

A perspective on protein microarrays

Proteins, not genes, are the true targets of medicines, but their analysis by array technology still poses significant challenges for drug developers.

Peter Mitchell

Despite great advancements in genomics, researchers remain frustratingly far, in most instances, from implicating specific proteins in disease. Researchers have long known that the concentration of an mRNA within a cell is poorly correlated with the actual abundance of that protein. The only solution is to measure the amount of the specific protein directly. However, when post-translational modifications are taken into account, human cells are a mixing pot of a million or more different protein variants, any of which could be altered in disease. The size of the problem is therefore enormous.

An attractive solution has arrived with the concept of the protein microarray, or protein chip. Using technologies employed in the manufacture of DNA microarrays, a glass slide (the "chip") could be printed with thousands of protein detectors, a biological sample spread over the chip, and any binding noted. The protein of interest could then be analyzed using standard techniques such as time-of-flight mass spectrometry and peptide mass fingerprinting. Such biochips promise a speedy and high-throughput means to profile disease-related proteins or to study protein-protein and protein-drug interactions, previously only achieved using painstaking and tedious standard methods such as western blotting and enzyme-linked immunosorbent assay (ELISA).

With few exceptions, however, the protein chip remains a drug discoverer's fantasy. Despite considerable investment by several companies, protein chips have yet to flood the market. Manufacturers have found that proteins are difficult beasts to handle, and the perceived market demands have not matched those initially anticipated.

The attraction of arrays

The runaway success of DNA chips in the past five years, with a current market of around \$300 million, is undoubtedly responsible for much of the enthusiasm within the protein chip arena.

First, the necessary tools and know-how were already in place: DNA arrayers and

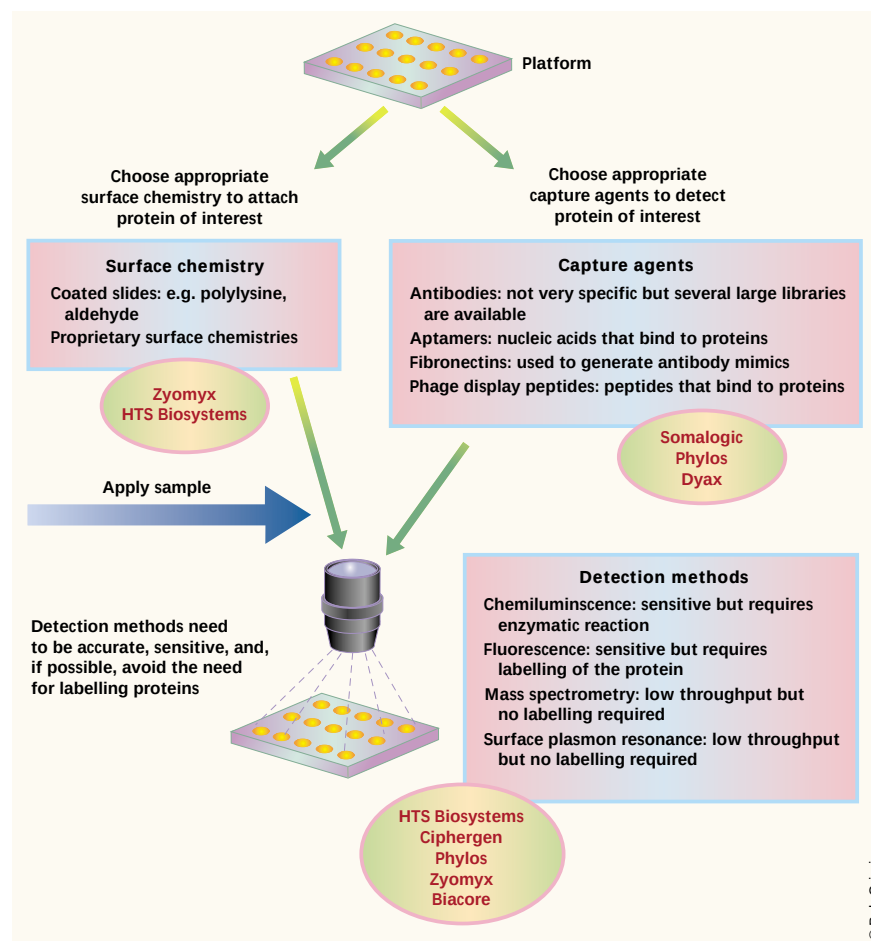


Figure 1. The schematic shows the different stages and components involved in the construction of protein chips. Companies focused on addressing some of these issues are highlighted in the green ovals. Note that many options are still in an experimental stage.

