	2		+	Tissue
527	Cytokeratin	Cytoskeleton	2			Tissue
529	Connexin	Cytoskeleton	2			Tissue
516	ARH	Enzyme	1			Tumor
523	KOC	Antigen	1	+++	+++	Tumor
524	PTHrP	Cell signaling	1	0	+++	Tumor
526	ATM	Cell signaling	1			Tumor
515	IGF-2	Cell signaling	2			Tumor
522	ADH7	Enzyme	5	0	++	Tumor
528	NMB	Antigen	14		+++	Tumor
514	Novel	Unknown	1	+++	+++	Unknown
531	Novel	Unknown	1	0	++	Unknown
530	Novel	Unknown	2	+++	++	Unknown
519	Novel	Unknown	10	++	++	Unknown

Source: From Wang, T. et al., *Oncogene*, 19, 1519–1528, 2000, Macmillan Publishers Ltd. With permission.

lated tracheas to reveal tumor growth and invasion of epithelial cells into the basement membrane. Satisfied with the ability to characterize invasiveness under defined conditions, the researchers then examined the gene expression profiles of these cells using the microarray.

Isolated mRNAs from cell lines were biotin-labeled, hybridized, and signal-generated using a colorimetric reagent and 8525 of 9600 "genes" showed significant levels of expression. They were arranged in 100-SOM (self organizing map) clusters, of which four clusters correlated with the promotion cell invasiveness (277 genes) and an additional four clusters (312 genes) correlated with inhibitory effects (Figure 5.15). These gene clusters were then rearranged for hierarchical clustering across the four cell lines used in the study and grouped in terms of cellular function, e.g., adhesion molecules, motility proteins, cell cycle regulators, signal transduction, and angiogenesis-related functions.

Many of the genes identified in this study were found to be consistent with results from other reports that recognized these genes as participating in various aspects of metastasis including the role of angiogenesis in blood vessel formation. One important outcome of this study was the finding that the tumor-associated surface antigen L6 was highly associated with tumor invasion in this lung metastasis model. L6 appears to be highly expressed in human lung, breast, and colorectal carcinomas and may serve as a useful diagnostic marker.

The results of the microarray gene expression studies from the CL cell lines were also confirmed by Northern analysis and the application of flow cytometry, providing an additional level of confidence to the data. For Northern analysis, five sequence-verified genes and 5 ESTs found to have positive correlations for metastasis were selected for comparison across the four cell lines. In addition, five known genes and one EST having a negative correlation were examined. The results of the 16 genes were found to be in agreement with the relative gene expression levels determined by the microarray experiments.

Finally, fluorescently labeled monoclonal antibodies directed toward L6 antigen, integrins α-3 and α-6 were used to interrogate the four cell lines. Antibody-labeled cell populations were analyzed in a flow cytometer. In confirmation of both the microarray and Northern blot analyses, the IL-6 antigen and integrins were most prominent in the cell line demonstrated to have the greatest degree of tumor invasion. In summary, this study represents one of the most thorough attempts at validating the utility of microarrays for clinical research: characterizing tumor cell line invasiveness by two independent methods with histochemical verification; performing replicate microarray experiments and cluster analysis; and then confirming the microarray results on gene expression at both the transcriptional and translational levels.

While the work of Chen et al. (2001) broadly defined the expression of genes associated with metastatic potential in the lung cancer model, others have more narrowly focused upon relationships for specific genes. For example, Pinheiro et al. (2001) examined gene expression profiles from patients

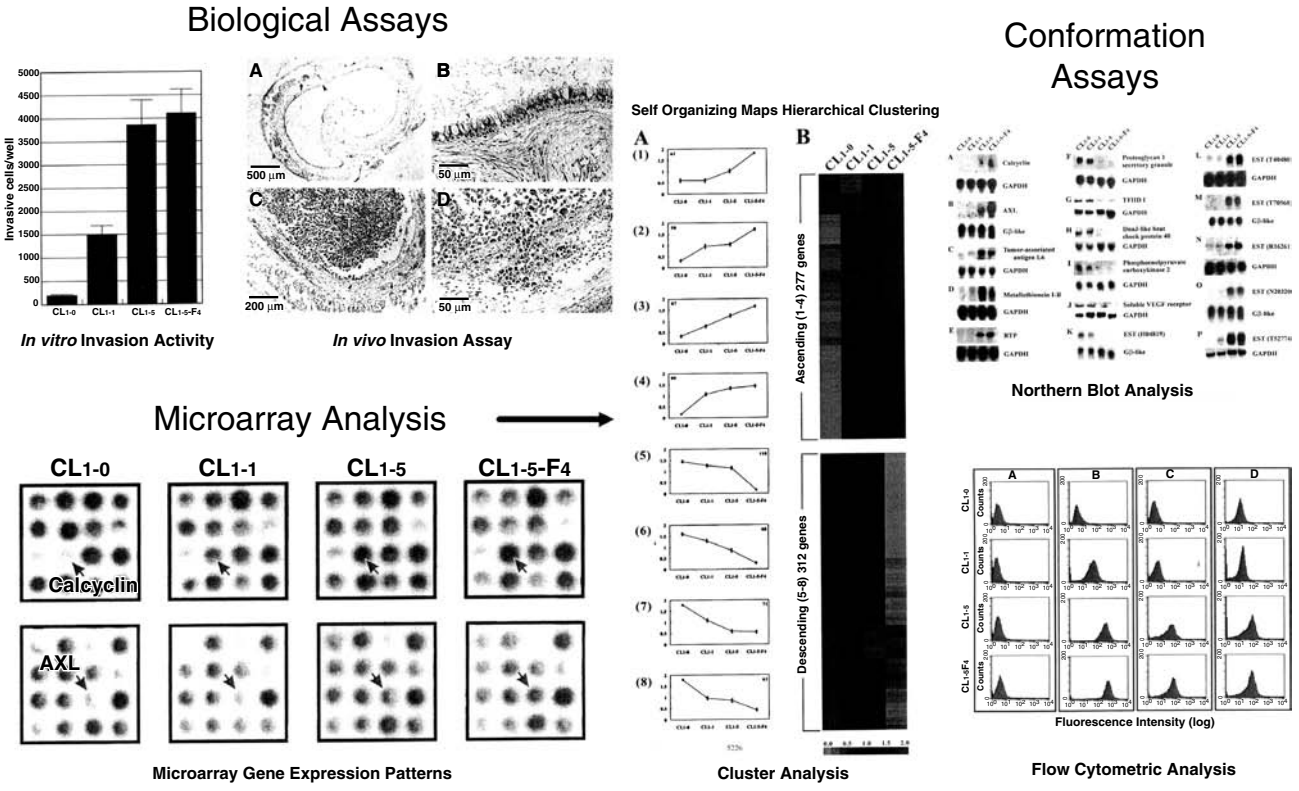


Figure 5.15 Validation of microarray-based gene expression analysis. (From Chen, J.J.W. et al., *Cancer Res.*, 61, 5223–5230, 2001. With permission.)

with colorectal adenocarcinomas and paired normal tissues. Using microarray filters comprising 18,376 cDNAs (Incyte-Genomics, Wilmington, DE), they focused on the overexpression of a single gene, oligophrenin-1.

Data from three independent microarray experiments showed changes from ~10- to 100-fold, depending upon the tumor pool used. The results confirmed by RT-PCR indicated that oligophrenin-1 was consistently overexpressed in colorectal tumors but not significantly detected in normal tissue. The sequencing of 10 PCR products verified a 100% identity with the gene. A subsequent inquiry of the SAGE (Serial Analysis of Gene Expression) database (National Cancer Institute, Cancer Genetic Anatomy Project) revealed overexpression in a prostrate cancer cell line as well. Surprisingly, the authors found no reports on the overexpression of this gene associated with colorectal tumors.

Oligophrenin-1 is known to be involved in X-linked mental retardation and encodes a protein having a rho GTPase-activating protein (rhoGAP) domain (Billuart et al., 1998). While at first glance mental retardation and colorectal adenocarcinoma would be seemingly unrelated, rho GTPase is involved in the regulation of rho and ras proteins (Chen et al., 2003). Activation of the K-ras oncogene is well known for its involvement in colorectal cancer (Matson, 2000).

Mullan et al. (2001) used an oligonucleotide G110 array (Affymetrix) comprising 6800 genes and ESTs to examine expression profiling in BRCA1-induced cell lines in an effort to identify downstream targets. The G110 array contained approximately 1700 cancer-associated genes.

The BRCA1 mutation is associated with the occurrence of 10% of all human breast cancers and thereby is implicated in the predisposition of breast and ovarian cancers. The mechanisms by which the tumor suppressor gene acts upon other genes are not well understood. In order to study these effects, two cell lines (one derived from an osteosarcoma cell line and the other from a breast cancer cell line) were established to exhibit inducible, tetracycline-regulated expression of BRCA1. The exogenous gene could be switched off with the addition of tetracycline (+ tet) or turned on by the removal (– tet) of the antibiotic without significantly altering genomic background expression levels.

In reality, 23 of 6800 genes (0.3%) were induced during the tet switch. However, the DNA damage-inducible gene [growth arrest after DNA damage (GADD45)] was found to be expressed 10-fold during BRCA1 induction (77-fold). Northern blots confirmed the increase in GADD45 expression 6 hr following BRCA1 induction in both cell lines. Fold changes in expression with GADD45 and BRCA1 showed a linear correlation ($r^2 = 0.96$) over 24 hr following the tet switch (Figure 5.16). This was confirmed over the same time course by Northern blot analysis. These results strongly suggest that GADD45 is a transcriptional target of the BRCA1 gene.

Bouras et al. (2002) undertook a comprehensive study of differential expression focused upon genes associated with estrogen responsiveness (ER+) in a human breast cancer cell line correlated with clinical tumor

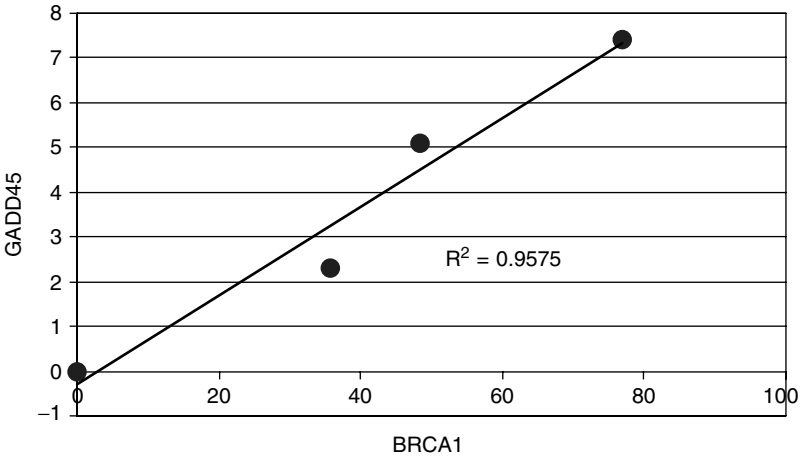


Figure 5.16 Association of DNA damage-inducible gene (GADD45) expression during tetracycline induction of the BRCA1 gene. (From Mullan, P.B. et al., *Biochem. Soc. Trans.*, 29, 678–683, 2001. With permission.)

samples. First, ER+ human breast cells (MCF-7 cell line, American Type Culture Collection, Manassas, VA) were grown in the presence of 17β-estradiol or an antiestrogen (ICI 182 780), and gene expression was measured from the mRNA populations. In addition, mRNAs from 25 primary tumors (13 ER+ and 12 ER–) were analyzed.

An Affymetrix 43K GeneChip set comprising 10,000 known genes and 25,000 ESTs was used to measure differential gene expression. The outcome of the cell culture experiments was that 299 genes (<1%) were significantly ($P \leq 0.0005$) regulated by estrogen and/or antiestrogen. The expression profiles for these 299 genes were then assessed in the 25 primary tumors and the 10 most highly differentially expressed subjected to hierarchical cluster analysis. Among these 10 genes, stanniocalcin-2 (STC2) was singled out as a potential diagnostic candidate for determining the estrogen responsiveness in breast tumors based upon:

1. STC2 differential expression between ER+ and ER– breast tumors
2. Correlation of STC2 mRNA levels with ER mRNA and protein levels
3. A threefold expression of STC2 in estrogen-stimulated MCF-7 cells within 3 hr and continuation of elevation up to 48 hr
4. A threefold reduction in expression of STC2 following antiestrogen treatment of cells with 6 hr of administration

Of particular interest was the observed strong correlation obtained between STC2 mRNA levels and its cognate protein levels. Using tissue microarrays from 216 breast tumor samples, *in situ* hybridization with a probe to STC2 mRNA was performed with 75 tumors and showed positive staining; 83% of the mRNA-positive tumors identified for STC2 protein by

immunohistochemical staining with STC2 antibody. Thus, microarray-based expression profiling of STC2 was corroborated by both *in situ* hybridization and immunostaining for ER+ tumors. A related gene, STC1, showed a similar correlation.

The study by Bouras et al. (2002) clearly demonstrated a rather strong correlation between transcription and translation of a single gene, stanniocalcin-2. The result is a high likelihood that STC2 may serve as a diagnostic or prognostic marker for breast cancer because it enables monitoring of both mRNA levels and protein products during the various stages of tumor growth.

In summary, we find gene expression profiling with microarrays to be an exceptionally powerful and profound analytical tool. Not only is the technique useful for global analyses, for example, of metabolic pathways and their interrelationships, but it also has the ability to focus upon (albeit assisted by clustering) and track important singular events that would otherwise remain hidden under a genomic backdrop.

In general, how well does gene expression (as measured by microarray) mirror biological outcome since microarrays offer only a transient (global) view of biology? Cellular function on the other hand is mostly the work of proteins, and it is well known that posttranslational modifications are important regulators. We now have the equivalent protein expression microarray tools available. Can we therefore rely upon the gene expression microarray as a surrogate tool to adequately track cellular endpoints from gene activities? Many advocate that gene expression does not correlate with protein expression, but in reality this is too simple an answer. The answer is more complicated.

Take for example the work of Chen et al. (2002) that involved an effort to correlate mRNA levels with protein expression levels in lung adenocarcinomas. They compared 57 stage I and 19 stage III lung adenocarcinoma tissues along with 9 non-neoplastic lung tissues. Gene expression was measured using the Affymetrix GeneChip HuGene FL oligonucleotide arrays comprising 6800 known human genes. Protein was estimated by spot densitometry after 2D polyacrylamide gel electrophoresis (PAGE) separation and silver staining. Proteins were identified from matrix-assisted laser desorption/ionization mass spectroscopy (MALDI-MS) of peptides obtained by tryptic digestion of protein spots from preparative 2D gel separations of extracts of a well-characterized lung adenocarcinoma cell line, A549.

Certain proteins were confirmed by Western blot analysis. The 2D PAGE analysis can resolve up to 2000 proteins. In this study, 820 spots were mapped to proteins but only 165 were used in the protein-to-gene analysis. Presumably these represent the highest fold expressed genes from the lung tumors that were also visible on the gels by silver staining.

Several approaches to the analysis were undertaken to achieve correlation: pair-wise individual protein to gene; protein isoform to gene; average protein to gene expression value; and tumor stage-related changes, protein to gene. As outlined in Figure 5.17, of the 165 protein spots 98 represented

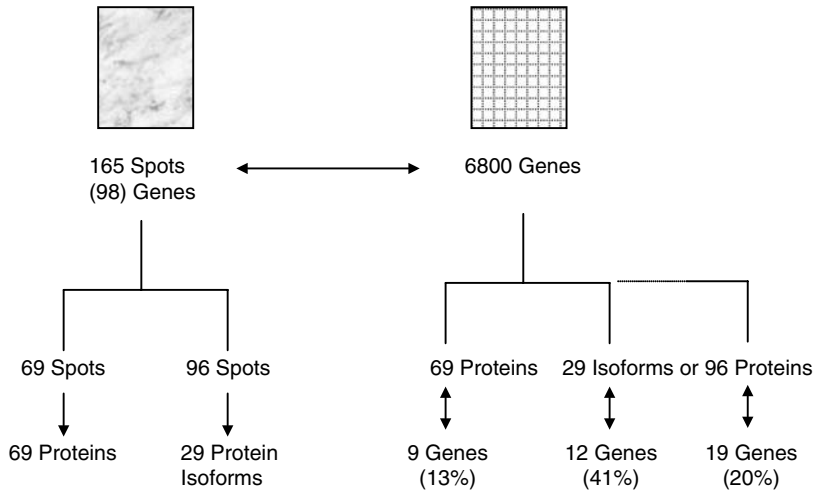


Figure 5.17 Correlations between mRNA and protein relative abundance. (From Chen, G. et al., *Mol. Cell. Proteomics*, 1, 304–313, 2002. With permission.)

genes. Of these, 69 proteins were mapped to single genes and 96 proteins appearing as two to five isoforms were mapped to 29 genes. However, only 9 of 69 (13%) of the single gene–protein pairs were significantly correlated ($P < 0.05$) in terms of relative abundance. For those proteins existing as isoforms, 19 of 96 (20%) showed correlation in relative abundance to gene expression level. Thus, ~29% (28 of 98) of the genes showed good correlations between mRNA abundance and protein abundance.

Proteins appearing as isoforms varied in their relationships to mRNA abundance. This most likely was a result of posttranslational modification of the nascent protein and possible degradation of the message. However, it is very difficult to generalize because subtypes showed different levels of correspondence to mRNA abundance. For instance, three of four isoforms of OP18 were significantly correlated, while one isoform did not correlate. In the case of cytokeratin-8, only one of five isoforms correlated. What was very clear from this study is that a generalized approach such as comparing averages of protein levels and mRNA abundance levels does not work. While many genes did not vary in their transcriptional-to-translational correlations relative to tumor stages, specific genes did show differentiation on the bases of tumor stages. Our discussion of these particular papers indicates that we cannot generalize about the relationships of transcriptional, translational, and posttranslational events.

Infectious disease

While the monitoring of infectious biological agents (bacteria and viruses) using DNA probe array technology is well known, many of the applications involve identification of allele-specific targets. Conversely, examining

differential gene expression in a host is useful to elucidate mechanisms leading to virulence.

For example, Lane et al. (2001) examined gene expression in *Candida albicans*, a common yeast leading to human commensal infections. Pathogenicity is related to morphological changes in the organism characterized by a switch from the yeast form to the filamentous hyphal form. These changes in phenotype are obviously under control of various regulatory genetic elements and should be recognizable by comparing the gene expression patterns of the two morphological states. For example, signaling pathways are involved in the regulation of the filamentation process. In particular, the Cph1 transcription factor is known to be involved in hyphal formation. Cph1 is regulated by the mitogen-activated protein (MAP) kinase pathway. Other regulatory elements include an Egf1-mediated cAMP protein kinase A and a Cph2 pathway, but how these multiple cascade pathways act to shift growth into the hyphal state is uncertain. Do they act independently or sequentially or are they convergent? What other genes may be important?

To more fully understand the mechanisms leading to the morphological switch, an array of PCR products representing ~1000 *C. albicans* genes (~10% of the genome represented by 7000 ORFs) was constructed on nylon membranes in triplicate. Gridding was accomplished using a Biomek 2000 equipped with a 384-pin HDRT (high density replicating tool) system. Each spot was overprinted three to five times to assure full surface saturation and uniformity at each probe location. Membranes were then UV cross-linked prior to use. Each membrane could be stripped in boiling 0.5% sodium dodecyl sulfate (SDS) and reprobed up to seven times without affecting its signal. As a result, the same filter sets could be used for multiple experiments with appropriate controls, thereby eliminating effects due to interfilter array variations.

A series of yeast mutants grown under conditions selectively inhibiting or promoting growth of hyphal forms were compared with the wild type. For example, Lee's medium at 25°C promotes the yeast form; at 37°C, it induces hyphal growth. In addition, SS (synthetic succinate) medium permitted the transformation of the yeast form into the hyphal form for the wild type but not for the single cph1/cph1 mutant. We know from other work that double mutants (cph1/cph1 efg1/efg1) are not virulent in mouse models while the single mutants cph1/cph1 and efg1/efg1 are virulent (hypha form).

These mutants were grown under the above conditions and their gene expression profiles compared with the wild type yeast form using the grid arrays. From this detailed study comparing mutant and wild type strains under a variety of growth conditions, it was found that Cph1, Cph2, and Efg1 regulations are convergent (Figure 5.18). On the basis of cluster analysis (Figure 5.19) and conformation by Northern blot analysis (Figure 5.20), these genes were found to regulate the expression of a set of hypha-specific genes, e.g., HYR1, ECE1, HWP1, SAP5, and SAP6 (Figure 5.21).

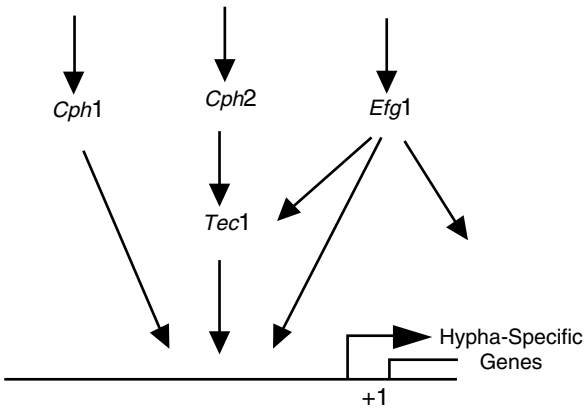


Figure 5.18 Convergent regulation of *Candida albicans* genes during switch from wild type to virulent form. (From Lane, S. et al., *J. Biol. Chem.*, 276, 48988–48996, 2001. With permission.)

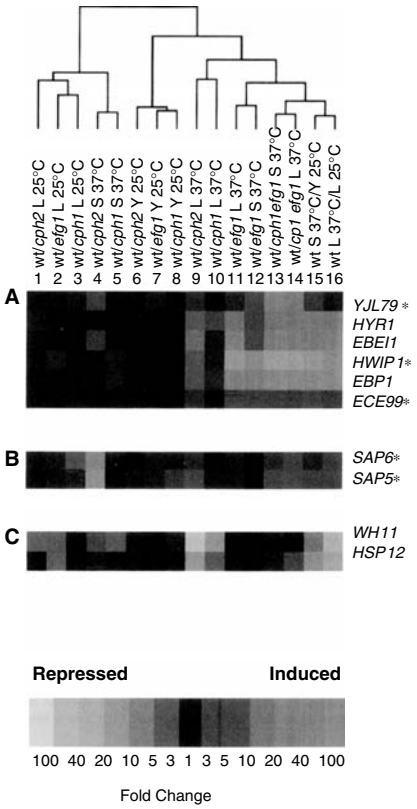


Figure 5.19 Gene expression cluster analysis. (From Lane, S. et al., *J. Biol. Chem.*, 276, 48988–48996, 2001. With permission.)

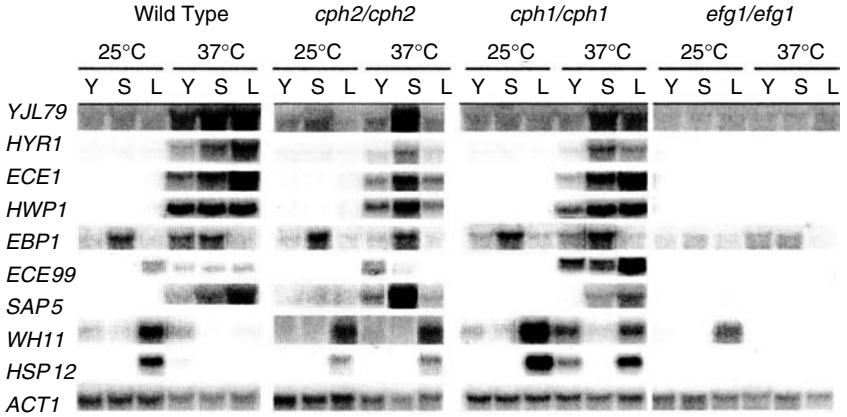


Figure 5.20 Northern blot analysis. (From Lane, S. et al., *J. Biol. Chem.*, 276, 48988–48996, 2001. With permission.)

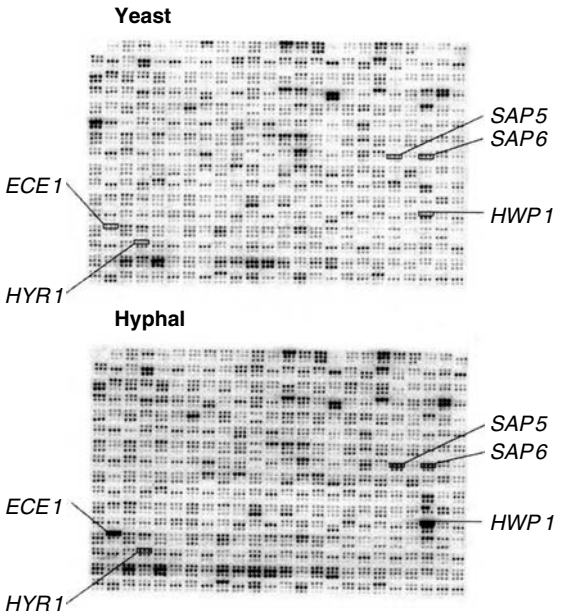


Figure 5.21 Comparison of wild type vs. hyphal (virulent) gene expression patterns. (From Lane, S. et al., *J. Biol. Chem.*, 276, 48988–48996, 2001. With permission.)

In addition, two new genes (DDR48, YPL184) were discovered to be differentially expressed and under regulation by the convergent pathways, while the key regulator TEC1 gene was found to be under the influence of Cph1 and Egf1. Thus, even with the limited representation of the *C. albicans* genome on the array, it was possible to probe into the regulation of virulence factors using a combination of well designed biological approaches

(mutations, growth media, temperature shifts) and gene expression tools (DNA arrays, Northern blots, clustering algorithms).

In a similar manner, Maeda et al. (2001) examined *Helicobacter pylori* alteration of gene expression in gastric cancer cells. *H. pylori*, a Gram-negative bacterium, is well known for its infection of the human gastric mucosa. However, the pathogenesis of the associated gastroduodenal disease (e.g., peptic ulcers) in the host is not well understood. The activation of various transcription factors such as nuclear factor B (NF-B) and induction of such inflammatory cytokines as IL-8 have been implicated. The activation of NF-KB in turn is believed to be under the regulation of the cag PAI (pathogenicity island) genes found in a high percentage of *H. pylori* strain isolates.

cDNA microarray gene expression analysis was utilized to further investigate *H. pylori*-mediated induction of signaling pathways leading to an inflammatory response in the host. The glass slide microarray of 2304 cDNAs comprised a human cDNA library of 2280 sequence-validated cDNAs (Research Genetics now Invitrogen, Carlsbad, CA) along with a number of housekeeping genes as internal controls and luciferase genes as negative controls. RT-PCR and Northern blots were used for conformation of the microarray results.

A human gastric cancer cell line, MKN 45, derived from a gastric adenocarcinoma was cocultured with either an *H. pylori* strain (cag PAI-positive) or a cagE-knockout strain. Of the 2300 genes, only 8 (0.3%) were considered up-regulated at more than a twofold expression change. These genes were not differentially expressed in the knockout strain, suggesting the importance of the cag PAI involvement. IL-8 showed the greatest level in fold expression (11.8 infected:control) followed by IκBα (5.0-fold). IκBα protein binds to NF-κB. In order to activate NF-κB, the phosphorylation of IκBα is required. Phosphorylated IκBα is subsequently degraded, resulting in the release of active NF-κB from the complex. NF-κB is regarded as one of the major transcriptional factors for IL-8 induction which leads to inflammatory response in the gastric mucosa. Of the remaining six genes showing at least twofold expression, the A20 gene (2.2-fold) was viewed as an important discovery in understanding the pathogenesis of *H. pylori*. The A20 protein inhibits NF-κB activation.

In summary, the gene expression analysis of the gastric cancer cell host revealed that a small but significant set of genes associated with inflammatory response within the host were induced by the bacterium. As was the case with the work of Lane et al. (2001), the analysis of differential gene expression between normal and diseased states uncovered key regulatory genes in specific pathways involved in pathogenesis.

Other disease states

While it would be difficult to provide a comprehensive overview of microarray applications, the following survey should provide ample proof of the continued expansion of microarray use into diverse fields.

Hearing loss

Lomax et al. (2000) explored the potential for microarrays to investigate the involvement of genes in the recovery of hearing loss following noise trauma. The chick basilar papilla model was used. Noise exposure is known to cause the loss of hair cells in the basilar papilla. However, birds have the ability to regenerate these hair cells on the auditory epithelium and thus serve as useful models for studying hearing loss and recovery.

A low density microarray containing 588 genes arranged in subgroups according to tissue (Clontech Rat Atlas cDNA array nylon membranes) was used to first examine which genes represented on the array were present in the cochlea and auditory regions of the brain. Although preliminary, the investigation revealed three genes differentially expressed between two neuronal regions of the auditory system: the inferior cochlea (IC) of the brain and the cochlea modiolus (MOD). Two of the three proteins identified are known to be present in high abundance (mRNAs, moderate abundance class). Myelin proteolipid protein (PLP) is abundant in the brain, and peripheral myelin protein (PMP22) is localized in the peripheral nerves.

PLP was observed to be differentially expressed 2.5-times higher in the IC (brain) region than the MOD, while PMP22 was 5.8 fold higher in MOD. The third gene, the plasma membrane calcium-transporting ATPase (PMCA2), was expressed two-fold higher in IC. Mutations to this rare gene are associated with deafness and imbalance in mice. The fact that the microarray could detect significant levels of PMCA2 was an unanticipated find. As the authors relate, "This exciting and gratifying result suggests that gene arrays may have a profound impact on the analysis of differential gene expression in the mammalian auditory system."

Bone pathology

Apert (Ap) syndrome is a craniofacial malformation thought to occur by mutation of the fibroblast growth factor 2 (FGFR-2) gene. This mutation has been reported to increase osteoblast differentiation that leads to premature calvaria ossification. Lomri et al. (2001) used cDNA microarrays to elucidate signaling pathways involved in osteoblast differentiation. Calvaria osteoblast cells were isolated from the bones of normal and Ap human fetuses and transformed using (simian virus) SV40 large T antigen. The immortalized cell lines were maintained in culture and polyA$^+$ mRNA isolated from confluent cells. The resulting cDNA was radiolabeled by incorporation of ^{32}P-dATP during the RT reaction and hybridized to a cDNA nylon membrane array (Atlas Human Expression Array, Clontech).

The array contained 588 PCR cDNAs arranged in various human gene families. More than 40% of the 588 genes produced signals under high stringency hybridization conditions. Of these, 27 genes (22 up-regulated and 5 down-regulated) were differentially expressed in Ap vs. control cells.

In particular, significant differential expression was observed for three genes in Ap cells vs. control cells: GTPase rhoA (3.6-fold), protein kinase C-α (2.9-fold), and the IL-1α cytokine (3-fold). All three putative proteins were confirmed to be overexpressed in the Ap mutant cell line relative to control cells by immunohistochemical staining (IL-1α) or Western blot (protein kinase C [PKC], Ras homologous protein A [Rhoa]; a member of one of the subfamilies of low-molecular-weight GTPase proteins). These proteins are now implicated as serving effector roles in osteoblast differentiation resulting from the initial FGFR-2 mutation.

Glaucoma

Glaucoma is one of the most common diseases of the eye and leads to destruction of the optic nerve if left unchecked. Astrocytes, the major cell types in the optic nerve head, are believed to undergo phenotypic changes during the onset of glaucoma. Hernandez et al. (2002) used microarrays to monitor changes in gene expression patterns during phenotypic shifts in cultured astrocytes in an effort to better understand the pathogenesis of the disease.

Optic nerve head astrocytes obtained from normal or glaucomatous human eyes were maintained in primary cell cultures. Total RNA was extracted from the cultured cells, purified and converted to labeled cRNA for hybridization to the U95A Human Genome GeneChip (Affymetrix). Differential gene expression levels of normal and glaucomatous astrocyte populations were compared. In many studies of this kind, we reported the differential expression of relatively few genes between 2- and 10-fold. However, in this case, at least 99 genes were overexpressed by at least 5-fold in astrocytes from diseased eyes and 53 genes were either absent or down-regulated to the same degree.

In fact, many genes were found to be at least 25-fold differentially over- or under-expressed in reactive astrocytes (Figure 5.22). In particular, prostaglandin D2 synthase was differentially expressed by greater than 100-fold. While the mechanism of action of prostaglandin during glaucoma is uncertain, it is known that prostaglandin D2 (PGD2) is abundant in the central nervous system (CNS), especially in spinal fluid and ocular tissues. Astrocytes are known to proliferate at sites of neural damage and may be the sources of the synthase enzyme. Another important finding from this study was that steroid metabolism is significantly up-regulated in reactive astrocytes, suggesting a possible relationship between glucocorticoid metabolism and glaucoma.

Multiple sclerosis

Multiple sclerosis (MS) is a devastating disease of the CNS that produces demyelination and inflammation. It is believed that its pathogenesis involves an immune reaction against various components of the myelin sheath. In a study by Lock et al. (2002), the gene expression patterns of brain lesions obtained during autopsies of MS patients were examined by microarray cluster analysis. A total of 1080 genes (from 7026 represented

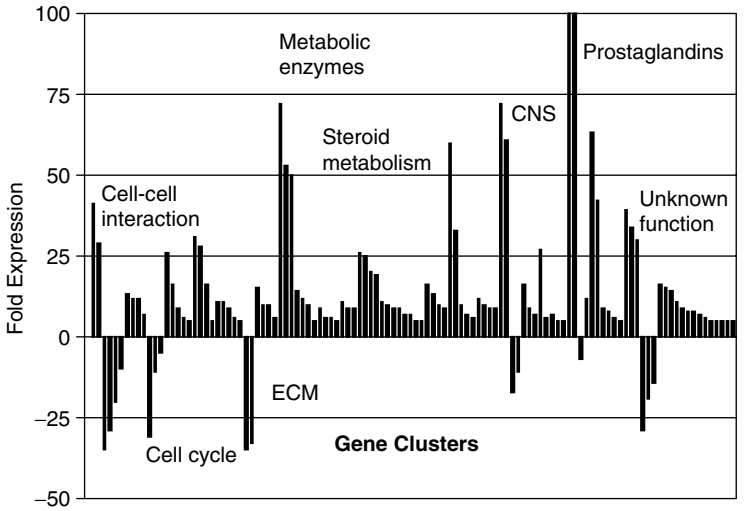


Figure 5.22 Investigating the pathogenesis of glaucoma based upon microarray gene expression clustering. (From Hernandez, M.R., *GLIA*, 38:45–64, 2002. With permission.)

on the GeneChip) were differentially expressed at twofold or greater levels. The following observations were made concerning differential gene expression in MS lesions (relative to normal tissue) from the cluster analysis:

1. Migration of lymphocytic cells — along with the presence of T cells, IL-17 a T cell transcript was found to be elevated.
2. Macrophage invasion — the up-regulation of the macrophage mannose receptor, cathepsin S, and macrophage capping protein.
3. Up-regulation of immune response genes — overexpression of major histocompatibility complex (MHC) and IgG.
4. Inflammatory cytokine activity — the IL-1 receptor and TNF receptor up-regulated.
5. Down-regulation of myelin synthesis pathway genes.

In addition, clustering revealed that several genes significantly were up-regulated and also differentially expressed for cases of acute and chronic MS.

In the next phase of the study, transcripts of up-regulated genes granulocyte colony-stimulating factor (G-CSF) and IgG were introduced into a mouse model commonly used to test potential therapies for experimental autoimmune encephalomyelitis (EAE). From microarray analysis, G-CSF was found to be up-regulated in acute MS but not in the chronic state of the disease. Subcutaneous injection of G-CSF prior to induction of EAE prevented onset of the disease in mice. The reversal of EAE by G-CSF has also been described (see Lock, 2002, Reference 40).

In the case of immunoglobulin, microarray data indicated the opposite effect: that the Fc-receptor is elevated in chronic MS but not in acute lesions. Using Fcγ-receptor knockout mice, the disease was found to be absent. Intervenous immunoglobulin therapy in the EAE mouse model was reported (see Lock, 2002, Reference 29). In summary, Lock et al. were able to apply the results of microarray-based gene expression clustering of a human disease pathological state (acute vs. chronic MS) to successfully identify therapeutic targets for an animal model (EAE) potentially applicable to the human condition.

References

Billuart, P., Bienvenu, T., Ronce, N., des Portes, V., Vinet, M.C., Zemni, R., Roest, C.H., Carrie, A., Fauchereau, F., Cherry, M., Briault, S., Hamel, B., Fryns, J.P., Beldjord, C., Kahn, A., Moraine, C., and Chelly, J., Oligophrenin-1 encodes a rhoGAP protein involved in X-linked mental retardation, *Nature* 392, 923–926, 1998.

Bouras, T., Southey, M.C., Chang, A.C., Reddel, R.R., Willhite, D., Glynne, R., Henderson, M.A., Armes, J.E., and Vender, D.J., Stanniocalcin 2 is an estrogen-responsive gene coexpressed with the estrogen receptor in human breast cancer, *Cancer Res*., 62, 1289–1295, 2002.

Chen, G., Gharib, T.G., Huang, C.C., Taylor, J.M.G., Misek, D.E., Kardia, S.L.R., Giordano, T.J., Iannettoni, M.D., Orringer, M.B., Hanash, S.M., and Beer, D.G., Discordant protein and mRNA expression in lung adenocarcinomas, *Mol. Cell. Proteomics*, 1, 304–313, 2002.

Chen, J.C., Zhuang, S., Nguyen, T.H., Boss, G.R., and Pilz, R.B., Oncogenic ras to rho activation by activating the mitogen-activated protein kinase pathway and decreasing rho-GTPase-activating protein activity, *J. Biol. Chem*., 278, 2807–2818, 2003.

Chen, J.J.W., Peck, K., Hong, T.M., Yang, S.C., Sher, Y.P., Shih, J.Y., Wu, R., Cheng, J.L., Roffler, S.R., Wu, C.W., and Yang, P.C., Global analysis of gene expression in invasion by a lung cancer model, *Cancer Res*., 61, 5223–5230, 2001.

Cleary, J.D., Rogers, P.D., and Chapman, S.W., Differential transcription factor expression in human mononuclear cells in response to amphotericin B: identification with complementary DNA microarray technology, *Pharmacotherapy,* 21, 1046–1054, 2001.

Dattagupta, N., Rae, P.M.M., Huguenel, E.D., Carlson, E., Lyga, A., Shapiro, J.A., and Albarella, J.P., Rapid identification of microorganisms by nucleic acid hybridization after labeling the test sample, *Anal. Biochem*., 117, 85–89, 1989.

DeRisi, J., Iyer, V.R., and Brown, P.O., Exploring the metabolic and genetic control of gene expression on a genomic scale, *Science*, 278, 680–686, 1997.

DeRisi, J., Penland, L., Brown, P.O., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y., Su, Y.A., and Trent, J.M., Use of a cDNA microarray to analyze gene expressioin patterns in human cancer, *Nat. Genetics*, 14, 457–460, 1996.

de Saizieu, A., Certa, U., Warrington, J., Gray, C., Keck, W., and Mous, J., Bacterial transcript imaging by hybridization of total RNA to oligonucleotide arrays, *Nature Biotechnol*., 16, 45–48, 1998.

Drmanac, R., Strezoska, Z., Labat, I., Drmanac, S., and Crkvenjakov, R., Reliable hybridization of oligonucleotides as short as six nucleotides, *DNA Cell Biol.*, 9, 527–534, 1990.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D., Cluster analysis and display of genome-wide expression patterns, *Proc. Natl. Acad. Sci. USA*, 95, 14863–14868, 1998.

Ferea, T.L., Botstein, D., Brown, P.O., and Rosenzwig, R.F., Systematic changes in gene expression patterns following adaptive evolution in yeast, *Proc. Natl. Acad. Sci. USA*, 96, 9721–9726, 1999.

Fountoulakis, M., Berndt, P., Boelsterli, U.S., Crameri, F., Winter, M., Albertini, S., and Suter, L., Two-dimensional database of mouse liver proteins: changes in hepatic protein levels following treatment with acetaminophen or its nontoxic regioisomer 3-acetamidophenol, *Electrophoresis* 21, 2148–2161, 2000.

Hernandez, M.R., Agapova, O.A., Yang, P., Salvador-Silva, M., Ricard, C.S., and Aoi, S., Differential gene expression in astrocytes from human normal and glaucomatous optic nerve head analyzed by cDNA microarray, *GLIA*, 38, 45–64, 2002.

Ivanov, I., Schaab, C., Planitzer, S., Teichmann, U., Machl, A., Therml, S., Meier-Ewert, S., Seizinger, B., and Loferer, H., DNA microarray technology and antimicrobial drug discovery, *Pharmacogenomics*, 1, 169–178, 2000.

Iyer, V.R., Eisen, M.B., Ross, D.T., Schuler, G., Moore, T., Lee, J.C.F., Trent, J.M., Straudt, L.M., Hudson, J., Jr., Boguski, M.S., Lashkari, D., Shalon, D., Botstein, D., and Brown, P.O., Transcriptional program in the response of human fibroblasts to serum, *Science*, 283, 83–87, 1999.

Jain, K.K., Applications of biochip and microarray systems in pharmacogenomics, 1, 289–307, 2000.

Katsuma, S., Nishi, K., Tanigawara, K., Ikawa, H., Shiojima, S., Takagaki, K., Kaminishi, Y., Suzuki, Y., Hirasawa, A., Ohgi, T., Yano, J., Murakami, Y., and Tsujimoto, G., Molecular monitoring of bleomycin-induced pulmonary fibrosis by cDNA microarray-based gene expression profiling, *Biochem. Biophys. Res. Commun.*, 288, 747–751, 2001.

Lane, S., Birse, C., Zhou, S., Matson, R., and Liu, H., DNA array studies demonstrate convergent regulation of virulence factors cph1, cph2 and efg1 in *Candida albicans*, *J. Biol. Chem.*, 276, 48988–48996, 2001.

Lapteva, N., Nieda, M., Ando, Y., Ide, K., Hatta-Ohashi, Y., Dymshits, G., Ishikawa, Y., Juji, T., and Tokunaga, K., Expression of renin-angiotensin system genes in immature and mature dendritic cells identified using human cDNA microarray, *Biochem. Biophys. Res. Commun.*, 285, 1059–1065, 2001.

Lashkari, D.A., DeRisi, J.L., McCusker, J.H., Namath, A.F., Gentile, C., Hwang, S.Y., Brown, P.O., and Davis RW, Yeast microarrays for genome wide parallel genetic and gene expression analysis, *Proc. Natl. Acad. Sci. USA*, 94, 13507–13062, 1997.

Lock, C., Hermans, G., Pedotti, R., Brendolan, A., Schadt, E., Garren, H., Langer-Gould, A., Strober, S., Cannella, B., Allard, J., Klonowski, P., Austin, A., Ladd, N., Kaminski, N., Galli, S.J., Oksenberg, J.R., Raine, C.S., Heller, R., and Steinman, L., Gene microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis, *Nature Med.*, 8, 500–508, 2002.

Lomax, M.I., Huang, L., Cho, Y., Gong, T.W.L., and Altschuler, R.A., Differential display and gene arrays to examine auditory plasticity, *Hearing Res.*, 147, 293–302, 2000.

Lomri, A., Lemonnier, J., Delannoy, P., and Marie, P.J., Increased expression of protein kinase-Cα, interleukin-1α, and RhoA guanosine 5´-triphosphatase in osteoblasts expressing the ser252trp fibroblast growth factor 2 apert mutation: identification by analysis of complementary DNA microarray, *J. Bone Mineral Res.*, 16, 705–712, 2001.

Maeda, S., Otsuka, M., Hirata, Y., Mitsuno, Y., Yoshida, H., Shiratori, Y., Masuho, Y., Muramatsu, M.A., Seki, N., and Omata, M., cDNA microarray analysis of *Helicobacter pylori*-mediated alteration of gene expression in gastric cancer cells, *Biochem. Biophys. Res. Commun.*, 284, 443-449, 2001.

Mahadevappa, M. and Warrington, J.A., A high-density probe array sample preparation method using 10- to 100-fold fewer cells, *Nature Biotechnol.*, 17, 1134–1136, 1999.

Marton, M.J. et al., Calcineurin signaling pathway in yeast, *Nature Med.*, 4, 1293–1301, 1998.

Matson, R.S., Oligonucleotide arrays for the detection of ras mutations, in *Nonradioactive Analysis of Bio-Molecules*, Kessler, C., Ed., Springer, Heidelberg, 2000, chap. 64.

Al Moustafa, A.E., Alaoui-Jamali, M.A., Batist, G., Hernandes-Perez, M., Serruya, C., Alpert, L., Black, M.J., Sladek, R., and Foulkes, W.D., Identification of genes associated with head and neck carcinogenesis by cDNA microarray comparison between matched primary normal epithelial and squamous carcinoma cells, *Oncogene* 21, 2634–2640, 2002.

Mullan, P.B., McWilliams, S., Quinn, J., Andrews, H., Gilmore, P., McCabe, N., McKenna, S., and Harkin, D.P., Uncovering BRCA1-regulated signaling pathways by microarray-based expression profiling, *Biochem. Soc. Trans.*, 29, 678–683, 2001.

Nishizuka, S., Winokur, S.T., Simon, M., Martin, J., Tsujimoto, H., and Stanbridge, E.J., Oligonucleotide microarray expression analysis of genes whose expression is correlated with tumorigenic and non-tumorigenic phenotype of HeLa X human fibroblast cells, *Cancer Lett.*, 165, 210-209, 2001.

Pinheiro, N.A., Caballero, O.L., Soares, F., Reis, L.F.L., and Simpson, A.J.G., Significant overexpression of oligophrenin-1 in colorectal tumors detected by cDNA microarray analysis, *Cancer Lett.*, 172, 67–73, 2001.

Reilly, T.P., Bourdi, M., Brady, J.N., Pise-Madison, C.A., Radonovich, M.F., George, J.W., and Pohl, L.R., Expression profiling of acetaminophen liver toxicity in mice using microarray technology, *Biochem. Biophys. Res. Commun.*, 282, 321–328, 2001.

Ross, D.T., Scherf, U., Eisen, M.B., Perou, C.M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S.S., Van de Rijn, M., Waltam, M., Pergamenschikov, A., Lee, J.C.F., Lashkari, D., Shalon, D., Myers, T.G., Weinstein, J.N., Botstein, D., and Brown, P.O., Systematic variation in gene expression patterns in human cancer cell lines, *Nature Genetics*, 24, 227–235, 2000.

Saiki, R.K., Walsh, P.S., Levenson, C.H., and Erlich, H.A., Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes, *Proc. Natl. Acad. Sci. USA*, 86, 6230–6234, 1989.

chapter six

Protein microarray applications

Introduction

This chapter will review some of the key applications presented by protein microarrays. The use of protein microarrays stems from works on gene expression arrays described earlier. However, unlike its predecessors whose process formats (mutation detection, polymorphism screening, gene expression analysis, etc.) are essentially based upon solid-phase hybridization of nucleic acid complementary strands, the protein array may play different roles and comprise a variety of formats.

For example, analogous to a gene array, we may print antibodies onto a solid support and capture the complementary antigen. However, we may also wish to immobilize an array of protein kinases and simultaneously measure their respective substrate levels in a cell supernatant. Thus, protein microarrays can be classified into at least two distinct categories, as Kodadek (2001) suggested: (a) protein function arrays that measure the activities of native proteins and (b) protein-detecting arrays that monitor protein levels. Several biomedical industry groups initiated work on protein-detecting arrays during the late 1980s (the microspot concept of Ekins, 1989) and 1990s (the slide-based microarray immunoassay of Silzel et al., 1998 and the high-throughput microarray-based enzyme-linked immunosorbent assay (ELISA) of Mendoza et al., 1999). However, the first publication describing the use of protein microarrays on a scale comparable to the gene expression arrays described by Schena et al. (1995) must be credited to MacBeath and Schreiber at Harvard (2000). The technology quickly became adopted for work on the proteome. The examples discussed in this chapter will highlight the applications of protein microarrays in biomedical research.