contact printers, for instance, could readily be adapted to create arrays of proteins. In theory, protein chips offer customers economies similar to those of DNA microarrays, consuming only tiny quantities of precious and expensive samples and reagents, stocks that might have taken researchers years to prepare. Moreover, from a commercial perspective, protein chips appeared to be a highly lucrative market: In December 2000, market research company BioInsights (Redwood City, CA) predicted that the worldwide market for protein microarrays would grow from \$45 million in 2000 to almost \$500 million in 2006. Julian Aberly,

head of Biacore's (Uppsala, Sweden) pharmaceutical and biotech business unit, is even more optimistic, predicting sales of at least \$2 billion within the next five years.

However, the analogy with DNA arrays is more apparent than real when it comes to proteins. It has proven far harder to make a marketable protein chip than anyone had expected. Constructing a protein array requires many more steps and more complex science than that used for DNA microarrays.

Among many challenges, to build a viable protein chip a manufacturer needs to:

- choose a surface chemistry that will allow

Peter Mitchell is a London-based freelance writer who specializes in pharmaceuticals and medical technology (pete@dmed.demon.co.uk).

immobilized proteins of diverse types to retain their secondary and tertiary structure, and thus their biological activity;

- identify and isolate an agent (e.g., antibody) with which to capture the protein(s) of interest;
- devise a means of measuring the degree of protein binding, assuring both sensitivity and a suitable range of operation;
- in some experiments, extract the detected protein from the chip and analyze it further.

A plethora of different technologies have been promoted, by dozens of companies, to address the needs highlighted in each of these steps (Fig. 1 and Table).

Contact sports

The first challenge is to fix or “print” the protein, or protein ligand, onto the chip in a biologically active form.

Most arrays are created on glass or silicon slides—standard in the production of DNA microarrays—treated with an aldehyde or other agent to immobilize the protein¹.

Other materials being investigated as immobilizing coatings include layers of aluminum or gold, hydrophilic polymers, and polyacrylamide gels. Each material demands its own chemistries (which must avoid denaturing the protein molecules) ideally to orient each molecule in the same direction while at the same time creating a hydrophilic environment in which reactions can take place.

Another alternative is being pursued by Zyomyx (Hayward, CA), which is using photolithography to etch miniature wells on the surface of silicon chips. The immobilized proteins or antibodies are located in the flow chambers on the chip, so that they are always kept in aqueous solution and do not denature through drying-out. The proteins will then be detected by fluorescence labeling, using a rapid-scanning reader developed by

the company. Using the photolithography technique, the company can manufacture high-density arrays capable of detecting up to 10,000 proteins in parallel, and will be suitable for target discovery and validation.

For the moment, a rather different approach to chip production has been taken by Large Scale Biology (LSB; Vacaville, CA). Hundreds of thin plastic rods, each doped with a particular antibody, are bundled together in a sheaf. The sheaf is then cut transversely into micrometer-thin slices, yielding chips containing antibody-loaded plastic circles (the rods’ cross-sections). A meter-long sheaf could generate a million low-density chips, each with ten to a few hundred elements—potentially a high-volume, low-cost product, says Leigh Anderson, LSB’s chief scientific officer (CSO). These chips may never be marketed, however: LSB is developing them for in-house assays and clinical diagnostics only.

Packard (Meriden, CT), also a contender in the protein chip arena, hopes to build on its continuing success in the DNA microarray market. Packard believes its hydrogel chip substrates will be suitable for protein microarrays, and is working with Oxford GlycoSciences (Oxford, United Kingdom) to obtain new capture proteins. Proteins are imprinted on a porous polyacrylamide gel, similar to that used in electrophoresis, and immobilized using a coupling reagent that forms a covalent bond with amine groups on the protein molecules. The gel is aqueous, so immobilized proteins can undergo binding reactions in solution.

Capture agents

Whereas DNA chips use the natural power of hybridization (between complementary nucleotide strands), protein microarrays commonly harness Nature’s own protein-

capturing agents—antibodies. Slides are printed with an array of different antibodies, which can be monoclonal, polyclonal, antibody fragments, or synthetic polypeptide ligands.

However, huge numbers of antibodies will be needed to populate these chips, and several protein chip companies have been forced to do deals with antibody providers to ensure an adequate supply. For example, Zyomyx has licensed phage-display libraries, which can display hundreds of millions of antibody fragments and proteins, from both Dyax (Cambridge, MA) and Cambridge Antibody Technology (CAT; Cambridge, United Kingdom).