Spot theory

The analytical concepts and strategies developed by Ekins and his coworkers (1990) to improve the sensitivity of the immunoassay heralded the appearance

of the antibody array. Ekins determined that only a small amount of capture (sensory) antibody was required to measure an antigen in solution. More precisely, only a few binding sites within the antibody spot had to be occupied by antigen in order to reveal antigen concentration. Ekins' ambient analyte immunoassay relied on this concept — that the fractional occupancy (of capture antibody with antigen) reflects the true "analyte" concentration of antigen in solution. This amount of antibody does not harvest a significant amount of antigen. Thus, the conventional practice of coating a microtiter plate well with excess antibody for use in an ELISA is not necessary. Only a small spot (microspot) is required.

The requirement for only a microspot of antibody for assay allows multiple immunoassays to be run simultaneously; microspots of different analyte-specific antibodies can be utilized in the same well. An early demonstration of the microspot immunoassay for determination of tumor necrosis factor (TNF) was achieved using a Lasersharp scanning confocal microscope (BioRad, Hercules, CA). See Figure 6.1. A ratiometric approach was used in which the capture antibody was labeled with Texas Red and the biotinylated secondary antibody labeled with avidin–fluorescein isothiocyanate (FITC). The FITC:Texas Red ratio for each microspot was determined and the specific binding activity determined under ambient analyte conditions (Figure 6.2).

While it was only possible to create low density microspot arrays then, Ekins postulated that a microspot array of 50-μ diameter spots could yield over a million immunoassays in a single square centimeter. However, implementation of the antibody microarray required additional considerations such as a means to deposit the microspots onto a substrate and the availability of high sensitivity detectors. As a result, Ekins entered into a collaboration with Boehringer-Mannheim GmbH in 1991 to pursue commercialization of the Microspot as a third-generation ultrasensitive immunoassay (Ekins, 1998).

Boehringer-Mannheim constructed microarrays on small, disposable single-well polystyrene carriers using piezo-electric "ink-jet" technology to print 100 to 200 spots on a 3-mm diameter well bottom (Figure 6.3). Each spot was approximately 80 μ in diameter. A prototype fully automated system built by the company in 1994–1995 produced 10,000 single-well carrier chips (200 spot arrays/chip) per hour. Lower limits of detection were reported to be approximately 0.01 detection molecules/μm^2, corresponding to an assay sensitivity approaching 10^{-17} M. High sensitivity was achieved for specific analytes, e.g., thyroid-stimulating hormone (TSH), 0.01 μU/mL, and total IgE, 0.01 U/mL (Ekins, 2001).

An alternative approach is the "mass-sensing" multianalyte microarray immunoassay first described by researchers at Beckman Coulter, Fullerton, CA (Silzel et al., 1998). As early as 1991, other groups at Beckman Coulter had adapted commercially available ink-jet printers (e.g., the Hewlett Packard Deskjet) for depositing oligonucleotides or proteins such as streptavidin onto substrates to create arrays (Matson, unpublished data). Piezoelectric

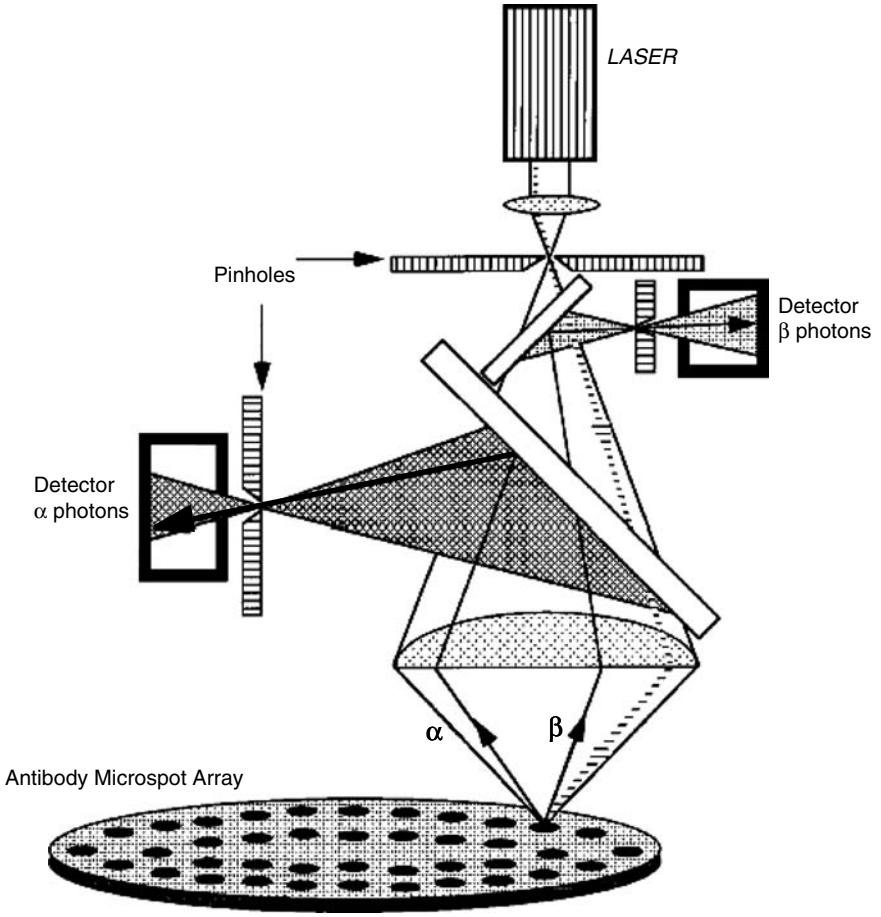


Figure 6.1 Microspot array detection. (From Ekins, R. et al., *Clin. Chim. Acta*, 194, 91–114, 1990. With permission.)

jet printing using a prototype drop-on-demand system (MicroFab Technologies, Plano, TX) was employed subsequently.

Silzel and coworkers utilized these print technologies to deposit avidin for immobilization of biotinylated antibodies and for the direct printing of monoclonal antibodies. For example, an array of four human IgG subclasses (IgG1, IgG2, IgG3, IgG4) was immobilized onto a polystyrene slide as a group of 200-μ diameter spots (Figure 6.4). In contrast to the ambient analyte immunoassay, these experiments demonstrated that capture antibody substantially depleted the sample of antigen analyte within a few hours. Thus, the antibody spot was thought to "harvest" the total analyte mass from solution — a process Silzel et al. termed "mass-sensing."

How different are the microspot ambient analyte and mass-sensing microarray immunoassays? Figure 6.5 compares these two formats. To be useful as clinical assays, Ekins insists that microspot arrays obey "ambient

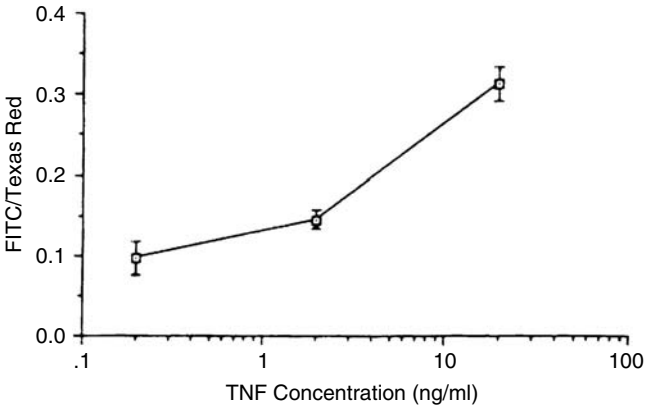


Figure 6.2 TNF microspot immunoassay based upon ratiometric signal detection. (From Ekins, R. et al., *Clin. Chim. Acta*, 194, 91–114, 1990. With permission.)

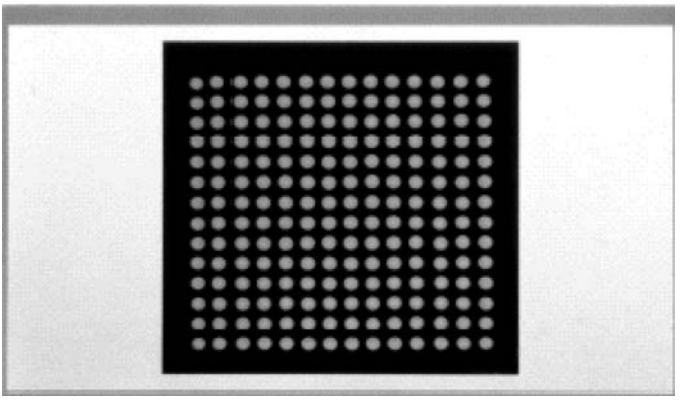


Figure 6.3 Boehringer-Mannheim's microspot array. (From Ekins, R.P., *Clin. Chem.*, 44, 2015–2030, 1998. With permission.)

analyte" conditions. That is, the solid-phase tethered capture (sensor) antibody must remove from the sample solution ≤1% of the analyte present in the sample for the assay to be valid. Provided the capture antibody equilibrium association constant, unit = Liter mole^{-1} (K_a) is not altered during attachment, the ambient analyte condition would be satisfied in most cases by a surface antibody concentration $<0.01/K_a$ M. If $K_a \sim 10^{11}$, then the antibody concentration required would be $\sim 0.01/10^{11}$ M or 0.1 pM. That corresponds to roughly 6×10^7 antibody molecules or binding sites from a 1-mL sample volume.

Adsorbed monolayers of antibodies (IgG) can be achieved in the range of 130 to 650 ng/cm^2 or from 10^9 to 10^{10} molecules/mm^2, depending upon the solid phase. Thus, antibody microspots having diameters of 50 to 120 μ would be sufficient. For a weaker binding antibody, $K_a \sim 10^{10}$, the ambient

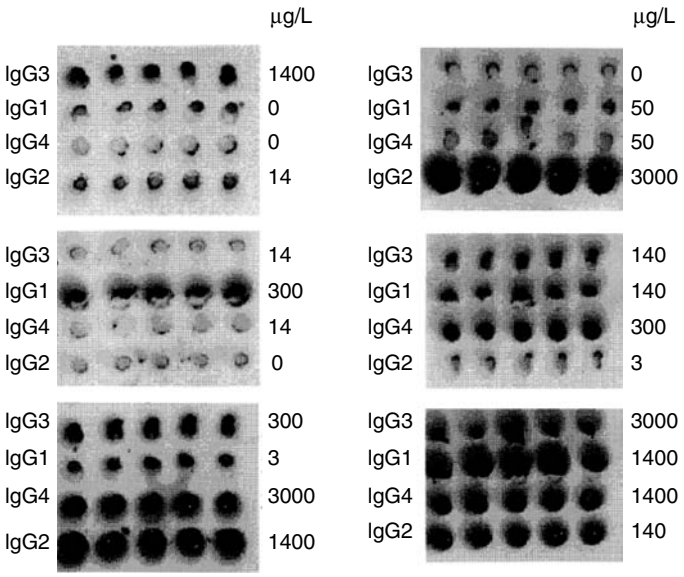


Figure 6.4 Raw image data from four-analyte human IgG subclass assay. Numbers beside microarray rows indicate expected concentration of each subclass in the four-component myeloma mixture assayed in a given image. (From Silzel, J.W. et al., *Clin. Chem.*, 44, 2036–2043, 1998. With permission.)

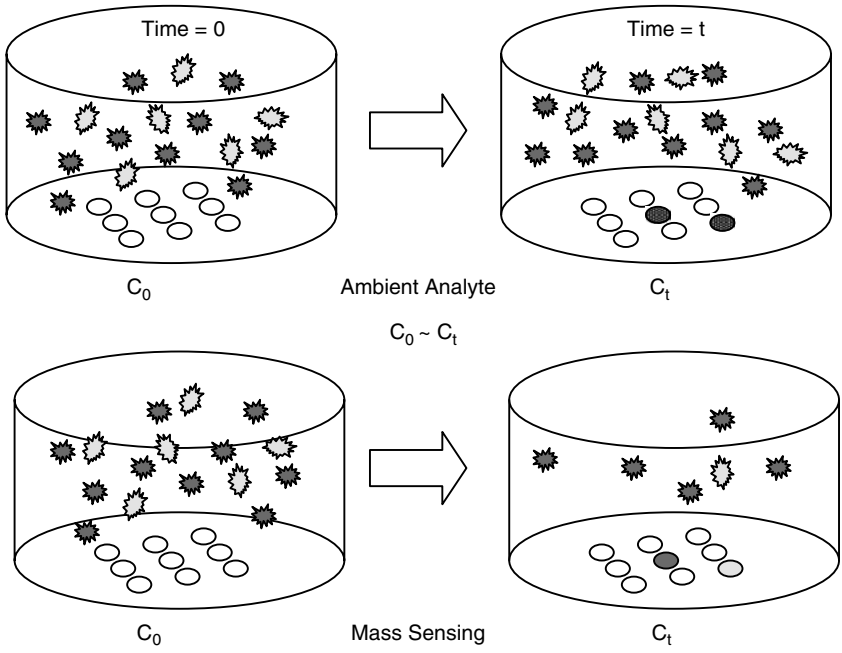


Figure 6.5 Ambient analyte vs. mass sensing immunoassay conditions.

analyte condition would require larger microspot diameters (~170 to 380 μ). However, it is more likely that the sampling volume for an immunoassay would be 100 μL or less, so that microspots on the order of 5- to 40-μ diameter would necessarily be utilized, depending upon antibody density or binding affinity. The smaller the diameter spot of antibody monolayer, the more likely the ambient analyte condition will be maintained. From a practical view, the spot diameter cannot be vanishingly smaller since that would mean a lower limit in detectability and acceptable capture rate.

In the work described by Silzel et al. (1998), a typical binding capacity for a 200-μ spot was reported on the order of 10^{10} analyte molecules based upon biotin DBCy5 dye binding to avidin spots. Based on a 100-μL sampling volume, this would correspond to an initial concentration of about 1.7×10^{-10} M or 170 pM, assuming full depletion of the sample. This suggests a binding condition much higher than that of the ambient analyte condition of $0.01/K_a$, where $K_a \sim 10^{10}$.

Silzel et al. then compared the mass assay relative to the ambient analyte assay based upon a hypothetical assay for TSH (Figure 6.6) and found a 60-fold improvement in absolute fluorescent signal when using the mass-sensing approach. For example, with TSH at a concentration of 10^{-15} M (60,000 TSH molecules) and an anti-TSH capture antibody with a $K_a \sim 10^{10}$, the ambient analyte condition would be satisfied with a total of 600 molecules (1%) bound to the microspot, while the mass-sensing assay would sequester 38,000 molecules (63%) for improved detection. However, under experimental conditions employing an antibody array, efforts to achieve such levels of sensitivity were compromised by nonspecific binding. The nonspecific signal was found to be associated with the capture antibody spot

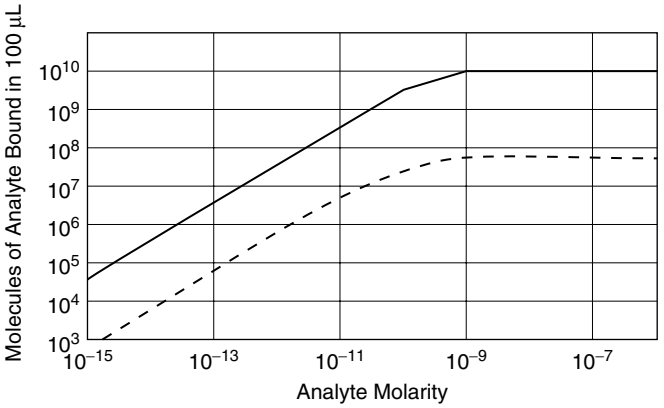


Figure 6.6 Hypothetical comparison of ambient analyte and mass-sensing analysis of thyroid-stimulating hormone. Solid line indicates mass assay; dashed line indicates ambient analyte assay. Antibody affinity of 10^{-10} liter per mole and volume of 100 μL are assumed. Mass assay assumes 10^{10} binding sites per 100 μL. (From Silzel, J.W. et al., *Clin. Chem.*, 44, 2036–2043, 1998. With permission.)

and estimated to be approximately 100-fold higher than the instrumental detection limit. IgG3 could be detected at approximately 15 µg/L (100 pM) providing an estimated capture of 3×10^8 IgG3 molecules/200-µ spot.

Sapsford et al. (2001) examined microarray-based antibody–antigen binding kinetics in real time to determine the effect of spot size. Capture antibodies were immobilized in an array pattern onto silver-clad microscope slides. Antimouse IgG was directly attached to the surface or attached via neutravidin capture of the biotinylated antibody. Cy5-labeled mouse IgG capture was monitored based upon the signal generated from the excitation of an evanescent wave guide (slide) with a 635-nm laser source; detection was achieved by a charge-coupled device (CCD) camera system. Both static and flow-through conditions were employed.

The binding kinetics were characterized in terms of the apparent time constant ($K_{app} = k_f C + k_r$) where C = analyte concentration; k_f = association rate constant; and k_r = dissociation rate constant. In closed loop experiments, a plateau value for K_{app} of 0.0024/s was reached at a linear flow rate of 2.67 mL/min. K_{app} was found to decrease with decreasing antigen concentration (C), with equilibrium achieved only at the highest level (1 µg/mL). The association rate constant K_f was calculated at 3.6×10^5 M/s for IgG binding.

To determine the effect of spot size, various photolithographic masks were used to create arrays of square patterns at spot widths from 80 to 1145 µ. The K_{app} (± 1 standard deviation) did not vary over the spot size range of 80 to 1145 µ under constant flow conditions at fixed levels of antigen concentration, and little effect of variation in spot size was noted on mean binding.

This study does not support Ekins' ambient analyte model in that under flow conditions the rate of analyte capture is not increased as the spot size is decreased. However, as the authors point out, this study is not necessarily a contradiction because Ekins' assumptions are based upon achieving equilibrium under static conditions in which diffusion is more likely an important limiting factor in influencing the binding rate. On a more operational note, Sapsford et al. also observed the problem in collecting sufficient signal from a very small spot relative to background noise.

The other interesting observation of this study was that K_f for directly immobilized antibody was only 50% of the avidin–biotin bridged antibody. This suggests that random coupling reduces the number of effective binding sites relative to the oriented coupling using avidin–biotin.

In practical terms, microarrays are operational by at least two means. One can (1) create very small spots of capture ligands that bind only appreciable amounts of analyte and measure binding with ultrasensitive signaling reagents and detectors or (2) create arrays of much larger spots of capture ligands that significantly deplete the sample of analyte and permit the use of less sensitive (and presumably less expensive) approaches to detection.

Applications demonstrating protein microarray utility

Researchers at Genometrix (The Woodlands, TX) provided one of the first examples of a high-throughput microarray-based ELISA (Mendoza et al., 1999). They created arrays of antigens within wells of an optically flat glass substrate. The 96 wells were formed using a hydrophobic Teflon mask. The wells were then chemically treated with N-hydroxysuccinimide (NHS)-active esters for covalent immobilization of proteins. Each well contained four replicates of a 6 × 6 element array (Figure 6.7). A uniquely designed capillary printer allowed for the simultaneous arraying of the 36 elements comprising various IgG antigens (Figure 6.8). Approximately 200 pL from each capillary was delivered to the glass surface, resulting in spot diameters of 275 µ at center-to-center spacings of 300 µ. The protein droplets were allowed to dry on the surfaces of the glass substrates prior to assay. While this allowed for visualization of the print run for quality control purposes, this practice may not be advisable for other proteins. In this case, the immobilized IgG antigens most likely provided recognizable epitopes in both the native and denatured states.

Assays utilized conventional ELISA processing except that the reagent volumes were greatly reduced, ranging from 25- to 50-µL well additions. Following a 1-hr block in casein, the monoclonal detection antibody was incubated an additional hour at room temperature. The assay sensitivity from the micro-ELISA was approximately 13.4 ng/mL for rabbit IgG — a result similar to that reported by Silzel et al. (1998). The benefit of the "array of arrays" approach is that 96 samples could be processed for multiple (36 to 144) analytes within the same time interval as a standard single-analyte ELISA.

Microtiter-based antibody arrays

Conventional plastic microtiter plates have also been adopted for use with protein microarrays in the "array of arrays" format. Matson et al. (Oak Ridge

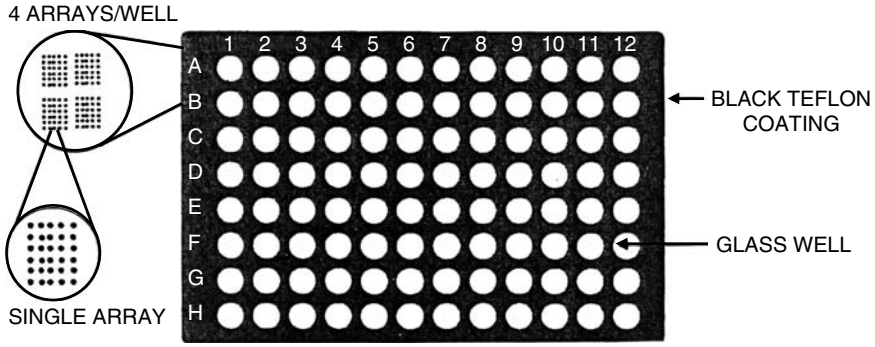


Figure 6.7 A 96-well microarray plate format. (From Mendoza, L.G. et al., *Biotechiques*, 24, 778–788, 1999. With permission.)

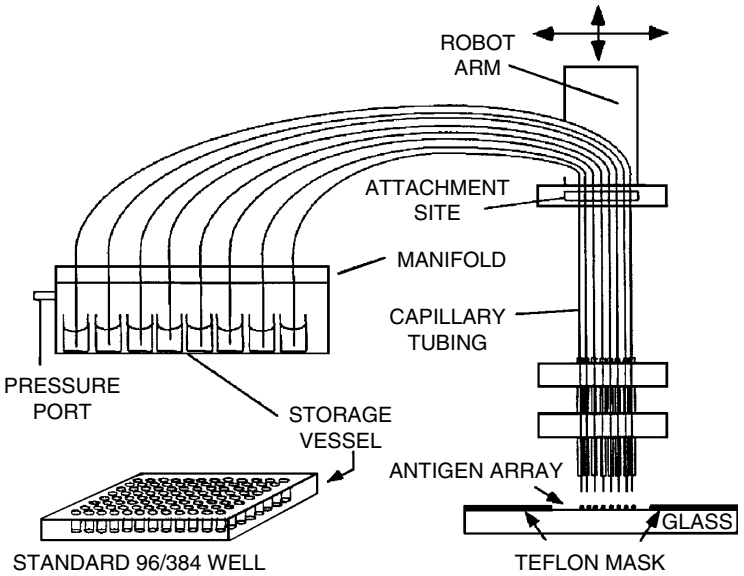


Figure 6.8 Capillary printer. (From Mendoza, L.G. et al., *Biotechiques*, 24, 778–788, 1999. With permission.)

Conference, 2001) printed antiinterleukin monoclonal antibodies in the bottom of a prototype shallow, vacuum-formed, 96-well polypropylene plate (Figure 6.9). The plastic was surface modified with acyl fluoride groups for rapid and efficient covalent attachment of the antibodies.

Similar results in terms of sensitivity and dynamic range were obtained with printing of macroarrays (~500-μ diameter spots) using a Beckman Coulter Biomek high-density replicating tool (HDRT) gridding tool or microarrays (200-μ features) prepared using an arrayer and quill pins. The micro-ELISA was performed with biotinylated secondary antibodies with streptavidin–alkaline phosphatase and (enzyme labeled fluorescence) ELF® (Molecular Probes, Eugene, OR) signal amplification. In the case of IL-8, a dynamic range from 16 pg/mL [7.2% coefficient of variation (CV)] to 1000 pg/mL (28.5% CV) was obtained with a minimal detectable dose estimated to be ≤1 pg/mL.

Moody et al. (2001) created a 3 × 3 "mini-array" of anticytokine monoclonal antibodies in wells of black Maxisorp 96-well plates (Nalge Nunc, Rochester, NY). The spots were approximately 0.4 mm in diameter and contained about 1 ng (7 fmol) of antibody. Assays were performed using biotinylated secondary antibodies with signal development using streptavidin–horseradish peroxidase (HRP) and SuperSignal™ (Pierce Endogen, Rockford, IL) chemiluminescent reagent.