Antibodies are not the only capture agents available: aptamers, oligonucleotides that bind specifically to proteins, can also be used. SomaLogic (Boulder, CO) is constructing arrays using light-sensitive “photoaptamers.” The aptamers capture proteins in the sample solution, and the chips are then exposed to ultraviolet light, covalently cross-linking the aptamers with their target proteins. The chips are washed vigorously to remove all unbound protein, and captured proteins are revealed using a universal protein stain.

Aptamer arrays have an advantage over antibody arrays because they can be printed using the same technologies used to create mRNA expression arrays. Manufacturers also claim that they are more specific and sensitive, because the aggressive washing removes much of the unwanted background (aptamer) binding. SomaLogic hopes eventually to put tens of thousands of aptamers onto a chip, although its progress toward this goal was limited by a shortage of proteins to raise aptamer ligands against. For this reason, the company struck a deal with Celera (Norwalk, CT) in which it will give Celera first use of its products in return for proteins and a significant amount of cash backing.

Detection is key

The third most pressing need in protein microarrays is detection. How can minute quantities of bound proteins on the chips be detected, bearing in mind that there is no equivalent of PCR for the amplification of proteins? The most obvious method is to use an ELISA-based assay in which the probes are tagged with enzymatic or fluorescent labels. An alternative is to apply a universal protein stain, although this is not an option if the capture molecules are themselves proteins.

There are disadvantages, however, to many of these methods. For example, ELISAs suffer from the nonspecificity of protein-antibody interactions, which cre-

Table: Selected companies working in the protein microarray business

Company (location)	Area of development
Affibody (Stockholm, Sweden)	Antibody-like ligands
Akceli (Cambridge, MA)	Live-cell microarrays
Biacore (Uppsala, Sweden)	Surface plasmon resonance for arrays
Cambridge Antibody Technology (Cambridge, United Kingdom)	Antibodies (phage-display libraries)
Ciphergen Biosystems (Fremont, CA)	Protein-capture chips for mass spectrometry analysis
Dyax (Cambridge, MA)	Antibodies (phage-display library)
HTS Biosystems (Hopkinton, MA)	Surface plasmon resonance arrays
Large Scale Biology (Vacaville, CA)	Plastic-based antibody arrays
Lumicyte (Fremont, CA)	Protein-capture chips for mass spectrometry analysis
Oxford GlycoSciences (Oxford, United Kingdom)	Protein discovery
Packard Bioscience (Meriden, CT)	Gel-based antibody arrays
Phylos (Lexington, MA)	Non-antibody capture agents
Prolinx (Bothell, WA)	Activated-glass array substrates
Protein Sciences (Meriden, CT)	Protein expression
SomaLogic (Boulder, CO)	Aptamer arrays
Zyomyx (Hayward, CA)	Silicon-based antibody arrays

ates “false positives”. Labeling proteins with, say, a fluorophore, reduces the quantitative accuracy of assays, because the label can change the way the molecule binds to other molecules. For these reasons, several protein chip companies are using label-free detection systems such as CIPHERGEN’s protein chip, which uses mass spectrometric analysis (see below).

Another unlabeled detection method—surface plasmon resonance (SPR)—is being developed for microarray application by Biacore, HTS Biosystems (Hopkinton, MA), and possibly others.

SPR uses a gold-coated glass chip on which test proteins or antibodies are immobilized. Unlabeled probes—antibodies, proteins from cell lysates, or drug candidates—are applied to the surface, while the chip is scanned from below by a light beam. The beam is reflected back by the gold layer, and the angle of reflection varies according to the mass of the molecules attached. Measuring this angle shows where ligands have bound to the immobilized test proteins. Moreover, it is possible to monitor the kinetics of the reaction in real time, providing a measure of the affinity and speed of binding between the capture agent and probe.

Biacore has marketed SPR-based protein chips for several years, albeit in a single-spot format. In June 2000, it began working with Millennium Pharmaceuticals (Cambridge, MA) to co-develop an array version. “We have prototypes working in labs that measure tens, hundreds, or even thousands of spots at a time,” said Biacore’s Abery. A commercial product is still at least two years from the market.

HTS Biosystems has products closer to market, having already shipped several beta versions of its system to Dyax and Mitsubishi Chemical, and a full commercial launch is planned for April 2002. The initial chips will contain 400 protein spots, using antibody assays supplied by Dyax, and future chips could be expanded to thousands of spots, possibly using aptamers and antibodies.