Dynamic ranges were reported as 0.8 to 200 pg/mL [IL-1α, IL-6, IL-10, interferon (INF)-α, and INF-γ]; 0.4 to 100 pg/mL (IL-1β); and 1.6 to 400 pg/mL (TNF-α). The average signal well-to-well intraplate CVs ranged

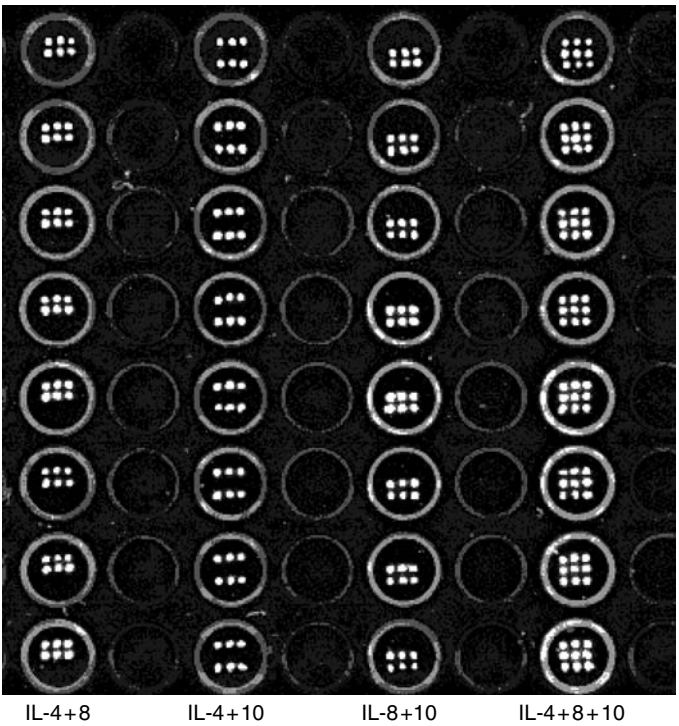


Figure 6.9 Monoclonal (antiinterleukin) antibody array in vacuum-formed 96-well microplate. (From Matson, R.S. et al., Poster 20, Oak Ridge Conference, 2001. With permission.)

from 8 to 13.1% while plate-to-plate CVs varied from 2.2 to 11.3%. The authors demonstrated the ability to measure cytokine levels in lectin-stimulated human peripheral blood mononuclear cells and in lipopolysaccharide-stimulated (human acute monocytic leukemia) THP-1 cells for antiinflammatory drug screening with dexamethasone. One drawback to the use of a chemiluminescent reagent was the blooming of light into adjacent spots. In order to expand dynamic range, it was necessary to increase the spot center-to-center distances to at least 1.25 mm and limit the number of spots in the array to nine spots/well.

Membranes

A somewhat less sophisticated approach is offset gridding onto membranes using robotic arms equipped with pin transfer tools (Figure 6.10). Lehrach's group at the Max Planck Institute in Berlin (Lueking et al., 1999) created grids of expressed histidine (His)-tagged protein from lysates (or purified by Ni–nitrilotriacetate metal affinity chromatography) on polyvinylidene fluoride (PVDF) membrane. [*Note:* His tagging of proteins not only allows for purification from the lysates, but also provides a useful quality control

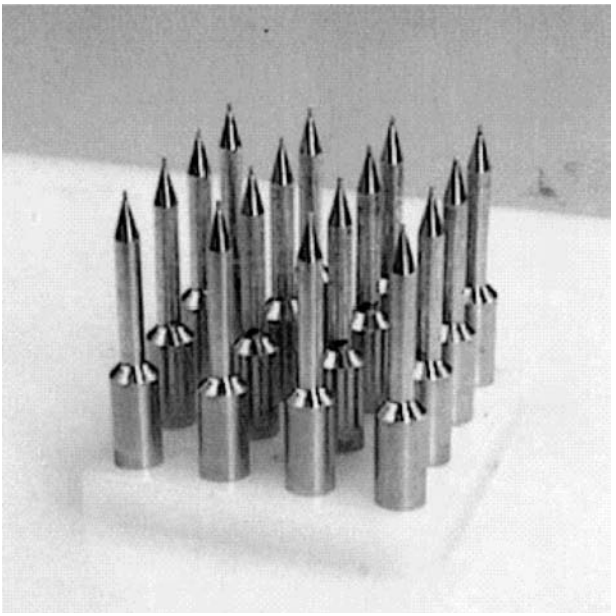


Figure 6.10 Early example of robotic arm pin transfer tool for printing proteins. (From Lueking, A. et al., *Anal. Biochem.*, 270, 103–111, 1999. With permission.)

method for determining printing efficiency, e.g., anti-RGS (regulators of gene protein signaling) His antibody.]

At a pin-to-pin center distance of 4.5 mm and with 250–μ tip pins, it was possible to array 4800 samples onto 25- × 75-mm filter strips. Developed spots appeared sharp and uniform (Figure 6.11). The reported threshold sensitivity was calculated as approximately 10 nM for purified G3PDH protein (10 pg/25nL) spotted at several dilutions. The drawback to this approach is the extensive washing required to remove nonspecific proteins and reagents from the membrane.

Joos and colleagues (2000) printed antigens onto nylon filters and glass slides and compared the titers for autoantibodies in a micro-ELISA format. Antigens were delivered to the various substrates stabilized in 10% glycerol, 0.1% sodium dodecyl sulfate (SDS), and 5 μg/mL bovine serum albumin (BSA). Protein arrays prepared in this manner were reported to be functional for up to a month when stored a room temperature in the dark (see exception below). According to the investigator's calculations, these arrays most likely operated in the mass-sensing mode rather than by the ambient analyte immunoassay constraints.

The influences of spot size and antigen density on sensitivity were examined. As predicted from Ekins' (1990) microspot model (but under presumed mass-sensing conditions), the signal intensity decreased upon dilution of the antigen independent of spot diameter (area) or antigen density (Figure 6.12). The comparison of sensitivities (lowest detectable

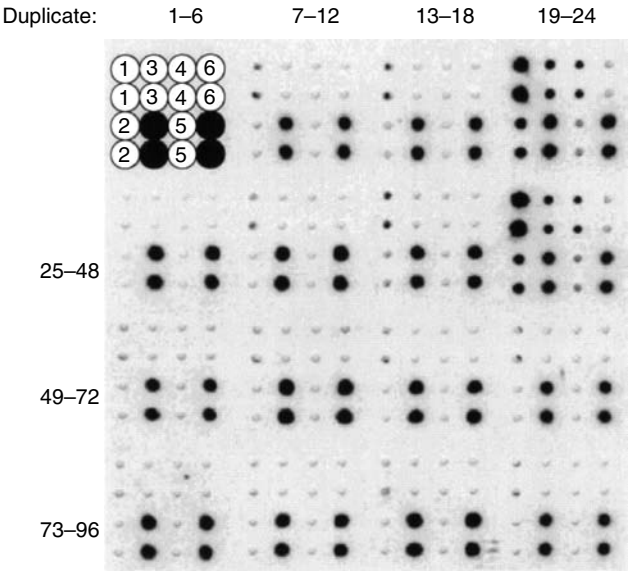


Figure 6.11 Membrane filter protein array. (From Lueking, A. et al., *Anal. Biochem.*, 270, 103–111, 1999. With permission.)

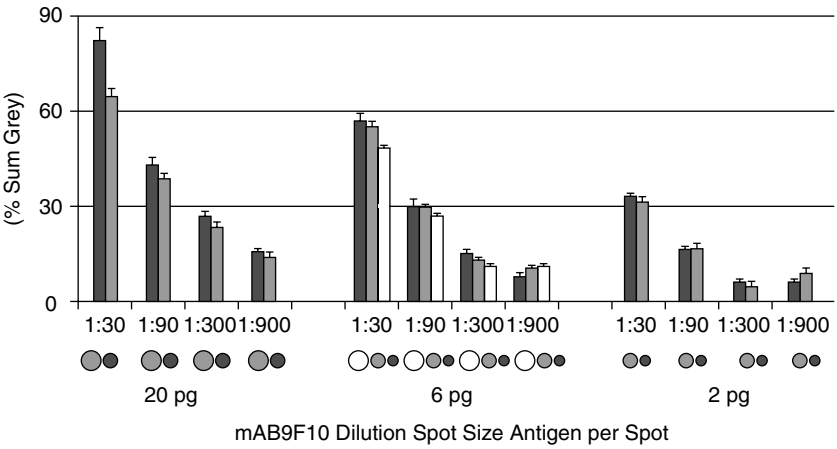


Figure 6.12 Antigen array: signal intensity vs. spot size and loading density. (From Joos, T.O. et al., *Electrophoresis*, 21, 2641–2650, 2000. With permission.)

signals above background) revealed that the nitrocellulose membrane provided a fivefold lower detection limit (8 fg huIgG/0.25nL spot volume ~0.2 nM) than an aldehyde-activated glass slide (40 fg huIgG). Understandably, most of the reported work was done using a nitrocellulose membrane with chemiluminescent signal detection. The membrane micro-ELISA gave comparable results to conventional ELISA with some notable exceptions in which

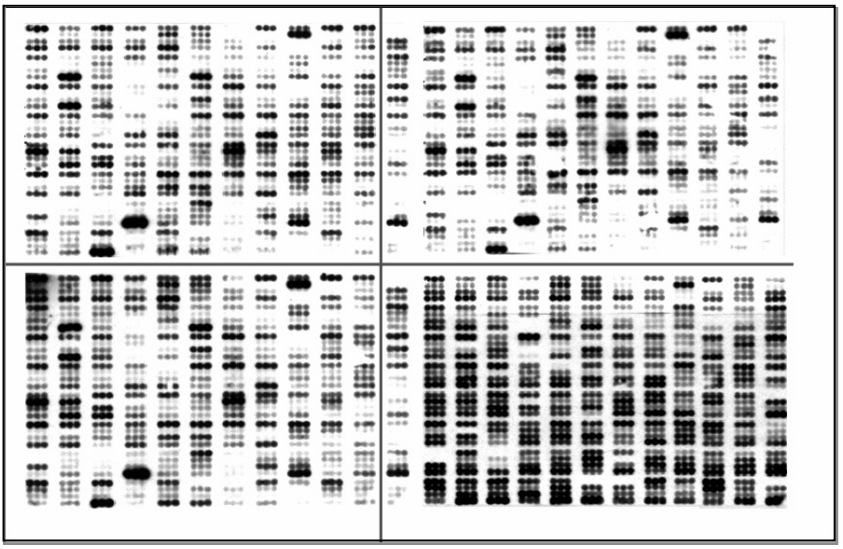


Figure 6.13 Milagen's antibody array; 768 antibodies are arrayed in duplicate on a solid support and visualized through enzyme-linked secondary antibodies and chemiluminescence. (Courtesy of Milagen, Inc., Emeryville, CA.)

the microarray showed lower titer due to denaturation of the antigens on the microarrays during storage. Freshly prepared microarrays were found to perform equivalent to the ELISA.

Milagen (Emeryville, CA) created antibody arrays for purposes of target discovery by a process designed ANTIBIOMIX™ (antibody against biological mixture) in which polyclonal antibodies are used as tools to screen for both known and unknown gene products from clinical samples (Valle and Jendoubi, 2003). Milagen reported having 61,000 polyclonal antibodies and expected to have 80,000 to 100,000 available by the end of 2002. The advantage of polyclonal over monoclonal antibodies in this application is that polyclonals have greater coverage for detection purposes, i.e., broader specificities and affinities provide a greater number of hits. This is particularly important in examining antigens with posttranslational modifications. The antibody array panels generated by the ANTIBIOMIX process can be used for differential display, for example, of diseased and normal states, and for monitoring the progression of a disease and its outcome (Figure 6.13).

Glass slides

Perhaps the first published demonstration of high density applications for protein microarrays came from the work of MacBeath and Schreiber at Harvard (2000). Proteins were arrayed onto aldehyde-activated glass slides and analyzed in much the same manner described for the creation of cDNA microarrays (Schena et al., 1995), including the use of dual-color label

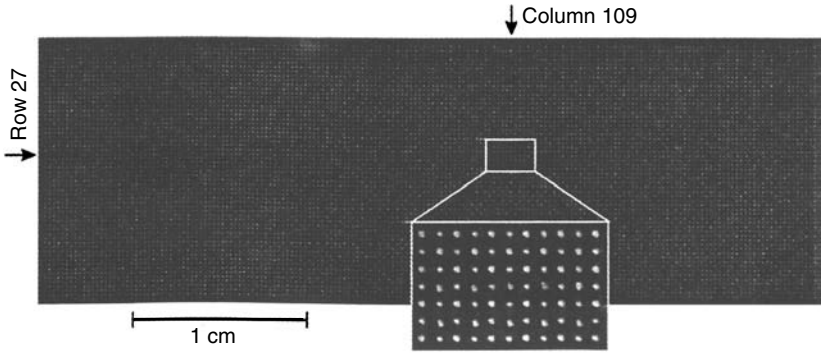


Figure 6.14 Protein microarray specificity. (From MacBeath, G. and Schreiber, S.L., *Science*, 289, 1760–1763, 2000. With permission.)

detection of specific proteins. Special care was taken to maintain proteins in their hydrated states to minimize surface denaturation. Proteins were printed in phosphate-buffered saline (PBS) containing 40% glycerol to prevent evaporation. Others including Delehanty and Ligler (2002) used a low salt buffer containing sucrose and BSA for this purpose.

The immobilization strategies are of particular interest. The authors reasoned that the use of aldehydes to tether proteins to the solid phase could be ideal for certain protein–protein interaction studies. Since many protein–lysine residues are available for coupling to aldehydes via Schiff's base, a number of spatial orientations are possible. Such random oriented attachments would permit exposure of various surfaces of a protein to the solution, and new protein–protein interactions would be potentially possible.

Another useful strategy is scaffolding. For example, immunoassays employ BSA both as a blocking agent to reduce nonspecific adsorption of other proteins and also as a scaffold. Essentially, BSA is first attached to the solid support and then further derivatized for the coupling of additional capture ligands. MacBeath and Schreiber first formed a monolayer of BSA and then printed proteins on top of the monolayer. In this manner, small proteins were expressed on the surface and not buried by the BSA.

The specificity that can be achievable with the protein microarray was demonstrated by a number of powerful examples. First, a single spot of FRB protein (that contains the binding domain for FRAP [rapamycin-associated protein]) printed down on the array was successfully detected among 10,799 spots of protein G (Figure 6.14). Second, enzyme–substrate reactions were possible using immobilized protein substrates of various kinases. The bound substrates were phosphorylated only by incubation of the slide with the specific kinase. Third, the microarray was used to screen for targets by immobilized protein–small molecule interaction studies. The immobilized protein receptors for steroid, biotin, and ketoamide ester were able to recognize their cognate partners. Finally, the ability to measure small molecule-induced

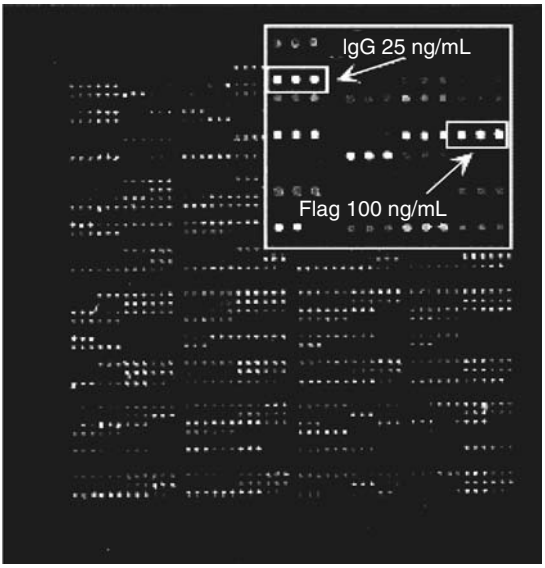


Figure 6.15 Applying dual-color ratiometric gene expression labeling approach to protein expression analysis. (From Haab, B.B. et al., *Genome Biol.*, 2, 0004.1–0004.13, 2001. With permission.)

conformational binding was shown by association of immunophilin with FRB only in the presence of the small rapamycin molecule.

Brown's group at Stanford, which introduced the slide microarray (Schena et al., 1995), turned to the large-scale immobilization of antibodies and antigens to study protein abundance (Haab et al., 2001). They examined the performances of 115 antibody–antigen pairs by printing arrays of antibodies and antigens onto polylysine-coated glass microscope slides. Six to twelve replicates of each protein were placed on the slides and comparative fluorescence labeling was used to measure performance.

Cy3 (green fluor)-labeled reference proteins were prepared at constant concentrations for all 115 proteins. The reference set was then mixed with the sample set labeled with Cy5 (red fluor) at specific abundance levels from 1 ng/mL to 1 μg/mL. The red-to-green ratio served as the calibrator for variation in binding between antibody and antigen on an array. The visual effect observed was variation in color. Red = higher concentration of protein in sample mix; yellow = equivalent concentration of sample and reference mixture; green = higher concentration in reference set, as shown in Figure 6.15. More quantitative information was obtained in plotting the $\log_{10}$ (red to green) vs. concentration to determine titer (Figure 6.16).

In most cases, the sample titer curves followed the predicted titer standard curve based upon known concentration ratios. Observed variations in linearity were traced back to inconsistencies in labeling rather than cross-reactivity or sample preparation issues such as pipetting or mixing errors. Overall, antigen arrays performed better than antibody arrays. While the exact reasons for the

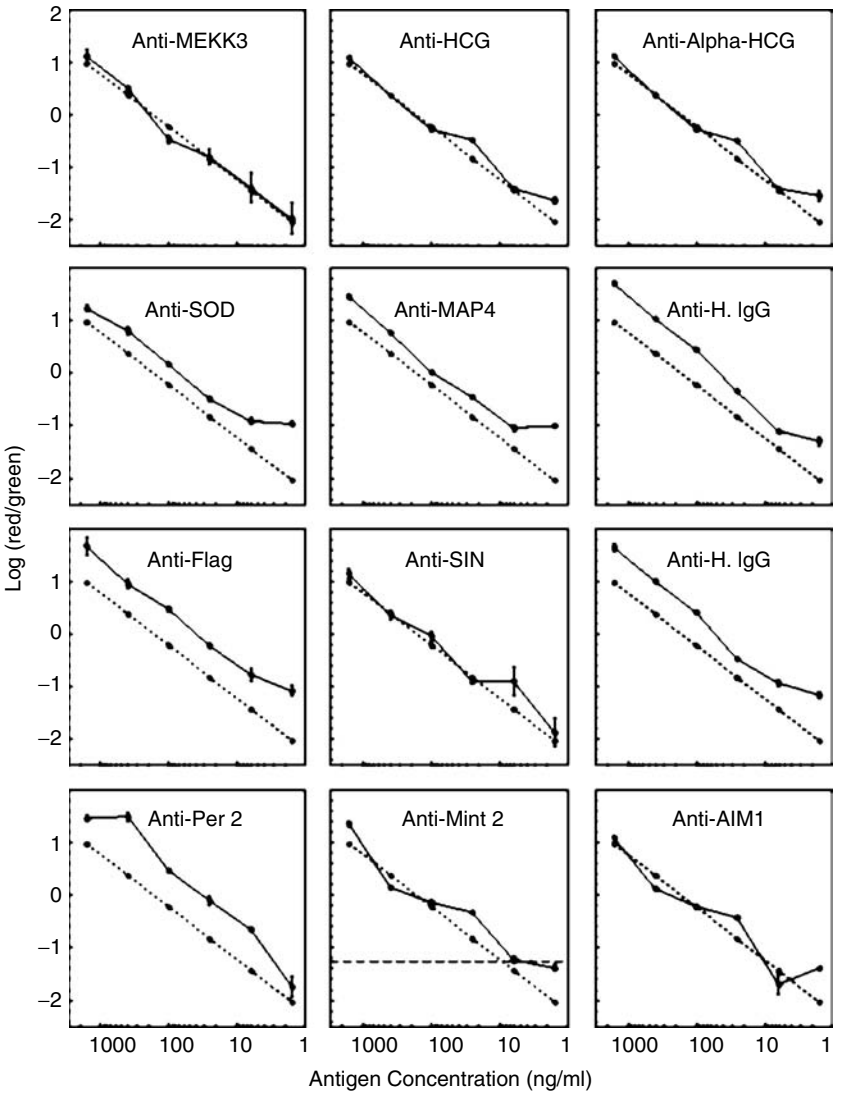


Figure 6.16 Microarray antibody–titer curves. (From Haab, B.B. et al., *Genome Biol.*, 2, 0004.1–0004.13, 2001. With permission.)

discrepancy between the two formats are not known, the discrepancy most likely relates to protein stability and inconsistencies in labeling.

The authors suggest that while antibodies are structurally similar and relatively stable proteins that can be easily labeled at lysine residues in the F_c region, antigens are much more variable both in structure and relative stability. An alternative explanation is that antibodies likely recognize surface-denatured antigens even better than native antigens. Thus, performance of antigen arrays may be more related to an increase in the number

of epitopes exposed upon surface denaturation that are recognized by the antibody.

In terms of sensitivity, the detection limits on the microarray slide using the two-color system approached 0.1 ng/mL for antigen arrays and 1 ng/mL for antibody arrays. Both were able to measure specific proteins in mixtures at 1 part per million in total protein (partial concentration).

Studies by Huang at Emory School of Medicine (2001a) underscore an important issue regarding future work with protein microarrays, that is, the selection of substrate. Membranes were selected based upon background and signal sensitivity for chemiluminescent detection (Table 6.1). For antigen arrays comprised of various IgG species, the MagnaGraph membrane (Micron Separations, Westboro, MA) provided excellent adsorption, negligible background, and moderate sensitivity. However, in a sandwich assay for cytokines, this membrane could not be used because of its very high background. Instead, Hybond ECL (Amersham, Uppsala, Sweden) was selected for cytokine assays based upon negligible background and good sensitivity.

In an effort to analyze cytokine expression at physiological levels directly from tissue culture or patient sera, Huang and coworkers modified their approach. To assay for specific cytokines in conditioned media at high sensitivity, membranes were precoated with the capture anti-cytokine antibody. Samples of conditioned media were printed and the cytokine detected by ECL using a biotin-conjugated antibody recognizing a different epitope. For MCP-1, the detection limit was 4 pg/mL, representing a 100-fold improvement in sensitivity over the standard ELISA using these "conditioned medium" arrays (Huang et al., 2001b).

Wiese et al. (2001) at Genometix designed an antibody microarray based upon immobilization of monoclonal capture antibodies to activated, silanized glass plates (Mendoza et al., 1999). They compared the performance of this micro-ELISA to that of a conventional ELISA in the detection of prostate-specific antigen (PSA), PSA-ACT (α_1-antichymotrypsin bound to PSA), and IL-6, all of which are indicators of prostate cancer. A good correlation ($r^2 = 0.88$) was obtained for 14 human serum PSA concentrations between the two assay formats (Figure 6.17). Woodbury et al. (2002) also correlated micro-ELISA with standard ELISA and observed a similar correlation ($r^2 = 0.90$) during an analysis for hepatocyte growth factor (HGF).

Certain design features of the Wiese study (2001) are noteworthy. First, capture antibodies were printed as a series of serial dilutions rather than at one concentration. This provided much needed flexibility in the dynamic working range for analyte analysis and reduced the likelihood of having to repeat the assay due to signal plateauing with out-of-range samples. Both high and low concentration ranges for all analytes could be addressed in the same well, and the researchers had no need to adjust sample volumes to measure antigens at equivalent levels. Another important feature of the microarray design is that since all analytes are represented within each well, the calibration curves for each analyte can be run simultaneously. Finally, sufficient redundancy is built into the well array to permit statistical

Table 6.1 Properties of Commercial Membranes Useful as Protein Microarray Supports

			Detection			
			IGGs		Cytokines	
Membrane	Manufacturer	Adsorption	Background	Sensitivity	Background	Sensitivity
Biotrans	ICN	Excellent	++	+++	+++++	?
Zeta-probe	Bio-Rad	Good	++	+++		
Colony/Plaque	NEN	Very good	++	+++		
Hybond N^+	Amersham	Very good	++	+++	+	+++
Magnacharge	MSI	Poor	+	+++	+++++	+
Magnagraph	MSI	Excellent	–	++	+++++	+
Hybond ECL	Amersham	Excellent	++	+++	–	+++

Source: From Huang, R.P. et al., *J. Immunol. Methods*, 255, 1–13, 2001a. With Permission.

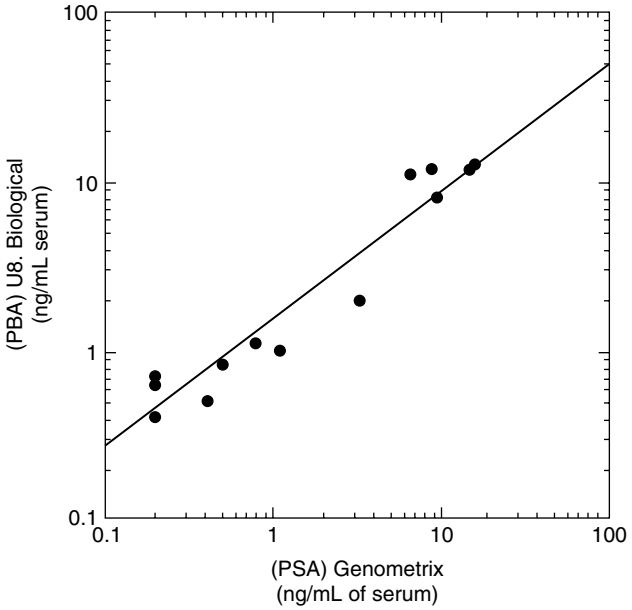


Figure 6.17 Concordance of standard ELISA and microarray ELISA formats in PSA determinations. (From Wiese, R. et al., *Clin. Chem.*, 47, 1451–1457, 2001. With permission.)

treatment of the data and minimize reliance or weighting due to single data points and outliers.

Seong (2002) investigated the performances of various commercial slides for protein microrarrays based upon different methods of protein immobilization. Two buffer systems (PBS at pH 7.4 and a carbonate buffer at pH 9.6) were studied for immobilization of IgG antigen. All slide chemistries performed well in terms of binding capacity. However, each slide chemistry displayed a slightly different loading isotherm profile. Epoxy-activated slides out-performed silylated (aldehyde) surfaces at the highest loading concentrations of IgG. The carbonate buffer marginally improved IgG surface loading over slides on which the protein was printed in PBS as the medium. This was essentially the case for either silylated (aldehyde) or silanated (amine) slides. One explanation for the enhanced loading in carbonate buffer may be related to the ionization state of the protein. At pH above the isoelectric point, the protein carries a slight negative charge, favoring stronger interaction with the positive amine surface.

Angenendt and coworkers (2002) at the Max Planck Institute undertook similar studies comparing the binding characteristics of five different antibodies arrayed onto eleven different substrates. Performance was measured in terms of sensitivity and intra- and intermicroarray slide variation. Using a QArray System (Genetix, New Milton, U.K.), antibodies were spotted down on various substrates from 25 to 40,000 amoles in eight replicates. The

antibodies included antihuman serum albumin (monoclonal and polyclonal), antifibrinogen (monoclonal and polyclonal), and antitubulin-α (monoclonal). Antigens were labeled with Cy5 or Cy3 for detection purposes, and microarrays were scanned with a laser confocal scanner (Model 428, Affymetrix). Substrates included plastic, glass, and hydrogel coatings.

The hydrogel coatings achieved the lowest limits of detection (LLD): ~1300 to 1600 amoles/spot, but exhibited significant assay variation (22% intra-slide to 37% inter-slide CV). LLD levels of surface-modified polystyrene slides (Maxisorb black, Nunc) equaled 1500 amole/spot at 15 to 32% CV, while reflective (mirror-like) slides coated with 3-aminopropyltriethoxysilane (Amersham) showed the lowest variation with CVs at 11 to 14%.

Polyacrylamide-coated slides showed wide variations in signals and lowered sensitivities. Preincubation in PBS buffer prior to spotting improved performance both in terms of LLD (~1875 amole/spot) and variance (14 to 15% CV). The use of immobilized ampholytes to alter the matrix $pK_a = -\log(K_a)$. Where K_a refers to the equilibrium acid ionization constant (i.e., charge) did not prove beneficial. All antibody arrays could be stored for 8 wk without loss in performance. Interestingly, all showed improved binding (maximal signal) after 2-wk storage.

Two important observations were noted. First, no single microarray substrate satisfied all performance criteria for all antibodies. Second, in other work, certain antibodies worked well in conventional ELISA but not in microarray format. Thus, results from ELISA do not necessarily qualify an antibody for use on microarrays. It is best to validate each antibody for use on the intended substrate in a microarray-based assay.

Avseenko et al. (2001) immobilized antigens onto aluminum-coated Mylar films by electrospray (ES) deposition. Various surface modifications of the metallized films were studied to determine their abilities to enhance sensitivity. The plastic surfaces were first cleaned by plasma discharge treatment, followed by coating with proteins (BSA and casein) or polymers such as poly (methyl methacrylate) or oxidized dextran, or they were exposed to dichlorodimethyl silane to create hydrophobic surfaces. Protein antigen was prepared in 10-fold excess sucrose and sprayed onto the surfaces to form arrays with spot diameters between 7 and 15 μm containing 1 to 4 pg protein.

ES deposition is performed using proteins prepared as a dry powder. Sucrose protects the proteins during drying and deposition. Once deposited, sucrose adsorbs moisture from the air, creating droplets that resolubilize the protein for attachment. Such microarrays can then be stored dry at –20°C for up to 8 mo without loss, as measured by ELISA. Of the various surface modifications studied, the greatest enhancement in sensitivity was achieved using dextran-coated Mylar. Dextran oxidation results in the formation of aldehyde groups that form a Schiff's base with primary amino groups on proteins. Oxidized dextran coating resulted in sensitivity to ~10 ng/mL, while reduction of the Schiff's base to form a stable covalent bond increased sensitivity to 1 ng/mL (~6.7 p*M* IgG). Contributing to this enhancement in

sensitivity was also the low background achieved using dextran rather than protein blockers.

Mezzasoma et al. (2002) created an array of microbial antigens to determine the levels of antibodies in human sera directed against *Toxoplasma gondii*, rubella virus, cytomegalovirus (CMV), and herpes simplex virus (HSV) ToRCH antigens. Microarray performance was measured against standard ELISA. Internal standards comprised of IgG and IgM were printed on each slide in duplicate. The linear range was 2 to 50 pg for IgG and 0.4 to 8 pg for IgM. The detection limit (rabbit myosin negative control, mean intensity +2 standard deviations) was interpolated at 0.5 pg or an LLD corresponding to approximately 0.04 pM. Signal was generated using Molecular Probes' Alexa 546-labeled antihuman IgG monoclonal and Alexa 594-labeled goat antihuman IgM μ chain. Detection was performed using a laser confocal scanner system and microarray performance metrics were studied.

Interslide variance was 1.7% to 8.6% CV; intraslide variance ranged from 2.6 to 15% CV. Batch-to-batch variability was between 5.2% and 18% CV for all serum IgG-reactive antigens. The antigen microarray was compared against a commercial ELISA (Radim, Pomezia, Italy) using well characterized human sera panels (54 to 56 samples) containing different levels of antigen-specific IgG. Good agreement (80 to 90% concordance) between the microarray and the ELISA determinations was obtained except for rubella virus. The microarray assay identified ~18% positive, while the ELISA scored 87.5% as positive for the virus. The discrepancy appeared to favor the results of the microarray assay when samples were evaluated by an independent method.

Measuring microarray performance

We described a number of early uses for protein microarrays. Others have suggested enhanced sensitivity when using vanishingly smaller spots to capture antigens (Ekins et al., 1990). How does the microarray perform relative to other well established technologies such as the ELISA? Which microarray format or detection process is most accurate and sensitive? In order to compare performance criteria (accuracy, precision, sensitivity, dynamic range), it is important to define these criteria in the same terms. Unfortunately, reporting values exhibit little consistency. Reporting is largely a matter of preference or accepted practice within a particular field of study. The following section reviews the results of studies reporting on sensitivity and using equivalent terms. Hopefully this will lead to a better understanding of what has been accomplished and what remains to be learned.

Sensitivity and dynamic range

It is most important to agree on the definition of assay sensitivity. Operationally this sensitivity is usually taken at the lower limit of detection (LLD

or LOD) or minimally detectable dose (MDD) whose value must be greater than the precision (standard deviation or coefficient of variation) in measuring zero dose value (Ekins et al., 1990). For example, sensitivity in a noncompetitive immunoassay with n = 3 replicates would be defined as LLD = B_0 + 3 SD_{B0}, where B_0 is the zero dose (minus analyte) sample value. For a competitive assay, LLD = B_0 – 3 SD_{B0}. However, fluctuation in the standard deviation due to issues such as pipetting errors or sample interferences can dramatically alter the LLD.

It is therefore essential that the variability of an assay be known precisely (Ezan and Grassi, 2000). The performance of an assay in terms of accuracy, reproducibility (CV, interassay variation), and repeatability (CV, intraassay variation) should be determined. For an ELISA, accuracy in the range of 85 to 115% of the standard value and CVs in the range of 15 to 20% are common. The assay limit of quantification is then taken as the lowest concentration of analyte that provides CVs under, e.g., ≤10% and accuracy within, e.g., ±15% of the standard value. A discussion of factors leading to imprecision of microarrays will be addressed later.

First, however, we will examine specific examples of measured sensitivity in microarray assays. These are predominately antibody or antigen microarrays in which a micro-ELISA has been evaluated. The antibody array format involves the immobilization of a library of different capture antibodies and is commonly employed in the standard sandwich immunoassay in which antigen is captured and detected by a labeled secondary antibody. This format requires that the capture antibody be capable of efficient binding of antigen, i.e., it has been immobilized largely in its native state. In case of the antigen array format, the antigen is attached directly to the surface and probed with a reporter (labeled) antibody. The assay requirements for this format are less stringent. Immobilized antigen may be surface denatured and still be recognized by the reporter. This is possible when surface denaturation leads to the linearization of previously unavailable epitopes permitting enhanced recognition, especially by polyclonals.

Table 6.2 summarizes many of the studies discussed in this review. For comparison, the analyte LLD values (pg/mL) have been recalculated in terms of analyte concentrations and reported in p*M* units in an attempt to easily identify assay formats leading to higher sensitivity. Where possible, we will also discuss dynamic range.

A standard sandwich ELISA for IL-4 is capable of detecting approximately 10 pg/mL at the MDD and may have a linear dynamic range from 30 pg/mL to about 2000 pg/mL (R&D Systems www.rndsystems.com). The assay is thus capable of detecting IL-4 analyte concentrations in working ranges from nM to p*M* levels. Most reported work indicates that antibody microarrays operate well within this range, and a number of examples demonstrate sensitivity achieved at the fM level of detection. Certainly, second and third generation ultrasensitive standard ELISA assays can achieve fM and aM sensitivity as well. The real advantages of the microarray-based assay over a standard assay are multiplexing and parallel processing

Table 6.2 Assay Sensitivity for Antibody–Antigen Microarrays and Standard Sandwich ELISA

Analyte	LLD (pg/mL)	LLD (pM)	Signal	Detection	Microarray support	Reference
			Antibody–Antigen Microarrays			
IgG	40 fg/spot	1,000	ECL	CCD camera	Glass	Joos
	8 fg/spot	200			Membrane	
IgG	15,000	100	DBCy5	CCD microscope	Plastic	Silzel
IgG	13,400	89	ELF	CCD camera	Glass	Mendoza
IgG	5,000	10.7	Cy5/Cy3	Confocal scanner	Glass	Haab
FKBP12	150	12.5	Cy5	Confocal scanner	Glass	MacBeath
Il-2	4	~2	ECL	Film	Membrane	Huang
Il-4	10	~0.7	Cy5	PWG, CCD camera	Glass	Pawlak
Il-8	1	0.15	ELF	CCD camera	Plastic	Matson
Il-6	4	0.15	ELF	CCD camera	Glass	Wiese
TSH	2	~0.1	Texas Red, FITC	Confocal microscope	Plastic	Ekins
IL-6	0.8	0.03	CL	CCD camera	Plastic	Moody
IL-8	0.25	0.03	RLS	CCD camera	Plastic	LaBrie
HGF	0.5	0.006	TSA-Cy3	Confocal scanner	Glass	Woodbury
PSA	0.1	0.0003	RCA	CCD camera	Glass	Schweitzer
			Standard Sandwich ELISA			
Il-4	10	~0.7	HRP colorimetric	Microplate reader	96-well plastic	R&D Systems
IL-6	0.04	~0.02				
Il-8	10	1.5				
IL-4	2	0.14	HRP colorimetric	Microplate reader	96-well plastic	Pierce Endogen
IL-6	1	~0.04				
IL-8	2	0.3				
TSH	0.2 μIU/mL	1.4	HRP colorimetric	Microplate reader	96-well plastic	Biocheck
TSH	0.05 μIU/mL	0.36				Biotech
TSH	0.01 to 0.02 μIU/mL	0.07 to 0.14	Third generation texts			

Note: ELF = enzyme labeled fluorescence. ECL = enhanced chemiluninescence. CL = chemiluminescence. RLS = resonance light scattering. TSA = tyramide biotin amplification, sAV-Cy3. RCA = rolling circle amplification.

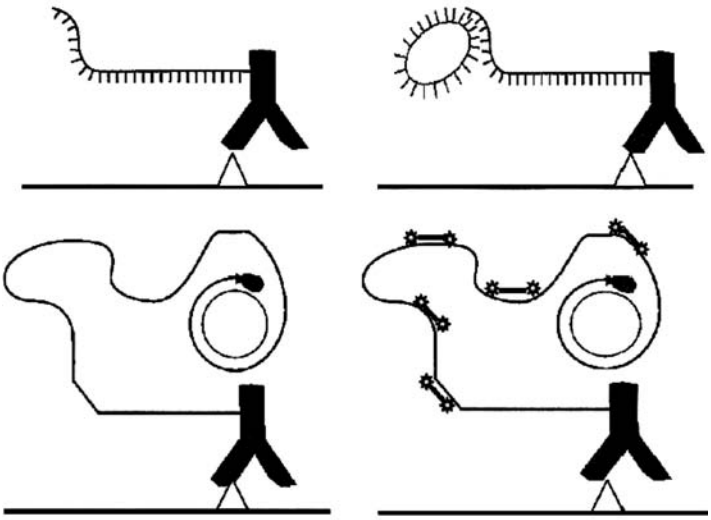


Figure 6.18 ImmunoRCA Assay. Top left: A reporter Ab conjugated to an oligonucleotide binds to a test analyte captured on a solid surface by a covalent attachment or by a capture Ab. Top right: DNA circle hybridizes to a complementary sequence in the oligonucleotide. Bottom left: Resulting complex is washed to remove excess reagents, and the DNA tag is amplified by RCA. Bottom right: Amplified product is labeled *in situ* by hybridization with fluor-labeled oligonucleotides. (From Schweitzer, B. et al., *Proc. Natl. Acad. Sci. USA*, 97, 10113–10119, 2000. With permission.)

abilities. Miniaturization that reduces the amounts of reagents and analyte is another attractive feature.

It is apparent that signal amplification provides increased sensitivity over direct labeling. This is especially true for fluorescent-based assays. One of the most sensitive signal detection technologies is the immunoRCA (Schweitzer et al., 2000). Rolling circle amplification (RCA) is combined with antibody detection. RCA involves the amplification of circularized oligonuceotide probes under isothermal conditions by DNA polymerase (Lizardi et al., 1998). With immunoRCA, the 5′ primer is attached to the reporter antibody. Initiation of the amplification starts when circular DNA template binds to the attached primer.

In the presence of DNA polymerase, the 5′ primer is extended as the circular DNA template rolls along the extended primer. The result is the formation of replicate copies of the original template laid out in linear fashion and still attached to the antibody. Smaller labeled oligonuceotide probes can then hybridize to portions of the amplified primer at multiple sites. The signal is thus amplified by the number of labeled probes (Figure 6.18). For example, PSA was detected at 0.1 pg/mL (300 zeptomoles) on a microarray corresponding to a sensitivity of about 0.3 fM. This is 3 logs more sensitive than a PSA ELISA. In other experiments, IgE was detected at 1 pg/mL with a dynamic range of about 5 logs (Figure 6.19).

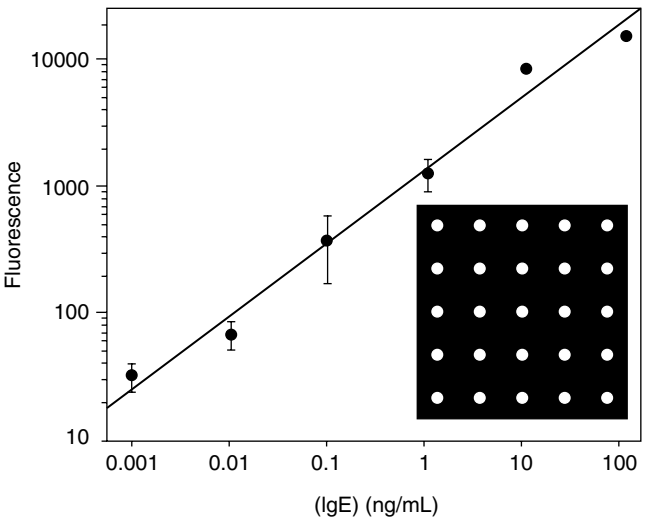


Figure 6.19 ImmunoRCA antihuman microarray dose response for purified IgE. Signals from six microarray spots were averaged for each point and the background (no IgE) signal was subtracted. Inset: Microarray scanner image of antihuman IgE array incubated with 1 ng/mL IgE. (From Schweitzer, B. et al., *Proc. Natl. Acad. Sci. USA*, 97, 10113–10119, 2000. With permission.)

Schweitzer and coworkers (2002) created antibody arrays on thiolsilane-coated glass slides immobilizing a library of monoclonal antibodies directed toward various cytokines in quadruplicate. Each glass slide was partitioned by Teflon barriers into 16 circular (0.5-cm diameter) subarrays. Each subarray contained 256 features, and 75 cytokines could be determined at ~10% CV for single analyte replicates and ~25% CV for various antibodies in terms of signal variation. Arrays stored dry at 4°C were stable for at least a month. Again, RCA was used for detection.

The authors noted that signal amplification was necessary to achieve sensitivities below 1 ng/mL (fluorescence signal without amplification) for cytokines. The RCA immunoassay was able to achieve in specific examples a 1000-fold sensitivity over direct detection (Figure 6.20). Of the 75 cytokines tested, the following sensitivities were noted: 45 = ≤10 pg/mL, 22 = ≤100 pg/mL, and 8 = ≤1000 pg/mL. The dynamic range was approximately three orders of magnitude with a precision of ~5% CV reported for four assays. These results were achieved by serial dilutions of purified cytokine antigens.

While in most instances the RCA microarray immunoassay results were comparable to published performance reports on commercial ELISA (Quantikine, R&D Systems), there were notable exceptions in which ELISA appeared to out-perform the microarray by 10- to 50-fold sensitivity. Such differences may be due to variations in binding affinities exhibited by the particular capture antibodies employed in the assay. In studies involving lipopolysaccharide (LPS)-induced secretion of cytokines from human

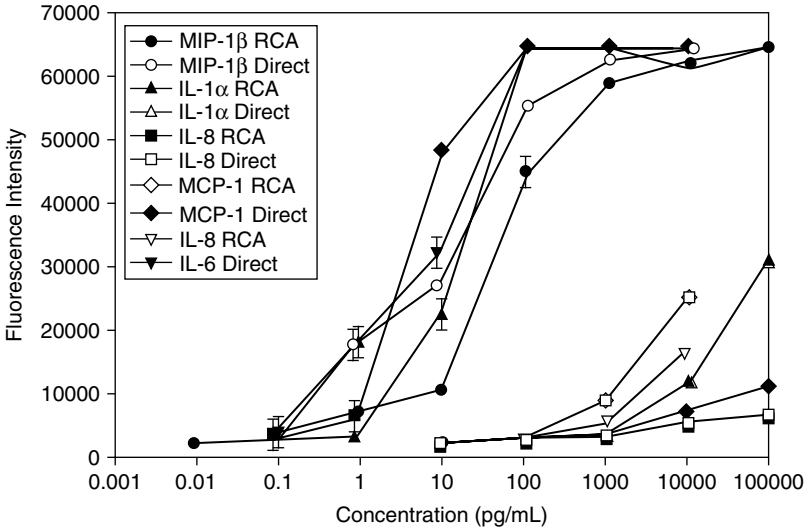


Figure 6.20 ImmunoRCA vs. direct detection. (From Schweitzer, B. et al., *Nature Biotechnol.*, 20, 359–365, 2002. With permission.)

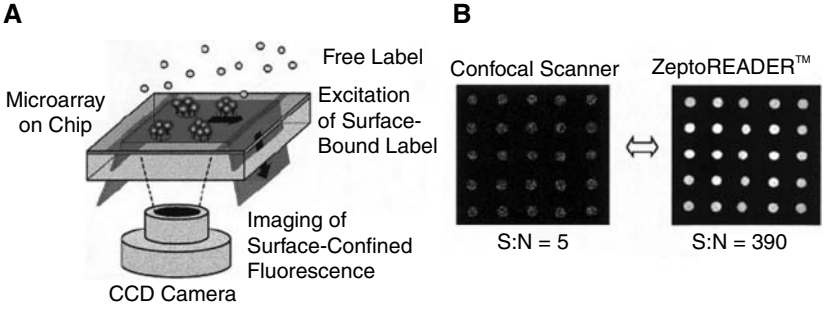


Figure 6.21 Planar waveguide technology. (From Pawlak, M. et al., *Proteomics*, 2: 383–393, 2002. With permission.)

dendritic cells, the RCA microarray immunoassay was found to be concordant with published data from standard ELISA. Unfortunately, these investigators did not include side-by-side comparisons of ELISA using LPS-induced samples except in monitoring a single analyte. The assay results were concordant for that single example.

Another approach to enhanced sensitivity relies upon improved detector designs such as those employing planar wave guides (Pawlak et al., 2002). Planar waveguide (PWG) technology involves the excitation and detection of surface-confined fluorescence (Figure 6.21). The net result is to improve the signal-to-noise ratio, thereby lowering the detection limit. For example, Cy5-labeled IgG spotted on a glass substrate was detected at a LOD of 2 p*M* and over a dynamic range of 3.5 logs. The LOD for IL-4 was determined to

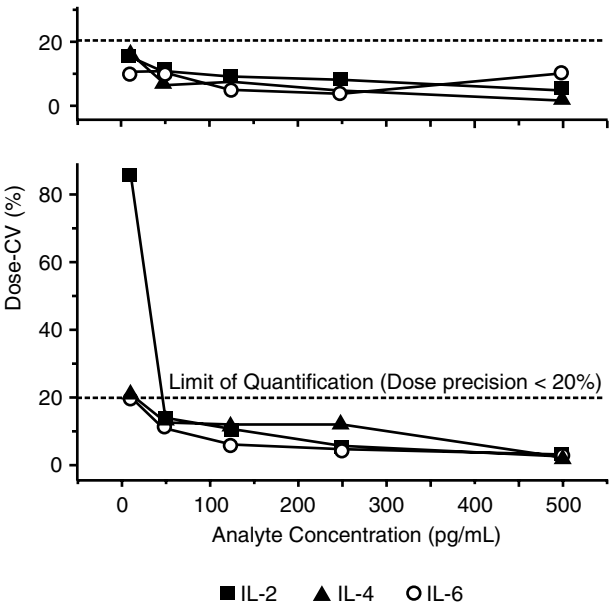


Figure 6.22 PWG assay dose precision profiles of multiplexed three-analyte immunoassay for two sets of experiments. Dose precisions correspond to standard deviations of analyte concentrations that were back-calculated using corresponding dose response curves. (From Pawlak, M. et al., *Proteomics*, 2, 383–393, 2002. With permission.)

be 10 pg/mL from a single Cy5 label. The limits of quantification (LOQ) were also estimated from the dose precision profile (Figure 6.22). The interassay variation was <10%, while the intraassay (chip-to-chip) variation was within 20%, providing an LOQ for IL-4 of 10 pg/mL.

Detection by resonance light scattering (RLS) involves the use of small gold beads that have been immobilized to reporters such as antibodies (Figure 6.23). Yguerabide and Yguerabide (1998 and 2001) discovered RLS when they observed that colloidal gold particles when illuminated by a narrow beam of white light were able to scatter light, giving rise to intense color. They found that 60-nm gold particle scattering power is equivalent to the fluorescence of 500,000 fluorescein molecules. Suspensions can be detected by eye in the fM range.