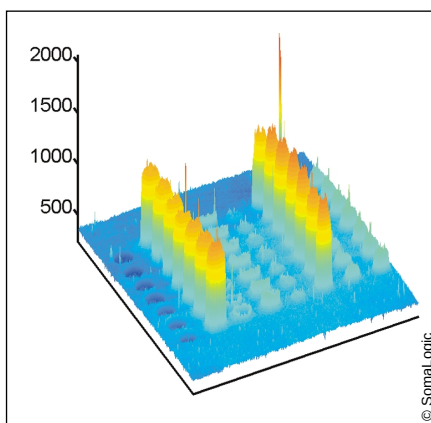
Chips to go

One of the few companies selling commercial protein microarrays is CIPHERGEN (Fremont, CA). However, its ProteinChips, which are marketed for expression profiling, are unlike most of the others in development. Bill Rich, CIPHERGEN’s chief executive officer (CEO), describes them as “small chromatographic surfaces”, carrying either 8 or 16 different spots of capture agents per array. When a biological sample is applied to the chip, the surface chemistry in each spot captures proteins with com-

mon properties, such as their affinity for a given ligand.

These spots are then analyzed directly using a laser that evaporates the spot into a benchtop time-of-flight mass spectrometer, a standard protein analysis technique. “By comparing the output from a variant sample and a normal sample, you can see which proteins are differentially expressed,” says Rich. The limit of the approach is the resolving power of the mass spectroscopy: “About 1,000 proteins is about the best we can separate in a reasonable amount of time,” he says.

CIPHERGEN has sold over 200 of its systems, generating about a half of the company’s 2001 revenues of \$19 million—a figure that is set to double in 2002, say analysts. The chips themselves cost \$50–\$75.



SomaLogic is using its photoaptamers to detect proteins on its chips, providing a simultaneous quantitative measurements of multiple analytes.

In December, CIPHERGEN introduced a device to interface its chips directly with the tandem mass spectroscopy instruments that are currently used in large protein analysis facilities, allowing for a much more detailed structural analysis of the discovered protein. According to Rich, Human Genome Sciences (Rockville, MD) successfully used this system to examine the structure of the receptor for its protein repifermin, a growth factor now in clinical trials as a wound treatment. “Everybody wants to do this [structure determination] at the moment in drug discovery,” he says. “[The chip and detector] is a massive value proposition where only micrograms of the protein are available.”

High expectations

Despite these few successes, Steven Bodovitz, a technology analyst at BioInsights, notes how little progress has been made since the company last surveyed the sector in December 2000. “Several companies that announced they were going to

launch products in 2001—SomaLogic, Phylos, Zyomyx, and HTS Biosystems, among others—but those products did not reach the market.”

SomaLogic originally expected to produce 1,000-aptamer arrays by the end of 2001, but has had to delay this. “High-throughput identification of photoaptamers ... took more time than we had hoped,” says Larry Gold, the company’s CEO. Meanwhile, Zyomyx has not disclosed either the reasons for the delays or what its new timetable might be. Phylos (Lexington, MA) has now dropped plans to commercialize protein microarrays. According to Al Collinson, Phylos’ vice president of business development, the decision was made because of the “uncertain dynamics” of the market for protein microarrays. Instead, it will supply its Trinectin ligand technology—based on high-affinity polypeptide “scaffold molecules” such as fibronectin—to a partner who will use the ligands to populate commercial protein arrays.

The bottlenecks ...

However, perhaps the heart of the problem is that it is fundamentally far harder to work with proteins than with nucleic acids. Bodovitz adds: “Obviously some of the initial enthusiasm was a little ahead of what could actually be delivered.” Indeed, there are still many outstanding obstacles for the industry to overcome, although there appears to be some disagreement as to exactly what those are.

Capture agents. According to Snyder, and many others, the biggest problem for the protein chip market is the shortage of sufficiently specific antibodies or other ligands. “All the antibodies at the moment are in the refrigerators of famous professors,” says CIPHERGEN’s Bill Rich. “They may have spent 20 years just getting enough volume of five antibodies.” Biacore’s Abery agrees: “Getting even a 100 proteins in a format pure enough to put on an array is a significant investment.”

Moreover, most potential buyers of protein chips do not want off-the-shelf arrays—they want to put their own precious antibodies or target proteins on a chip surface. Array makers must therefore work closely with pharmaceutical partners to build access to useful proteins and ligands.

Currently, off-the-shelf chips have just a few capture agents, typically aimed at a well-defined class of proteins (e.g., cytokines). “For a target-validation assay, you don’t need too many proteins on a chip,” says Bodovitz. “If you are looking at cancer genes, cytokines, or signal-transduction proteins, you can just study a smaller subset of them at much lower density. The anti-

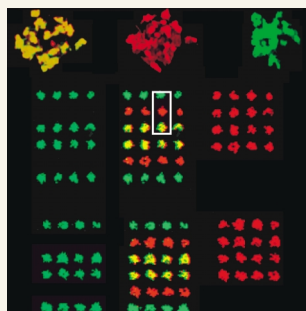
Microarrays go live

More exotic variants on the protein chip theme include the "living protein chip". David Sabatini, of the Whitehead Institute for Biomedical Research (Cambridge, MA), affiliated with the Massachusetts Institute of Technology (MIT), spotted cDNA onto glass slides and cultured human embryonic kidney (HEK) cells on the surface. The HEK cells take up the cDNA, expressing the corresponding proteins, and can then be used for functional studies of, for example, the interactions of the protein with small molecules (ref. 4, and see *Nat. Biotechnol.* **19**, 535, 2001).

According to Sabatini, living arrays have the advantage that the proteins do not need to be purified, because they are expressed at concentrations so high that they swamp any signal from the cell's endogenous proteins. Also, cells can express any protein, even membrane proteins like G-protein-coupled receptors (GPCRs), which are otherwise difficult to solubilize.

Under a license granted by MIT, the technology will be commercialized by a start-up company called Akceli (Cambridge, MA). Akceli's arrays, which could be launched next year, will be populated with thousands of cDNAs—a significant investment, as full-length cDNA constructs are time-consuming to make and expensive to buy.

However, Sabatini is more interested in using living protein chips for "expression monitoring", examining the effects of expressed proteins on the cells themselves (e.g., to identify proteins that induce apoptosis). Here the arrays, bearing 10,000–15,000 cDNAs on a single slide, provide much faster throughput than a traditional microtiter plate, and allow for direct examination of the cells, for example under a fluorescence microscope. *PM*



bodies for these tend to already exist, and you don't have to worry so much about cross-reactivity."

Because antibodies are generally glycosylated, and have large surface areas for interactions, they show significant cross-reactivity between target proteins. This leads to large numbers of false positives and negatives. So, many companies are trying to develop ligands with better specificity, such as aptamers or Phyllos's non-antibody-based peptide scaffolds.

Additionally, many proteins appear in complexes, and it is therefore necessary to find a ligand that can bind to the protein in this new form. "There are several diseases that are related to the composition of a complex variant protein," says Stefan Stahl, CSO of the ligand company Affibody (Uppsala, Sweden). "This must be very difficult to understand with protein chips."

Others, however, such as Kevin Johnson of antibody maker CAT, disagree: "We have got antibodies working on all the platforms we have tried, so there doesn't yet seem to be a technological limitation that says antibodies are not going to be a good idea as content for microarray platforms."

The nature of proteins. Other problems may result in attempts to "stick" proteins on microarray surfaces: proteins often denature at solid-liquid and liquid-air interfaces, and a small protein spot has a disproportionately large surface area, further

encouraging denaturation. Furthermore, the surface chemistry *per se* may denature proteins or result in loss of activity. The problem is confounded as the number of different protein spots on the chip increases, because they may not all behave alike when exposed to the same surface chemistry. "You need the proteins in a functional format, and for 100 proteins there may be 100 different ways you need to immobilize them," says Aberly.

Optimizing these chemistries demands know-how in many fields, according to Zyomyx, one of the most advanced of the protein chip companies: the solution could include organic thin films, self-assembled monolayers, nanotechnology, biomaterials, and surface biochemistry. Because surface chemistries are key to products reaching the market early, companies are understandably keen to keep trade secrets, and there is little detail available on the exact nature of these products.

Solubility is also a serious difficulty. "Many proteins that are soluble in their native environments might not be soluble when you want to measure them on the chip," says Stahl. Should the chemical environment change, soluble proteins may precipitate onto the chip, wrecking the assay.

Moreover, G-protein-coupled receptors (GPCRs), which make up a large percentage of all targets for current therapeutics, are particularly a problem, because purifying

these proteins with detergents can destroy their function. "We need improved technology to get them expressed and assayed in a functional state," says Snyder. "Some people have succeeded in doing this, but not in a high-throughput fashion for purifying hundreds of proteins." (See "Microarrays go live".)