The nanometer diameter beads possess intrinsic vibrational resonances that transfer from one bead to another when they are in close proximity. The result is a change in color from red to blue when the beads are in close proximity, for example, when the reporter (bead) antibodies are localized on a surface during binding to antigen. The intensity (I) of the scattered light is particle size (diameter)-dependent, increasing by the sixth power of the radius, $I \propto r^{6}$. The resulting spectra obey Rayleigh's theory of small particle (≤40-nm diameter) light scattering.

Silver and gold have the highest light scattering powers. Silver scatters at 380 nm (purple), while gold scatters at 520 nm (green). In terms of molar

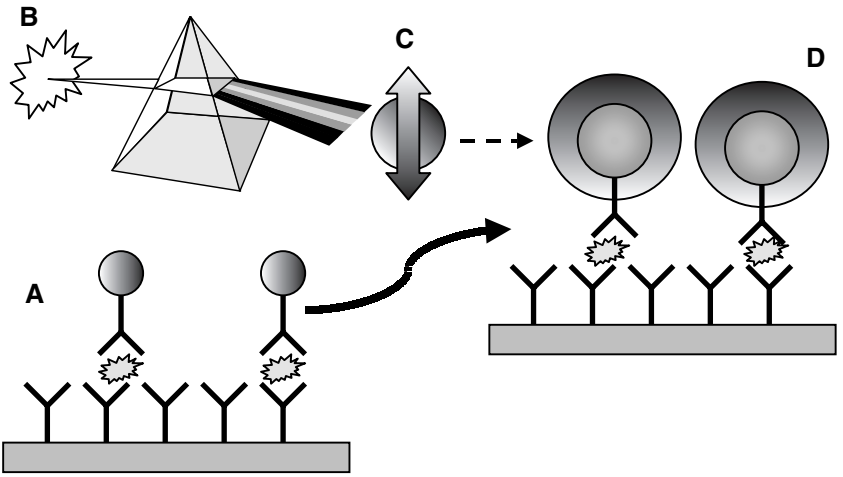


Figure 6.23 Resonance light scattering. (From Yguerabide, J. and Yguerabide, E.E., *Anal. Biochem.*, 262, 137–156 and 157–176, 1998. With permission.)

extinction coefficient (ε) for silver, $\varepsilon = 1.68 \times 10^{10}$; for gold, $\varepsilon = 5.88 \times 10^{9}$; for copper, $\varepsilon = 2.59 \times 10^{9}$; and for fluorescein, $\varepsilon = 6 \times 10^{4}$. Thus, RLS is capable of ultrahigh sensitivity relative to fluorescein. In terms of particle size, ε for gold at 30 nm-diameter = 5.9×10^{9}; at 40 nm, $\varepsilon = 1.6 \times 10^{10}$; at 60 nm, $\varepsilon = 5.3 \times 10^{10}$; and at 80 nm, $\varepsilon = 1.1 \times 10^{11}$. The dynamic range for the light scattering intensity is approximately 3 logs with fM sensitivity.

In an ELISA for toxin A (the cause of dysentery), a sensitivity was reached at 10^{-14} M with 1 hr incubation at >95% specificity. Researchers at Incyte Genomics demonstrated sub-p*M* sensitivity in an antibody array employing RLS (LaBrie, 2001). Antibodies were arrayed in 96-well microtiter plates in a 10×10 pattern with 290-μ center-to-center spacing. Assays were developed using antibiotin RLS particles for direct detection of biotinylated sample proteins (cytokines) or by biotinylation of a secondary (detection) antibody in a sandwich assay format. Similar results were obtained in terms of sensitivity and dynamic range. For IL-8, a limit of detection was achieved at 0.25 pg/mL with a linear dynamic range from 0.5 to 10,000 pg/mL. This was compared to a standard ELISA found to have a sensitivity at 1 pg/mL with a linear dynamic range < 2 logs.

Tyramide signal amplification (TSA; PerkinElmer Life Sciences, Boston) and enzyme-labeled fluorescence (ELF; Molecular Probes) are related detection technologies. In the tyramide amplification process, a tyramide–biotin complex is produced by the action of horseradish peroxidase. The complex precipitates near the binding site and accumulates. The complex is detected by the use of streptavidin–Cy3/Cy5.

ELF involves a soluble substrate (ELF–phosphate) for alkaline phosphatase that cleaves the substrate into an insoluble and highly fluorescent product (ELF–alcohol). For immunoassay purposes, the secondary antibody

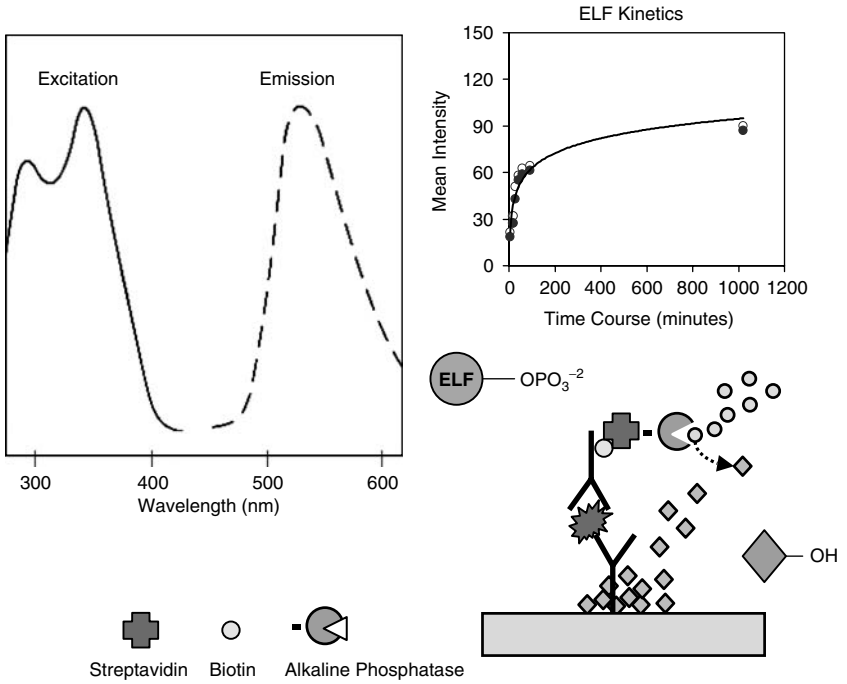


Figure 6.24 Enzyme-labeled fluorescence. ELF-97 is a soluble phosphorylated substrate cleaved by alkaline phosphatase into a highly fluorescent, insoluble product. (Molecular Probes, Inc., Eugene, OR.)

is first labeled with biotin which allows for binding of a streptavidin–alkaline phosphatase conjugate. Addition of the ELF substrate results in the accumulation of the ELF–alcohol at the site of attachment, precipitating over the captured antigen (Figure 6.24).

Various chemiluminescence (CL) signal amplification methods may also be used with microarrays. Most of these rely upon the formation of an unstable intermediate that decays to release light. The intermediate is a charged species that is sequestered as the result of a surface charge. The generated light becomes localized on the surface and can be detected with films, phosphorimagers, or CCD camera systems. Most CL detection is employed for work with membrane arrays involving intrinsic fluorescent backgrounds and light scattering issues.

Other microarray formats useful for proteomic applications

mRNA–protein fusions

While the popularity of antibody–antigen microarrays remains on the increase, other approaches are also useful in proteomic studies. Phylos technology (Weng et al., 2002) makes use of mRNA–protein fusion products that

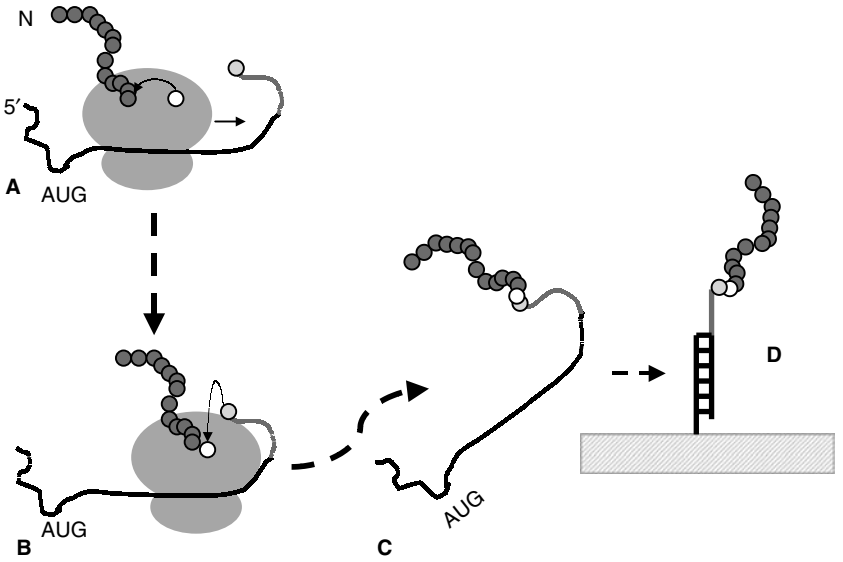


Figure 6.25 PROfusion. (Phylos, Inc., Lexington, MA.)

self-assemble onto microarrays comprised of complementary oligonucleotide capture probes (Figure 6.25).

The PROfusion process involves *in vitro* translation of modified mRNAs using the rabbit reticulocyte system. The mRNA species are conjugated to puromycin via an oligonucleotide linker. The nascent polypeptide is assembled on the ribosome (Figure 6.25, Step A). Protein synthesis proceeds unchecked until reaching the RNA–oligonucleotide junction. At this point the puromycin and C terminus of the peptide are linked — halting further extension of the nascent peptide (Figure 6.25, Step B). The result is the formation of the mRNA–protein fusion product that is then released from the ribosome following the addition of metal chelators (Figure 6.25, Step C).

Capture oligonucleotide probes complementary to the mRNAs are arrayed onto a glass substrate with immobilization via a 3′ terminal amino group. Since the mRNA is 5′ to the polypeptide, capture results in orientation of the polypeptide away from the surface (Figure 6.25, Step D). The capture probes are tethered to the substrate indirectly via a hydrophilic spacer arm composed of hexaethylene oxide–phosphodiester units that terminate with the 3′ amino group.

The oligonucleotide array is converted into a protein array by the addition of mRNA–protein fusion products to the support. The proteins are presented in proper orientation for efficient binding. The hydrophilic spacer arm further extends the protein away from the surface, allowing greater access to binders and potentially lowering the level of nonspecific adsorption. Self-assembly can be accomplished from the crude reticulocyte product or from a partially purified form to further reduce nonspecific background. The investigators were able to demonstrate specific protein capture in the

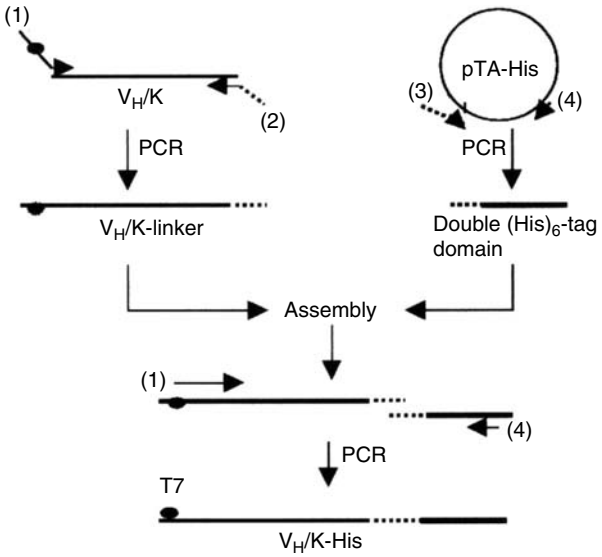


Figure 6.26 Protein *in situ* array (PISA) constructs. (From He, M. and Taussig, M.J., *J. Immunol. Methods*, 274, 265–270, 2003. With permission.)

range of 1 to 4000 attomoles/200–μ diameter spot. The authors note how "tremendous is the resolving power of the biochip format" to accomplish such efficient self-assembly even from a crude extract.

Protein in situ *array (PISA)*

Another approach to creating arrays of proteins is PISA, a technique first described by researchers at the Babraham Institute, Cambridge, U.K. (He and Taussig, 2001). The gene for a specified protein is first amplified by polymerase chain reaction (PCR) and linked, for example, to a His tag domain sequence to create a tagged gene PCR construct. An upstream primer (G/back) for gene-specific amplification containing a T7 promoter overlap is used with a downstream primer (G/for) containing a tag domain overlap sequence. The resulting amplicon therefore contains the gene sequence plus the T7 and His tag overlaps.

The tag sequence is amplified from plasmid pTA-His that provides a flexible linker and a double $(His)_6$ tag. The two amplicons can then be assembled and amplified by using an upstream T7 primer containing the promoter plus Kozak sequence and start codon for translation. A T-term/for primer is used downstream to amplify across the tag and linker. The resulting construct contains the gene, the promoter, and the tag (Figure 6.26). This is key to the technology for the construct is then used for cell-free expression of the protein.

Since the protein is already tagged with His, for example, it can be directly immobilized to a Ni–NTA support such as the bottom of a microtiter

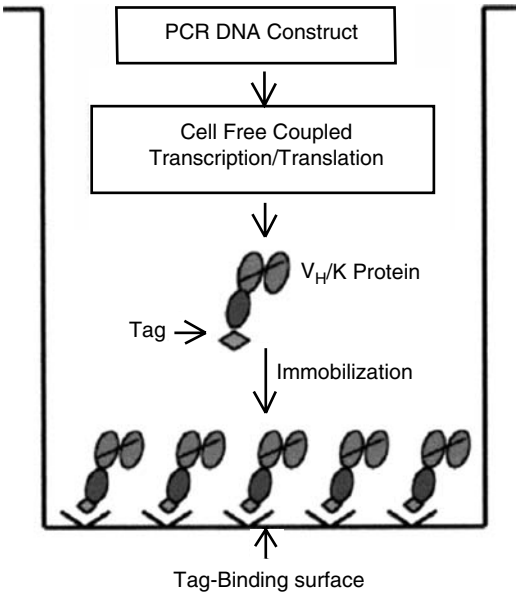


Figure 6.27 In-well PISA and His tag-expressed protein immobilization. (From He, M. et al., *J. Immunol. Methods*, 274, 265–270, 2003. With permission.)

plate well (Figure 6.27). The entire process from cDNA (or mRNA) reverse transcriptase (RT)-PCR amplification coupled to *in vitro* transcription–translation (e.g., using a rabbit reticulocyte system) and immobilization of tagged protein requires 10 hr. In one example, a human single-chain antibody fragment (V_H/K) anti-progesterone was obtained by this process. Approximately 120 ng from a 25-μL reactin volume was generated; approximately 50% of the tagged antibody was immobilized (He and Taussig, 2001) These researches recently provided a detailed protocol for the DiscernArray™ commercial product (2003).

Aptamers

Another nucleic acid technology that has demonstrated utility in proteomics involves the use of aptamers and photoaptamers (Zichi et al., 2002). Aptamers (a word derived from the Latin *aptus*, meaning *to fit*) are oligonucleotides that selectively bind proteins. Jayasena (1999) prepared a comprehensive review of aptamer technology focusing on diagnostic utility.

The pioneering work on aptamers was done by Gold and collaborators at the University of Colorado, Boulder (1990) and by Ellington and Szostak at the University of Texas, Austin (1990). These ligand-binding oligonucleotides are selected by a combinatorial process known as SELEX (systematic evolution of ligands by exponential enrichment). A random library of oligonucleotides is prepared with additional fixed sequences for amplification. A purified antigen is mixed with the library and the complexes are captured on nitrocellulose

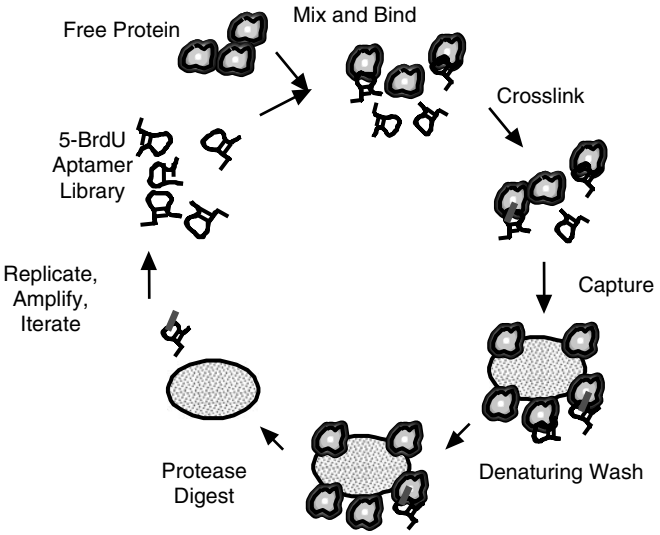


Figure 6.28 The photoSELEX process. (Copyright SomaLogic, Inc., Boulder, CO.)

membrane. The bound oligonucleotide–protein complexes are denatured and the aptamer candidates amplified for a second round with the putative protein. The selection rounds continue until the population of aptamers is reduced to high-binding candidates that are then fully sequenced.

Enrichment of high affinity candidates is usually achieved in 8 to 15 rounds of SELEX. Each round takes approximately 2 days to perform. The process has been automated using robotic liquid handlers both for DNA (SomaLogic) and RNA aptamers (Cox, 2002). Next, the sequenced aptamer is prepared in bulk by conventional DNA synthesis chemistry and purified, then the aptamer arrayed onto a solid support. Thus, an aptamer is ready for application within 2 to 3 mo. Because the sequence is known, preparation of additional aptamer is easily accomplished using conventional oligonucleotide chemical synthesis.

How well aptamers perform relative to antibodies remains an open debate. In earlier work involving a survey of 100 aptamers, more than 75% were characterized by $K_d \leq 1$ nM (Brody et al., 1999; Brody and Gold, 2000). Such affinities are well within the range exhibited by antibodies, so it is quite plausible that libraries of aptamers could replace antibodies as general tools for certain applications such as antigen screening. Specific aptamers with higher K_d values could be selected for use in affinity chromatography purification processes. However, for diagnostic applications, sensitivity and specificity must be rigorous. K_d values in the p*M* to aM range are required and nonspecific binding to the oligonucleotide aptamer must be minimal.

A more recent version of this process known as photoSELEX addresses these concerns (Figure 6.28). Aptamers with photolabile Br-dUTP incorporated

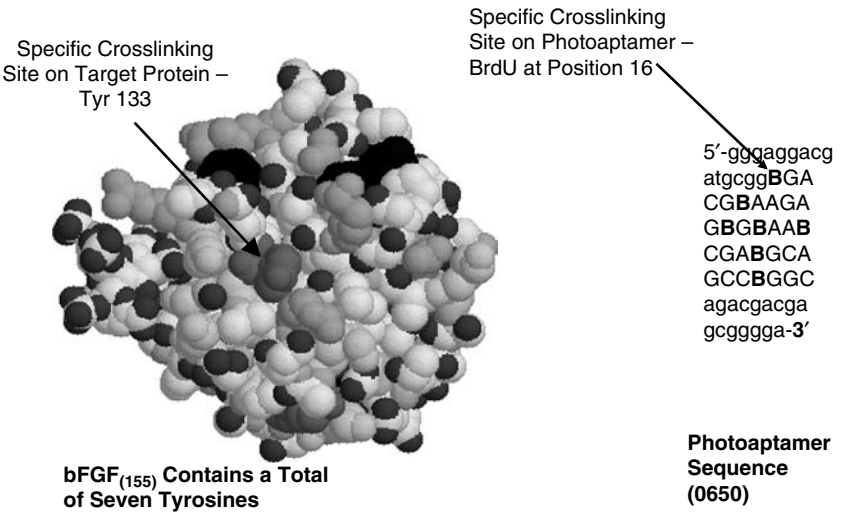


Figure 6.29 (See color insert following page 116.) Photoaptamer crosslink. (Copyright SomaLogic, Inc., Boulder, CO. Reproduced with permission only.)

into specific regions on the oligonucleotides are produced (Willis et al., 1994). When the protein becomes bound to the photoaptamer, the aptamer and protein are crosslinked by laser excitation of the Br-dUTP groups that are in close proximity to specific amino acid residues of the protein (Golden et al., 2000) as shown in Figure 6.29.

Since the cognate protein is now covalently linked to the aptamer, nonspecific protein can be extensively washed from the array under harsh conditions. With the ability to substantially reduce the background, the array can be read with greater sensitivity. Since only the cognate protein remains after the rinse, it is possible to use general protein staining reagents for detection.

Researchers at SomaLogic created aptamer arrays on glass slides and used NHS-Alexa 555 dye staining of the captured cognate proteins following a harsh rinse under denaturing conditions. As an example, thrombin could be detected at 100 p*M* in a multiplexed assay involving a total of eight different proteins. Total protein concentration was 11.1 nM with the concentrations of individual proteins ranging from 0.01 to 10 nM. Endostatin and VEGF were both detectable below 100 p*M* (Figure 6.30).

The combination of the photoSelex process with the denaturing rinse steps provides for increased specificity. Improvements in general protein staining are needed to further increase sensitivity into the low- to sub-p*M* ranges in order for the technique to be competitive with current ELISA and antibody microarray assays (see Table 6.2). In recent studies, proteins have been detected in the p*M* range using an ALONA (antibody-linked oligonucleotide assay) sandwich aptamer–antigen–reporter antibody. The

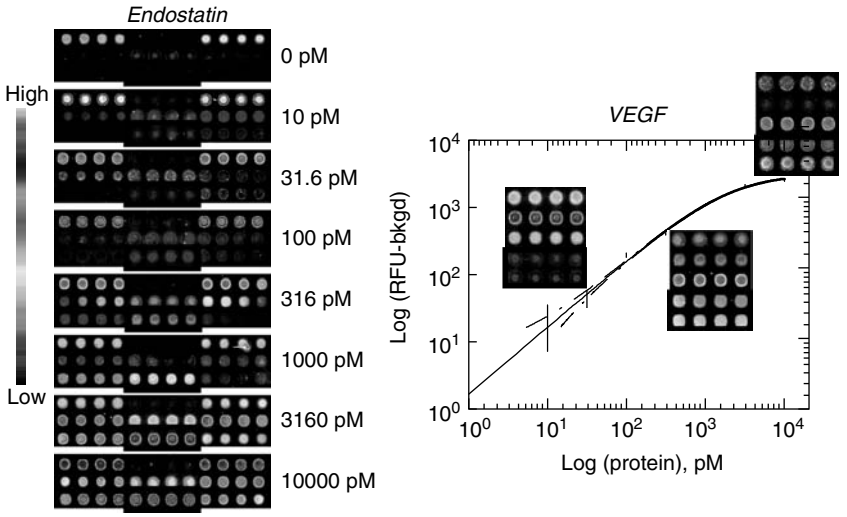


Figure 6.30 (See color insert following page 116.) Aptamer dose response. (Copyright SomaLogic, Inc., Boulder, CO. Reproduced with permission only.)

aptamer provides capture specificity while the secondary antibody coupled with signal amplification allows increased sensitivity (Matson, unpublished data).

Universal protein array

Ge (2000) carefully designed a universal protein array (UPA) system based on the use of various transcription factors, their activators, and cofactors as probes. A total of 48 different, highly purified factors were used to create the UPA on nitrocellulose filters. Protein–protein interactions of various binding affinities could be assessed using different ionic strength buffers (e.g., 100 mM KCl vs. 1000 mM KCl). The relative binding of radiolabeled (^{32}P) GST-K-p52 proteins to various transcription factors was studied.

The p52 specifically interacted with nucleolin but not with topoisomerase I, thereby supporting the observation that p52 associates with nucleolin as a multiprotein complex in HeLa cells. Nucleolin is thought to be involved in pre-mRNA splicing and the unwinding of DNA–RNA duplexes as well as mediating cell doubling time in human cancer cells. The p52 protein is a general transcriptional cofactor capable of potentiating activated transcription of class II genes. It also serves as a pre-mRNA splice regulator. The UPA system may serve as a new tool in discovering new components involved in gene regulation and in the evaluation of therapeutic protein targets.