David Sabatini, of the Whitehead Institute for Biomedical Research (Cambridge, MA), raises a further obstacle. "Proteins are notoriously unstable," he warns. Unlike DNA microarrays, which have a shelf life of many months, protein arrays may not be stable in the long term, as the proteins lose their native conformation. "The only point in having an array is to do many experiments quickly. If they are so unstable that you need to print more every couple of days, then that limits their utility." The current literature on the issue is inconclusive. "I expect the answer is that we don't know," says Sabatini. "A lot of the proteins on those big arrays have unknown functions, so you could never actually test whether they are still functional."

But CAT's Johnson adds that antibodies are more robust capture agents, and eventually it will be possible to dry and store them at room temperature.

Robust detection. One more difference between protein and DNA microarrays is that protein concentrations in a single biological sample vary far more than do mRNA levels. So protein-chip detector systems must have very large ranges of operation—up to a factor of 10^{14} , compared with 10^4 for mRNA. For example, an antibody with nanomolar affinity for a specific protein would be swamped by target protein present in micromolar concentration. On the other hand, very rare proteins present at picomolar or femtomolar concentrations would not be detected.

"This means you can choose to detect low-abundance proteins or high-abundance proteins, but to fit them all into one chip might be difficult," says Affibody's Stahl. As a result, some chip companies are considering having separate array product lines—some to detect very rare proteins, and others to detect more abundant types.

But according to CAT's Johnson, this "dynamic range" problem is not insuperable. The solution is to estimate the concentration of the target protein in the sample, and then choose an antibody with an affinity that can match that concentration. If necessary, two or three antibodies to the same target, but with different affinities, can be used in different dots on the chip to cover the whole concentration range.

However, even choosing multiple antibodies may not overcome the problems

inherent in detecting proteins modified after translation, for example by glycosylation or phosphorylation. A ligand designed to bind to the unmodified protein will not necessarily bind to the ten or twenty different translationally modified variants. "Often the assay produced from the open reading frame protein just does not work with the protein as it appears in real life," says CIPHERgen's Rich. "This means that there is a strong probability of a large number of failed ELISA assays."

There are other, perhaps even more serious, criticisms of the currently popular ELISA method of protein detection. In ELISA, the primary antibody is fixed to the chip and captures the target protein from the sample. This complex is then detected either via a label that was attached to the target protein during expression, or by introducing a labeled secondary antibody that will bind to the complex. It turns out to be very difficult to make this method robust in microarrays, where many different ELISAs have to work on the same chip under the same conditions.

"To get quantitative expression profiling from this, you need a way of uniformly labeling the proteins," says Affibody's Stahl.

"But this is a major challenge because they have such diverse properties. You would more likely label one protein to 100% and another to only 10%." As a result, he says, a lot of effort is being put into finding smart non-labeled detection systems, such as mass spectrometry, SPR, and (Affibody's own solution) fluorescence resonant energy transfer (FRET; ref. 2).

Small steps forward

It is a formidable list of problems to solve, but at least some of the companies in the game are confident that they can find solutions.

Indeed, the protein chip field was encouraged by the publication of a milestone paper by researchers at Yale University (New Haven, CT) and North Carolina State University (Raleigh, NC) in 2001 (ref. 3). The paper proved that it was feasible to make an array large enough to investigate the function of almost the entire yeast proteome.

Thirteen thousand protein samples were spotted onto half the area of a standard microscope slide. Snyder says: "You can globally see all the proteins that bind a particular other protein or lipid or small

molecule, so we think they will be enormously powerful for drug discovery." They will also help to elucidate protein function and map out pathways of protein-protein interactions—work that is currently done by two-hybrid assays. But the microarrays, says Snyder, are between ten to a hundred times as fast as two-hybrid methods, because of the very small amounts of material needed.

The bottom line, Johnson insists, is that all the evidence to date says the technology will work. "The real challenge is how to make it commercially attractive," he adds. However, the million-dollar question is whether the products will reach the market this year, says BioInsights' Bodovitz. "Without products, there is not much of a market."

1. MacBeath, G. & Schreiber, S. Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760–1763 (2000).
2. Karlstrom, A. & Nygren, P. Dual labeling of a binding protein allows for specific fluorescence detection of native protein. *Anal. Biochem.* **295**, 22–30 (2001).
3. Heng, Z. *et al.* Global analysis of protein activities using proteome chips. *Science* **293**, 2101–2105 (2001).
4. Sabatini, D. & Ziauddin, J. Microarrays of cells expressing defined cDNA. *Nature* **411**, 107–110 (2001).