Zyomyx introduced a microfabricated microarray system based upon the use of pillars that serve as platforms for depositing capture antibodies. Relying on substrate features such as ultraflatness and surface modification, the

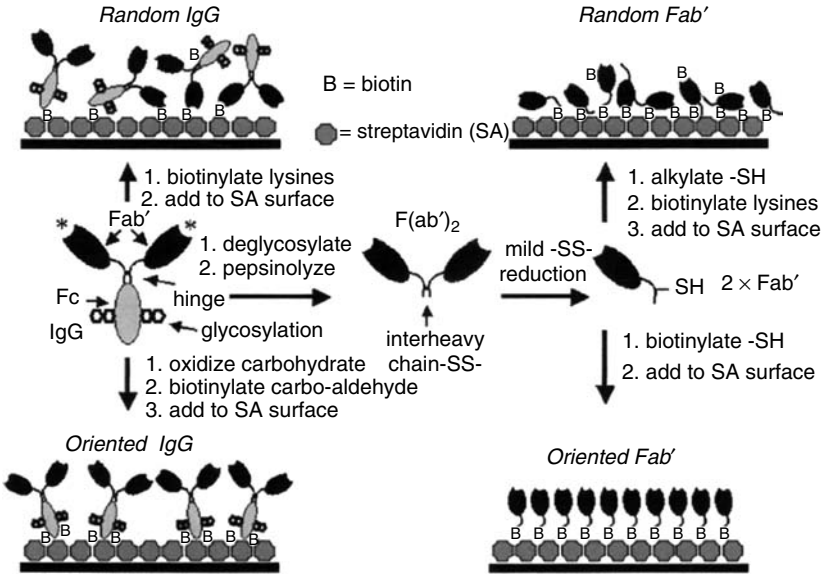


Figure 6.31 Solid-phase antibody density and steric hindrance effects. (From Matson, R.S. and Little, M.C., *J. Chromatogr.*, 458, 67–77, 1988. With permission.)

company created an environment for optimal protein–protein interaction. Precise delivery of reagents to each pillar enables robust assay development.

Peluso et al. (2003) investigated strategies for antibody immobilization using their chip format. Four immobilization strategies for placement of capture antibody were studied: random vs. oriented coupling of IgG and random vs. oriented coupling of antibody fragments (reduced) from pepsin digest (Fab′) (Figure 6.31). Oriented coupling of IgG antibody to solid supports is well documented in the literature as having distinct advantages over random tethering of antibodies. Foremost is the loss of antigen-binding activity due to steric hindrance caused by the antibody being too near the surface or crowded by adjacent antibodies, thereby blocking the binding region (Figure 6.32).

Surface denaturation of the antibody by excessive attachment at multiple sites on the surface can also occur. Oriented coupling of the antibody at an optimal surface density has been demonstrated to overcome many of these issues (Matson, 1988). Streptavidin (sAV) was used as means to orient antibodies that had been site-specifically modified with biotin vs. randomly biotinylated IgG and Fab′. In addition, two different surface modifications that yield sAV monolayers were studied: b-SAM (biotinylated self-assembled monolayer) and poly-L-lysine (PLL)–polyethylene glycol (PEG)–biotin (Figure 6.33). The b-SAM was formed on gold-coated glass slides treated with a SAM of oligo-(ethylene glycol) containing alkane disulfides terminating with N-hydroxysuccinimide (NHS) reacted with tri-(ethylene glycol) amino biotins.

Incubation with sAV led to the creation of the sAV monolayer. This surface was used exclusively with the BIAcore system (Biacore International

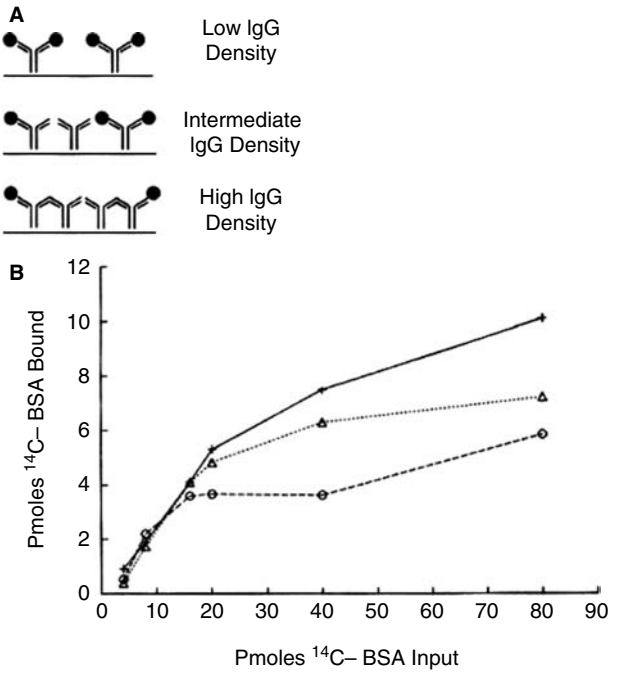


Figure 6.32 Strategies for immobilization of capture antibodies. (From Peluso, P. et al., *Anal. Biochem.*, 312, 113–124, 2003. With permission.)

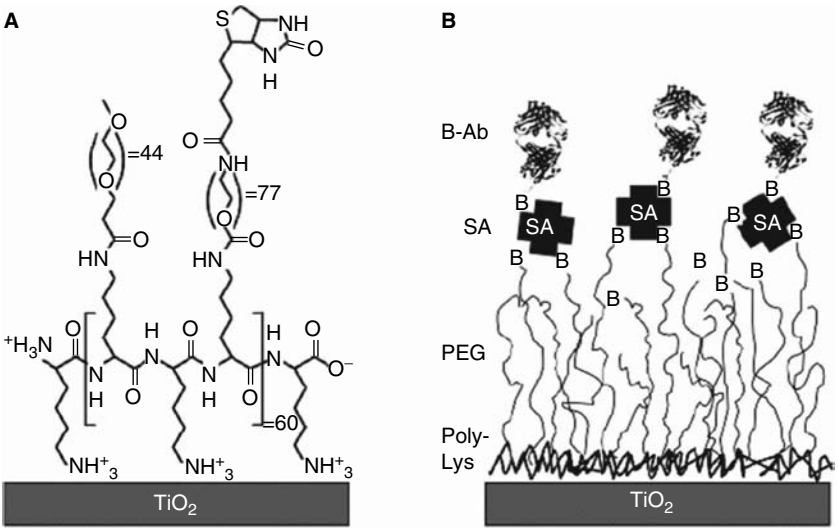


Figure 6.33 Zyomyx's b-SAM and PLL–PEG–biotin surface coatings. (From Nock, S. and Wilson, D.S., *Angew. Chem. Int. Ed.*, 42, 494–500, 2003. With permission.)

AB, Switzerland) in SPR (surface plasmon resonance) studies. The PLL–PEG–biotin–sAV was used for microarray studies as described for the commercial product. A detailed description of the design and physical and chemical characterization of the PLL-grafted PEG monolayer is provided by Ruiz-Taylor et al. (2001).

sAV could be reproducibly tethered to the b-SAM surface at about 2×10^{12} molecules/cm^2 (~4 pmoles/cm^2). Not surprisingly, oriented Fab′ could be immobilized at the highest density with a 70% retention of the calculated specific binding activity. For the three antibodies studied, the oriented Fab′ was 1.8-fold to 5.6-fold more active than Fab′ biotinylated by a random process. In the case for immobilization of the full-length antibodies, the oriented antibodies exhibited higher binding activity on average but the increase over randomly biotinylated IgG was only ~1.3-fold. With the b-SAM, surface-oriented Fab′ bound 49% more antigen than oriented IgG antibody.

For microarray work, PLL–PEG–biotin–sAV monolayers adsorbed onto underlying titanium dioxide surfaces (pillars) replaced the b-SAM gold-coated glass surface. While oriented capture agents out-performed their random counterparts, distinct differences in Fab′ and antibody performance between this new surface and the b-SAM surface were apparent.

Under certain conditions, the oriented antibody out-performed the oriented Fab′ system: At high antigen loading (100 nM), oriented antibody was threefold higher than oriented Fab′. At 100-p*M* loading (K_a antibody ~100 pM), the Fab′ was threefold higher than antibody. One explanation offered by the authors is that the tethered antibody provided a higher number of antigen binding sites per surface area but were of lower antigen affinity ($K_a' >> K_a$) than the Fab′ sites. This could be the result if the antibody were damaged during periodate oxidation while Fab′ biotinylation was less destructive. However, this does not explain why these differences were not observed on the b-SAM surface. The PLL–PEG surface provided a higher fold change in random vs. oriented tethering of the capture antibodies and Fab′ fragments. Perhaps the spatial distribution and relative density of antibody vs. Fab′ are more important.

The Protein Profiling Biochip™ is composed of six chips assembled in a flow cell cassette device. Each chip provides 200 data points (200 pillars per chip) for a total of 1200 data points per cassette. Pillars are 50 μ in diameter. The mesa on each pillar is covered with a self-assembled monolayer (20 to 25 Å thickness) of biotin-derivatized PLL-g–PEG groups. A constant grafting ratio of 3.5 parts Lys to 1 part PEG is maintained with variable biotin–PEG content. sAV antibody is immobilized at 0.5 to 2 pmole/cm^2 and Fab fragments at 4 pmole/cm^2.

Each pillar is addressed by a separate capillary to apply sample and processing reagents. This is a noncontact process in which the capillary seats over a peg that has a hydrophobic coating, thereby allowing wetting only of the mesa. Oriented coupling of antibody to the surface allowed for a lower LOD of IL-8 to 1 pg/mL (0.15 pM) at total signal above twofold background (Wagner, 2002). A multiplex assay for 30 human cytokines in each of the 6

channels has recently been introduced for analysis using sample volumes as small as 40 µL.

Ciphergen (Fung et al., 2001) described the ProteinChip™ proteomics platform that combines the use of chromatographic support materials to capture proteins with the resolving power of mass spectroscopy. The ProteinChip in its current format more closely resembles a macroarray. Various surfaces borrowed from chromatographic use are prepared as large spots in a $1 \times N$ array that can be read in a surface-enhanced laser desorption time of flight (SELDI-TOF) mass spectrophotometer. The Ciphergen system is especially useful for analysis of protein differential display (profiling) in which changes in the cellular contents of proteins vary. For example, one can compare different growth conditions and then focus on the appearance or disappearance of cellular proteins isolated onto the ProteinChip.

The elucidation of differentially expressed proteins is based upon well-known protein isolation and characterization strategies. For example, Thulasiraman et al. (2001) were able to identify virulence factors in *Yersinia pestis* (a bacterium causing human plague) by examining the m/z (mass-to-charge ratio) profile differences in bacteria grown at two different temperatures. The temperature shift from 26 to 37°C induces up-regulation in the virulence genes.

Cell extracts from the bacteria grown at the two temperatures were compared for protein content by first applying them to different chip surfaces. Strong anion exchange (SAX) spots were used to adsorb negatively charged proteins, while metal affinity resin (IMAC) loaded with Cu^{+2} was used to capture His proteins. SAX isolation identified the occurrence of 14.9- and 78.8-kDa peaks present only in *Y. pestis* grown at the higher temperature.

Since both proteins bound at pH 7.4, the result suggests that both had isoelectric point (pI) values <7.4. Furthermore, the 78.8-kDa protein bound to the IMAC spot, while the 14.9-kDa protein did not bind. This suggests that the 78.8-kDA protein had a free His imidazole group on its surface, while the 14.9-kDa protein did not have an available metal binding domain.

Following a scale-up purification of bacterial lysate, the two proteins were excised from sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) bands and subjected to tryptic digestion. SELDI analysis of the peptide fragments with reference to the National Center for Biotechnology Information (NCBI) protein database resulted in a match of the 78.8-kDa protein with the KatY protein, which also produced a good fit with the protein characterization induced at 37°C: 78.8-kDa mass, 6.4 pI, and active site motifs containing His imidiazoles. Similarly, the 14.9-kDa protein was identified by tryptic digestion as the Antigen 4 protein containing a fibrillar structure required for full virulence.

Boyle et al. (2001) used the Ciphergen SELDI protein chip to analyze the secretion and autoactivation of a cysteine protease (SpeB) from *Streptococcus pyogenes* that has been implicated in the onset of group A streptococcal infections and may contribute to toxic shock symptoms. SpeB could be detected at ~0.75 ng protein in a 30-min assay based upon SELDI-TOF

analysis. This represents a >6,000-fold increase in sensitivity over conventional Western immunoblot analysis. Zymogen and activated intermediate forms of the protease could be identified and easily monitored during the culture growth. This will enable future studies on the activation process to elucidate the mechanism of action of the virulence factor.

The resolving power of mass spectroscopy is further exemplified in posttranslational modification studies of acyl carrier protein from *Mycobacterium tuberculosis*. Researchers were able to determine both the nature of the modification and also the mechanism by which the protein was regulated (Fung et al., 2001, Reference 7).

The protein profiling approach also provides the use of pattern recognition for discrimination of disease states. Biomarkers for prostate cancer were profiled and a panel assembled that could differentiate cancer patients from noncancer populations (see Fung et al., 2001, Reference 11). Poon et al. (2003) utilized the ProteinChip to obtain tumor-specific proteomic signatures to detect hepatocellular carcinoma (HCC) in patients having chronic liver disease (CLD).

The proteomic signature is based upon the separation, detection, and profiling of low molecular mass proteins in terms of their relative abundance using the SELDI-TOF (m/z) spectrum. Serum fractions from anion exchange fractionation were placed on IMAC3 ProteinChip Arrays loaded with Cu^{+2} ion as well as WCX2 (weak cation exchanger) arrays. Following processing of the arrays, each spot was subjected to SELDI-TOF and various m/z peaks identified as unique to either HCC or CLD cases. A total of 2384 mass (m/z) peak assignments were identified in samples with 1087 from IMAC3 and 1297 from WCX2 arrays.

Of these, 250 markers were found to be significantly different in HCC and CLD. This was accomplished by applying two-way hierarchical clustering and artificial neural network (ANN) for differentiation. Because the number of peptides, proteins, and other polymers within a sample can lead to generation of 8000 to 10,000 m/z values in the SELDI-TOF spectrum, the ANN has proven very useful in predicting tumor grades (Ball et al., 2002). Both methodologies exhibited specificities and sensitivities of approximately 90% in detection of HCC. This finding is similar to results of earlier work that identified ovarian cancer at 100% sensitivity and 95% specificity and detected prostate cancer at 83% sensitivity.

Wellmann et al. (2002) also examined prostate carcinoma with the aid of laser-assisted microdissection and SELDI to analyze protein extracts from about 500 cells. A number of differentially expressed proteins in the 1.5- to 30-kDa range were found between normal prostate gland cells and prostate tumors. A prominent (threefold) up-regulated protein peak at mass 4299 Da was observed for prostate tumor cells. The average relative intensities were 24.37 for tumor cells, 9.99 for transitional zone cells, and 7.26 for normal prostate cells.

Batorfi et al. (2003) used a similar approach combining laser capture microdissection and SELDI-TOF of the ProteinChip to investigate the pathogenesis of gestational trophoblastic disease by differential protein expression. In their studies of normal and molar trophoblast (tumor) cells in placentas, they identified three metal-binding polypeptides (11.3, 13.8, and 15.2 kDa) present at statistically significant lower levels in tumors.

Peptide arrays for antibody detection

Melnyk et al. (2002) created peptide arrays for detection of antibodies in blood raised against various infectious agents [hepatitis C virus (HCV) core protein, for example]. Peptides were synthesized and terminated with diglyoxyl groups that were in turn oxidized to aldehydes using periodate. Glass slides were surface treated with 3-aminopropyl trimethoxysilane and converted to the semicarbazide form by a triphosgen-mediated reaction with Fmoc-hydrazine. The glyoxylyl peptides were printed onto semicarbazide slides and the arrays were held overnight at 37°C in a humid chamber. The peptide microarrays were stored at 37°C for 6 mo without significant variation in signal or background. Slide-to-slide variation was 4.5% (standard deviation). Oxidized antibodies were also immobilized without issue.

Diluted human serum (1 μL) was incubated with the peptide microarray and bound antibodies were detected using a rhodamine-labeled anti-human IgG. Signal was detected using a slide scanner (Affymetrix model 418) with data collection in the Cy3 channel. A reported eightfold gain in sensitivity at 100% specificity over standard ELISA was achieved using the peptide microarray.

Comparison of the peptide microarray and a standard ELISA was made using a reference collection of HCV-infected and normal human sera. Concordance was observed in 117 of 130 sera. The remaining 13 samples were reported as false positives or false negatives by the ELISA. The length and position of spacer groups on the glyoxylyl peptides did not appreciably affect the binding of serum antibodies from HCV patient sera. However, the manner in which Epstein–Barr virus (EBV) peptides were immobilized had a significant impact on sensitivity. Thus, generalizations regarding tethering of peptides to a solid support appear to be unwarranted in this case. This study clearly demonstrates the difficulty in selecting attachment chemistries for optimizing binding conditions.

Phage display antibody selection

De Wildt et al. (2000) utilized phage display techniques to screen for large antibody populations. The phages were selected to express scFv antibody fragments with binding regions for protein A and protein L. Capture and detection of the antibody fragment were possible without interference with the antigen-binding domains. Selected clones were gridded onto a large square tray, then coated with a nitrocellulose (NC) filter and the phage grown

up overnight. A second NC filter coated with protein L was overlaid with the colony filter to capture the expressed scFv phage. This filter could then be probed with protein A HRP to determine expression levels.

Up to 18,342 different colonies could be arrayed in duplicate onto a 22- × 22-cm NC filter. Fifteen of these filters could be produced by robotic gridding. The scFv grids could be used as antibody arrays for large scale protein (antigen) expression analysis. Selected scFv fragments could be released from the membrane and purified by protein A Sepharose affinity chromatography. From the purified fragment, the affinity binding constants could be determined by solution phase competition experiments using a BIAcore biosensor system.

Protein kinase microarray

We discussed applications involving ligand binding (such as using antibodies to capture antigens), but other interesting and potentially valuable protein array formats have potential. A case in point is the enzyme microarray in which the activity of the sequestered enzyme is used to assess a sample for substrate content. In the array format, a library of enzymes acting on a variety of substrate analytes would be immobilized, and the metabolic activity of cells could be assessed globally simply by measuring the content of cell extracts or spent culture media.

Perhaps the best-known example is the work on yeast protein kinase chips by Zhu et al. at Yale (2000). Disposable microwell microarrays (later described as nanowell chips) were fabricated from a poly(dimethyl siloxane) (PDMS) elastomer onto glass slides. An optimal design was preparation of wells 1.4 mm in diameter and 300 μ deep to hold approximately 300 nL. The wells were arranged in a 10 × 14 pattern on a 1.8-mm pitch between wells. Two arrays could be mounted onto a single slide so that 288 wells could be analyzed. Proteins were covalently attached to the wells via epoxide using 3-glycidoxypropyl trimethoxysilane (Figure 6.34). Conversely, substrates could be attached to the wells and the presence of enzyme in the sample determined.

For the studies involving yeast kinases, entire coding regions for the kinases were cloned into an expression vector generating glutathione-S-transferase (GST) fusion proteins rescued in *Escherichia coli*. One hundred nineteen GST–kinase fusion proteins were successfully prepared in this manner, overexpressed in yeast, and purified on glutathione Sepharose beads in a 96-well format. The purified forms were immobilized in nanowells for assay consisting of incubation of specific substrates with ^{33}Pγ-adenosine triphosphate (ATP). A total of 17 different assays were performed to simultaneously characterize the activity of 119 kinases. Following the reactions, the microarrays were rinsed and the phosphorylation signal of each kinase quantified using a phosphorimager. The investigators found that a >10-fold improvement in signal-to-noise ratio was achieved with the nanowell array over a conventional microtiter plate assay with reaction volumes at 1/20 to 1/40 the amounts needed for 384-well assays.

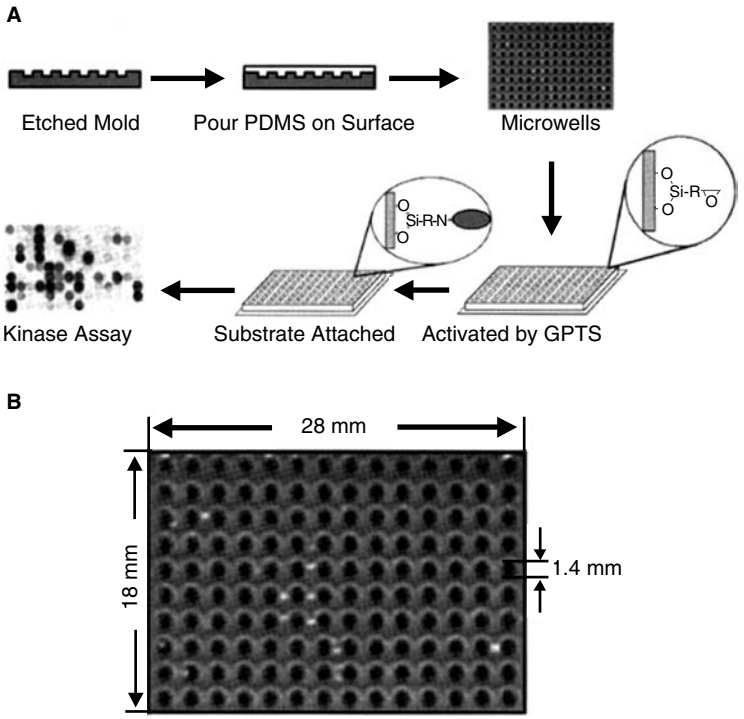


Figure 6.34 Elastomer-based microwell protein kinase assay array. (From Zhu, H. et al., *Nature Genetics*, 26, 283–289, 2000. With permission.)

Second generation proteomics

Proteomics is a means to identify and classify all proteins within an organism, tissue, or biological matrix such as serum. Functionally speaking, we must be able to accurately measure protein expression at the global level of the proteome, but there lies the problem. Protein abundance is estimated to vary over six logs of dynamic range (Tyers and Mann, 2003). Technologies currently employed to address the challenge (2D gel electrophoresis and mass spectroscopy) are regarded as lacking the necessary resolution, sensitivity, and throughput needed to successfully confront the proteome (Humphrey-Smith, 1999; Kodadek, 2001; Albala, 2001; Huels et al., 2002).

In fact, large format 2D gel electrophoresis has the ability to resolve 10,000 proteins although the process is rather tedious and suffers from low throughput (Klose and Kobalz, 1995). It is also conceivable that additional proteins may have been resolved but remain undetectable due to a lack of more sensitive protein stains. However, even with the aid of higher sensitivity staining, 2D gel electrophoresis may only be able to resolve 1 to 10% of the human proteome.

Mass spectroscopy with its tremendous mass resolving power would be the detector of choice, but it also suffers from sensitivity and throughput issues. For these reasons, Humphrey-Smith asserts that a paradigm shift toward second generation proteomics (SGP) is needed to tackle the proteome (1999). SGP is the application of array-based technologies to call out the total protein content of the proteome without reliance upon separation science (2D gel electrophoresis, mass spectroscopy, high performance liquid chromatography, etc.). Rather, the intent is to wed the power of the microarray as a parallel processor with more traditional molecular biology approaches to generate the necessary capture ligands and develop more sensitive labeling strategies.

The protein microarray represents an emerging technology. While we have described its potential utility, several key problems remain to be overcome before this tool is fully adopted by the research and biopharmaceutical communities. The most likely first embodiment will be an antibody "protein-detecting" microarray. This is understandable based upon the availability and suitability of antibody libraries originally developed for ELISA. We have discussed many demonstrations of antibody arrays in this chapter but commercial introductions (Pierce, Beckman Coulter) have been limited.

In reality, these forays represent miniaturization of the standard sandwich ELISA to attain higher throughput assays by multiplexing a limited number (<50) of analytes, e.g., cytokine panels. Even at these low densities, quantification problems arise in part due to a lack of robustness in the printing process and also in the selection and stability of monoclonal antibodies that must be highly specific and of high binding affinity to be useful for microarrays.

The issue of which antibody to select for an assay is not a new problem. Certainly anyone involved in the development of an immunoassay has been faced with this choice. Consider attempting to create a multianalyte, microarray-based micro-ELISA of modest density (10 to 100 analytes) and determining which capture antibodies to use based upon their affinities, stabilities, and cross-reactivities. For a sandwich assay, add in the 10 to 100 analyte-specific secondary (reporter) antibodies and determine their levels of cross-reactivity with each other and with the specified antigens and capture antibodies. In other words, achieving high performance for all analytes with a microarray immunoassay is indeed a formidable challenge.

How do we then envision the protein microarray as a proteomics tool? We now estimate the human genome to comprise around 30,000 genes. For gene expression analysis using DNA microarrays, ~1000 to 10,000 gene elements are often used. Since proteins undergo posttranslational modification (>200 different types; see McDonald and Yates, 2000, Reference 40) and can occur as isoforms and multiprotein complexes, the number of protein expression elements needs to be much larger.

The proteome may approach a million proteins. How will we be able to apply protein microarrays? Where and how will we obtain the necessary libraries of capture ligands on such a scale? Thomas Kodadek (2001) wrote,

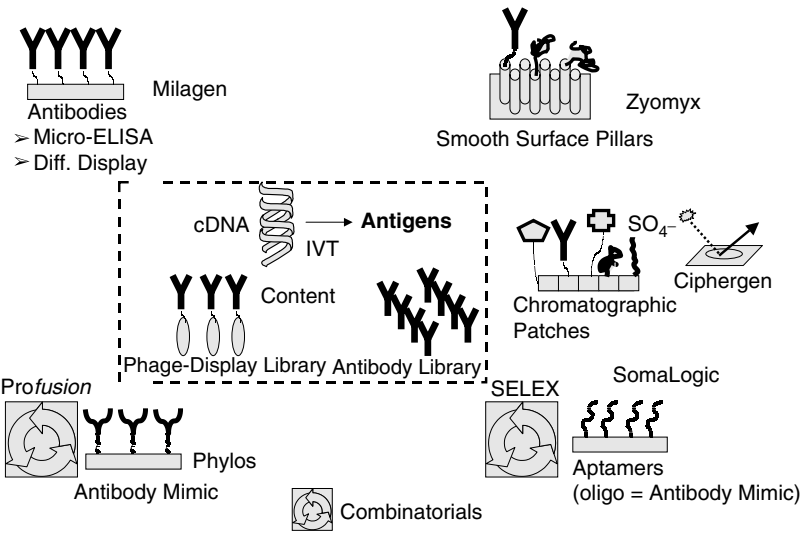


Figure 6.35 Proteome-based microarray technologies.

"While the first protein-detecting chips will almost certainly be based on monoclonal antibodies produced by traditional means, these will be largely for show." As we noted, reliance on other combinatorial-based technologies such as *in vitro* transcription, phage displays, and aptamers will undoubtedly be needed to provide the huge number of protein capture ligands required (Figure 6.35).

The other major issue for array-based proteomics is sensitivity — it is really the daunting problem with the large dynamic range in relative abundance of proteins. In addition to specific (high affinity) protein–protein interactions, most proteins exhibit some degree of cross-reactivity with low affinity binders. If such low affinity binders are of higher relative abundance than the specific analyte, it would be possible for the binders to compete with the specific, less abundant analyte (Zhu and Snyder, 2003; Phizicky et al., 2003). This problem could be resolved by a reduction in the sample complexity by fractionation or amplification steps prior to microarray analysis. However, no PCR equivalent process for amplifying low abundance proteins presently exists (Albala, 2001). Protein tagging technologies that incorporate affinity tags (e.g., biotin, His) would allow separation and enrichment of certain protein classes (Hanash, 2003).

It is also well known that proteins differ in their abilities to be chemically stained; and it would be very difficult to devise a set of calibration curves for every protein on an array. The heterogeneity of protein structure makes it unlikely that a general protein labeling strategy can be applied to arrays (Kodadek, 2001). Direct labeling may also alter protein structure, leading to denaturation or the inability to form a complex with a specific capture agent. While antibody stability on arrays appears adequate, other proteins may

become fully or partially surface denatured by the inability to fully refold into the native conformation after printing (Cahill, 2001).

In spite of these technical challenges we have shown how protein microarrays have worked very well. For instance, MacBeath and Schreiber's work (2000) on the specific detection of a FRB at a single spot on the microarray placed among 10,799 protein G spots speaks to the high degree of specificity that can achieved. While we do not have the PCR equivalent of mass amplification, we have the opportunity to use rolling circle amplification to fish out low abundance protein targets (Schweitzer et al., 2002). Finally, we have also seen Ciphergen's successful implementation of array technology with mass spectroscopy (SELDI-TOF) to identify potential cancer markers (Poon et al., 2003).

References

Albala, J.S., Array-based proteomics: the latest chip challenge, *Expert Rev. Mol. Diagn.*, 1, 145–152, 2001.

Angenendt, P., Glokler, J., Murphy, D., Lehrach, H., and Cahill, D.J., Toward optimized antibody microarrays: a comparison of current microarray support materials, *Anal. Biochem.*, 309, 253–260, 2002.

Avseenko, N.V., Morozova, T.Y., Ataullakhanov, F.I., and Morozov, V.N., Immobilization of proteins in immunochemical microarrays fabricated by electrospray deposition, *Anal. Chem.*, 73, 6047–6052, 2001.

Ball, G., Mian, S., Holding, F., Allibone, R.O., Lowe, J., Ali, S., Li, G., McCardle, S., Ellis, I.O., Creaser, C., and Rees, R.C., An integrated approach utilizing neural networks and SELDI mass spectrometry for classification of human tumours and rapid identification of potential biomarkers, *Bioinformatics*, 18, 395–404, 2002.

Batorfi, J., Ye, B., Mok, S.C., Cseh, I., Berkowitz, R.S., and Fulop, V., Protein profiling of complete mole and normal placenta using ProteinChip analysis of laser capture microdissected cells, *Gynecol. Oncol.*, 88, 424–428, 2003.

Boyle, M.D.P., Romer, T.G., Meeker, A.K., and Sledjeski, D.D., Use of surface-enhanced laser desorption ionization protein chip system to analyze streptococcal exotoxin B activity secreted by *Streptococcus pyogenes*, *J. Microbiol. Meth.*, 46, 87–89, 2001.

Brody, E.N. and Gold, L., Aptamers as therapeutic and diagnostic agents, *Rev. Mol. Biotechnol.*, 74, 5–13, 2000.

Brody, E.N., Willis, M.C., Smith, J.D., Jayasena, S., Zichi, D., and Gold, L., The use of apatamers in large arrays for molecular diagnostics, *Mol. Diagn.*, 4, 381–388, 1999.

Cahill, D.J., Protein and antibody arrays and their medical applications, *J. Immunol. Meth.*, 250, 81–91, 2001.

Cox, J.C., Hayhurst, A., Hesselberth, J., Bayer, T.S., Georgiou, G., and Ellington, A.D., Automated selection of aptamers against protein targets translated *in vitro*: from gene to aptamer, *Nucleic Acid Res.*, 30(20), 1–14, 2002.

Delehanty, J.B. and Ligier, F.S., A microarray immunoassay for simultaneous detection of proteins and bacteria, *Anal. Chem.*, 74, 5681–5687, 2002.

De Wildt, R.M., Mundy, C.R., Gorick, B.D., and Tomlinson, I.M., Antibody arrays for high-throughput screening of antibody–antigen interactions, *Nature Biotechnol.*, 18, 989–994, 2000.

Ekins, R.P., Ligand assays: from electrophoresis to miniaturized microarrays, *Clin. Chem.*, 44, 2015–2030, 1998.

Ekins, R.P., Multi-analyte immunoassay, *J. Pharm. & Biomed. Analysis*, 7(2), 155–168, 1989.

Ekins, R.P. and Chu, F., Ultrasensitive microarray-based analytical methods: principles and practice, Presentation, International Business Communication, Protein Microarray Technology, San Diego, CA, 2001.

Ekins, R.P., Chu, F., and Biggart, E., Fluorescence spectroscopy and its application to a new generation of high sensitivity, multi-microspot, multianalyte, immunoassay, *Clin. Chim. Acta*, 194, 91–114, 1990.

Ellington, A.D. and Szostak, J.W., *In vitro* selection of RNA molecules that bind specific ligands, *Nature*, 346, 818–822, 1990.

Ezan, E. and Grassi, J., Optimization, in *Immunoassays: A Practical Approach*, Gosling, J.P., Ed., Oxford University Press, Oxford, 2000, chap. 7.

Fung, E.T., Thulasiraman, V., Weinberger, S.R., and Dalmasso, E.A., Protein biochips for differential profiling, *Curr. Opinion Biotechnol.*, 12, 65–69, 2001.

Ge, H., NAR UPA, a universal protein array system for quantitative detection of protein–protein, protein–DNA, protein–RNA and protein–ligand interactions, *Nucleic Acid Res.*, 28(2), i–vii, 2000.

Golden, M.C., Collins, B.D., Willis, M.C., and Koch, T.H., Diagnostic potential of photoSELEX-evolved ssDNA aptamers, *J. Biotechnol.*, 81, 167–178, 2000.

Haab, B.B., Dunham, M.J., and Brown, P.O., Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions, *Genome Biol.*, 2, 0004.1–0004.13, 2001.

Hanash, S., Disease proteomics, *Nature*, 422, 226–232, 2003.

He, M. and Taussig, M.J., DiscernArray technology: a cell-free method for the generation of protein arrays from PCR DNA, *J. Immunol. Meth.*, 274, 265–270, 2003.

He, M. and Taussig, M.J., Single step generation of protein arrays from DNA by cell-free expression and *in situ* immobilization (PISA method), *Nucleic Acid Res.*, 29(15), 1–6, 2001.

Huang, R.P., Detection of multiple proteins in an antibody-based protein microarray system, *J. Immunol. Meth.*, 255, 1–13, 2001a.

Huang, R.P., Huang, R., Fan, Y., and Lin, Y., Simultaneous detection of multiple cytokines from conditioned media and patient's sera by an antibody-based protein array system, *Anal. Biochem.*, 294, 55–56, 2001b.

Huels, C., Muellner, S., Meyer, H.E., and Cahill, D.J., The impact of protein biochips and microarrays on the drug development process, *Drug Dev. Technol.*, 7, S119–S124, 2002.

Humphrey-Smith, I., Replication-induced protein synthesis and its importance to proteomics, *Electrophoresis*, 20, 653–659, 1999.

Jayasena, S.D., Aptamers: an emerging class of molecules that rival antibodies in diagnostics, *Clin. Chem.*, 45, 1628–1650, 1999.

Joos, T.O., Schrenk, M., Hopfl, P., Kroger, K., Chowdhury, U., Stoll, D., Schorner, D., Durr, M., Herick, K., Rupp, R., Shon, K., and Hammerle, H., A microarray enzyme-linked immunosorbent assay for autoimmune diagnostics, *Electrophoresis*, 21, 2641–2650, 2000.

Klose, J. and Kobalz, U., Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome, *Electrophoresis*, 16, 1034–1059, 1995.

Kodadek, T., Protein microarrays: prospects and problems, *Chem. Biol.*, 8, 105–115, 2001.

LaBrie, S., Protein analysis in an array format using a novel, highly sensitive detection system, Presentation, International Business Communication, Drug Discovery Technologies, Session F1, Boston, MA, 2001.

Lizardi, P.M., Huang, H., Zhu, Z., Bray-Ward, P., Thomas, D.C., and Ward, D.C., Mutation detection and single-molecule counting using isothermal rolling circle amplification, *Nat. Genetics*, 19, 225–232, 1998.

Lueking, A., Horn, M., Eickhoff, H., Bussow, K., Lehrach, H., and Walter, G., Protein microarrays for gene expression and antibody screening, *Anal. Biochem.*, 270, 103–111, 1999.

MacBeath, G. and Schreiber, S.L., Printing proteins as microarrays for high-throughput function determination, *Science*, 289, 1760–1763, 2000.

Matson, R.S. and Little, M.C., Strategy for the immobilization of monoclonal antibodies on solid-phase supports, *J. Chromatogr.*, 458, 67–77, 1988.

Matson, R.S., Milton, R.C., Cress, M.C., and Rampal, J.B., Microarray-based cytokine immunosorbent assay, Oak Ridge Conference, Seattle, WA, 2001, Poster 20.

McDonald, W.H. and Yates, J.R., Proteomic tools for cell biology, *Traffic*, 1, 747–754, 2000.

Melnyk, O., Duburcq, X., Olivier, C., Urbes, F., Auriault, C., and Gras-Masse, H., Peptide arrays for highly sensitive and specific antibody-binding fluorescence assays, *Bioconjugate Chem.*, 13, 713–720, 2002.

Mendoza, L.G., McQuary, P., Mongan, A., Gangadharan, R., Brignac, S., and Eggers, M., High-throughput microarray-based enzyme-linked immunosorbent assay (ELISA), *Biotechiques*, 24, 778–788, 1999.

Mezzasoma, L., Bacarese-Hamilton, T., DiCristina, M., Rossi, R., Bistoni, R., and Crisanti, A., Antigen microarrays for serodiagnosis of infectious diseases, *Clin. Chem.*, 48, 121–130, 2002.

Moody, M.D., Van Arsdell, S.W., Murphy, K.P., Orencole, S.F., and Burns, C., Array-based ELISAs for high-throughput analysis of human cytokines, *Biotechniques*, 31, 1–7, 2001.

Nock, S. and Wilson, D.S., Recent developments in protein microarray technology, *Angew. Chem. Int. Ed.*, 42, 494–500, 2003.

Pawlak, M., Schick, E., Bopp, M.A., Schneider, M.J., Oroszlan, P., and Ehrat, M., Zeptosen's protein microarrays: a novel high performance microarray platform for low abundance protein analysis, *Proteomics*, 2, 383–393, 2002.

Peluso, P., Wilson, D.S., Do, D., Tran, H., Venkatasubbaiah, M., Quincy, D., Heidecker, B., Poindexter, K., Tolani, N., Phelan, M., Witte, K., Jung, L.S., Wagner, P., and Nock, S., Optimizing antibody immobilization strategies for the construction of protein microarrays, *Anal. Biochem.*, 312, 113–124, 2003.

Phizicky, E., Bastiaens, P.I.H., Zhu, H., Snyder, M., and Fields, S., Protein analysis on a proteomic scale, *Nature*, 422, 208–215, 2003.

Poon, T.C.W., Yip, T.T., Chan, A.T.C., Yip, C., Yip, V., Mok, T.S.D., Lee, C.C.Y., Leung, T.W.T., Ho, S.K.W., and Johnson, P.J., Comprehensive proteomic profiling identifies serum proteomic signatures for detection of hepatocellular carcinoma and its subtypes, *Clin. Chem.*, 49, 752–760, 2003.

Ruiz-Taylor, L.A., Martin, T.L., Zaugg, F.G., Indermuhle, P., Nock, S., and Wagner, P., Monolayers of derivatized poly(L-lysine)-grafted poly(ethylene glycol) on metal oxides as a class of biomolecular interfaces, *Proc. Natl. Acad. Sci. USA*, 98, 852–857, 2001.

Sapsford, K.E., Liron, Z., Shubin, Y.S., and Ligler, F.S., Kinetics of antigen binding to arrays of antibodies in different sized spots, *Anal. Chem.*, 73, 5518–5524, 2001.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O., Quantitative monitoring of gene expression patterns with a complementary DNA microarray, *Science*, 270, 467–470, 1995.

Schweitzer, B., Roberts, S., Grimwade, B., Shao, W., Wang, M., Fu, Q., Shu, Q., Laroche, I., Zhou, Z., Tchernev, V.T., Christiansen, J., Velleca, M., and Kingsmore, S.F., Multiplexed protein profiling on microarrays by rolling circle amplification, *Nature Biotechnol.*, 20, 359–365, 2002.

Schweitzer, B., Wiltshire, S., Lamber, J., O'Malley, S., Kukanskis, K., Zhu, Z., Kingsmore, S.F., Lizardi, P.M., and Ward, D.C., Immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection, *Proc. Natl. Acad. Sci. USA*, 97, 10113–10119, 2000.

Seong, S.Y., Microimmunoassay using a protein chip: optimizing conditions for protein immobilization, *Clin. Diagn. Lab. Immunol.*, 9, 927–930, 2002.

Silzel, J.W., Cerecek, B., Dodson, C., Tsong, T., and Obremski, R.J., Mass-sensing, multianalyte microarray immunoassay with imaging detection, *Clin. Chem.*, 44, 2036–2043, 1998.

Thulasiraman, V., McCutchen-Maloney, S.L., Motin, V.L., and Garcia, E., Detection and identification of virulence factors in *Yersinia pestis* using SELDI ProteinChip system, *Biotechiques*, 30, 428–432, 2001.

Tuerk, C. and Gold, L., Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, *Science*, 249, 505–510, 1990.

Tyers, M. and Mann, M., From genomics to proteomics, *Nature*, 422, 193–197, 2003.

Valle, R.P.C. and Jendoubi, M., Antibody-based technologies for target discovery, *Curr. Opin. Drug Disc. Dev.*, 6, 197–203, 2003.

Wagner, P., Protein biochips as powerful new tools in proteomics, Presentation, International Business Communication, Protein Microarrays Conference, San Diego, CA, 2002.

Wiese R., Belosludtsev, Y., Powdrill, T., Thompson, P., and Hogan, M., Simultaneous multianalyte ELISA performed on a microarray platform, *Clin. Chem.*, 47, 1451–1457, 2001.

Wellmann, A., Wollscheid, V., Lu, H., Ma, Z.L., Albers, P., Schutze, K., Rohde, V., Behrens, P., Dreschers, S., Ko, Y., and Wernert, N., Analysis of microdissected prostate tissue with ProteinChip arrays: a way to new insights into carcinogenesis and to diagnostic tools, *Int. J. Mol. Med.*, 9, 341–347, 2002.

Weng, S., Gu, K., Hammond, P.W., Lohse, P., Rise, C., Wagner, R.W., Wright, M.C., and Kuimelis, R.G., Generating addressable protein microarrays with PROfusion covalent mRNA-protein fusion technology, *Proteomics*, 2, 48–57, 2002.

Willis, M.C., LeCuyer, K.A., Meisenheimer, K.M., Uhlenbeck, O.C., and Koch, T.H., An RNA–protein contact determination by 5-bromouridine substitution, photocrosslinking and sequencing, *Nucleic Acid Res.*, 22(23), 4947–4952, 1994.

Woodbury, R.L., Varnum, S.M., and Zanger, R.C., Elevated HGF levels in serum from breast cancer patients detected using a protein microarray ELISA, *J. Proteome Res.*, 1, 233–237, 2002.

Yguerabide, J. and Yguerabide, E.E., Light-scattering submicroscopic particles as highly fluorescent analogs and their use as tracer labels in clinical and biological applications, *Anal. Biochem.*, 262, 137–176, 1998.

Yguerabide, J. and Yguerabide, E.E., Resonance light scattering particles as ultrasensitive labels for detection of analytes in a wide range of applications, *J. Cell. Biochem.*, 37 (Suppl.), 71–81, 2001.

Zhu, H. and Snyder, M., Protein chip technology, *Curr. Opin. Chem. Biol.*, 7, 55–63, 2003.

Zhu, H., Klemic, J.F., Chang, S., Bertone, P., Casamayor, A., Klemic, K.G., Smith, D., Gerstein, M., Reed, M.A., and Snyder, M., Analysis of yeast protein kinases using protein chips, *Nature Genetics*, 26, 283–289, 2000.

Zichi, D., Koga, T., Greef, C., Ostroff, R., and Petach, H., Photoaptamer technology: development of multiplexed microarray protein assays, *Clin. Chem.*, 48, 1865–1868, 2002.

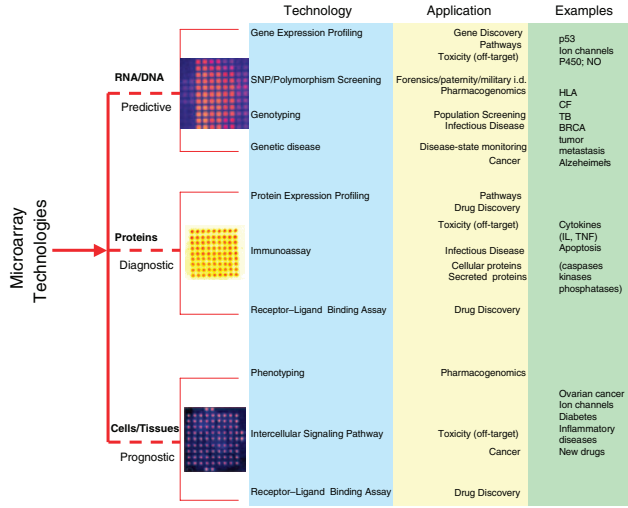


Figure 1.18 Applications of microarray technologies.

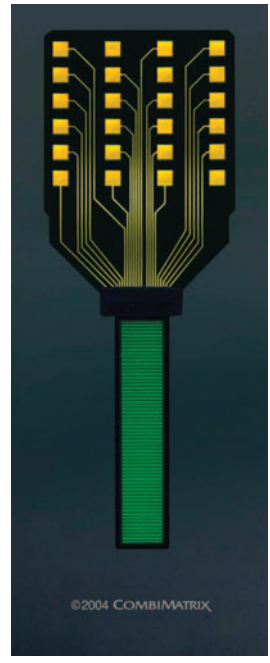


Figure 2.7 CustomArray™. (Photo courtesy of CombiMatrix Corporation, Mukilteo, WA.)

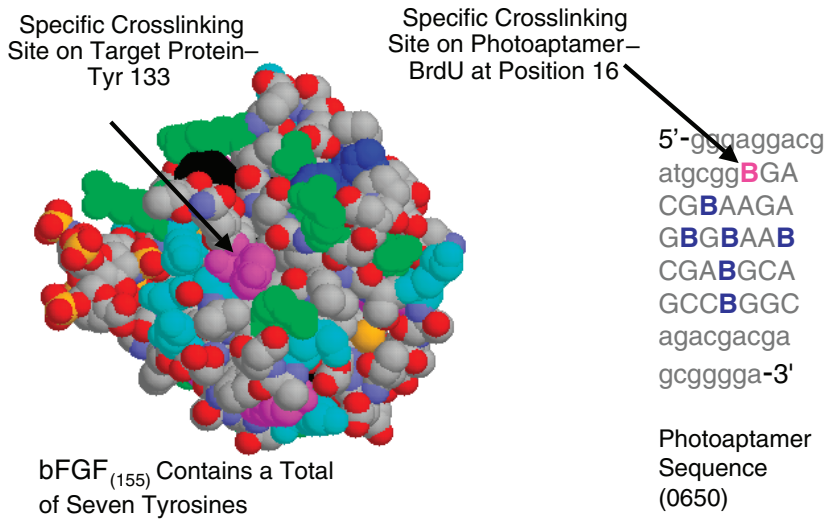


Figure 6.29 Photoaptamer crosslink. (Copyright SomaLogic, Inc., Boulder, CO. Reproduced with permission only.)

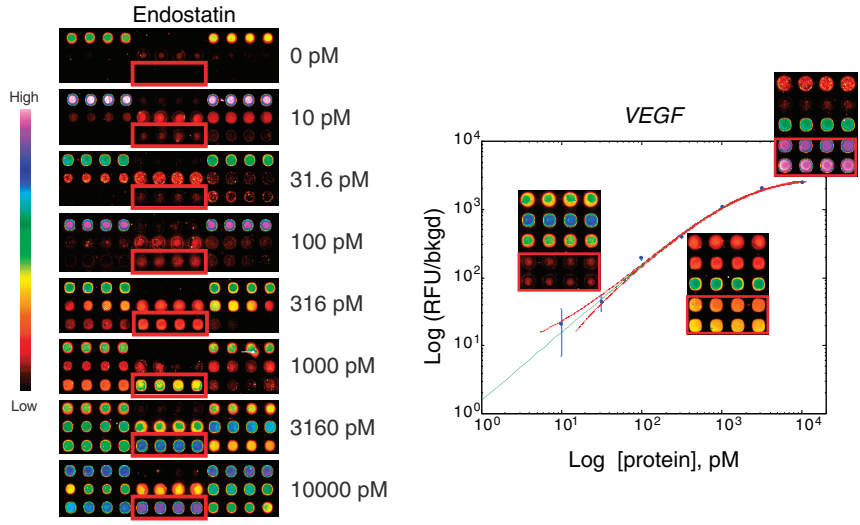


Figure 6.30 Aptamer dose response. (Copyright SomaLogic, Inc., Boulder, CO. Reproduced with permission only.)