DRUG DISCOVERY TODAY TARGETS

Protein biochips: the calm before the storm

Steven Bodovitz, Thomas Joos and Jutta Bachmann

The growth of protein biochip technology is on a different trajectory than other drug discovery and development technologies, such as DNA sequencing and high-throughput screening, where output per experiment has grown exponentially. By contrast, experimentation with protein biochips immediately hit barriers in output because of the limited availability of content and the challenges of running biochemical experiments on the surface of a biochip. Nevertheless, the industry has been making significant progress recently by launching new platforms with focused content and new multiplexed biochemical assays. However, this success might only represent the calm before the storm. Over the long-term, protein biochips have the potential to change the drug discovery and development process at the molecular level. The output and throughput of protein biochips could enable researchers to change from the traditional model of one target-one drug to a new model of evaluating one or more potential drugs against a panel of relevant molecular targets from a complex disease state.

The rate of advance of drug discovery and development technologies in the past decade has been exponential. If we turn the clock back ten years, GenBank (http://www.ncbi.nlm.nih.gov/) only contained 217 million base pairs of sequence (compared with 42 billion base pairs today), DNA microarrays had not been commercialized, protein identification by mass spectrometry had just become a reality [1,2], and automation was just beginning to enable high-throughput screening. Now the whole human genome has been sequenced and can be studied on a single DNA microarray, thousands of proteins can be identified in a single mass spectrometry run, and entire compound libraries can be screened in a matter of days. When protein biochips (Box 1) [3] began to emerge as the most probable candidate for the next great class of drug discovery and development technology at the beginning of this century, the common perception within the research community was that the rate of advances in protein biochip technology would match that of DNA microarrays, mass spectrometry and/or gene sequencing. This, however, has not proven to be correct. The amount of data per experiment (defined as output) that could be obtained from the earliest protein biochips [4–7] is still the standard several years later, with significant but not exponential increases in output expected within the next few years.

The primary reason for the slower-than-expected progress of the protein biochip industry has been the limited availability of content [8]. This problem is defined by a lack of high-quality capture agents that can be immobilized on the surface of a biochip – defined as a capture protein biochip – to bind to proteins from complex mixtures for the purpose of detection and quantification. In contrast to DNA microarrays, in which cDNA capture sequences can be predicted by Watson–Crick base-pairing, protein

Steven Bodovitz
BioPerspectives,
2040 Hyde Street,
San Francisco, CA 94109,
USA
e-mail:
bodovitzs@bioperspectives.com
Thomas Joos
NMI Natural and Medical Sciences Institute,
Markwiesenstr. 55,
72770 Reutlingen,
Germany
Jutta Bachmann
3 Bachmann Consulting,
Nøkkefaret 12,
1450 Nesoddtangen,
Norway

1359-6446/04/$ – see front matter PII: S1359-6446(04)03373-8

BOX 1

What is a biochip?

A biochip is a collection of miniaturized test sites – also commonly referred to as a microarray – arranged on a solid substrate that permits many tests to be performed at the same time to achieve higher throughput and speed. The miniaturized test sites can be placed onto a 2-dimensional surface, such as a glass slide, or a 3-dimensional surface, such as a bead. In the case of the 2-dimensional surface, multiple spots are arrayed, each containing a different immobilized protein or capture agent (or duplicates for control purposes). In the case of 3-dimensional surfaces, each has a unique immobilized protein or capture agent (or duplicates for control purposes), and are pooled into groups for parallel analysis.

capture agents have significantly more complex interactions with their ligands. Protein binding involves charge, hydrogen bonds and/or weak hydrophobic forces. Moreover, binding often requires specific 3-dimensional conformations, post-translational modifications and/or co-factors. Because of this complexity, each capture agent must be optimized empirically, presenting a significant challenge for large-scale production. The limited availability of content precludes experiments analogous to the comprehensive coverage of the human genome by DNA microarrays, but not high-value experiments addressing specific biological questions. Furthermore, even in the cases where more than 100 high-quality antibodies are available for capturing 100 different proteins, only approximately 30 to 40 of these can be placed together on the same biochip before the cross-reactivity between sandwich pairs becomes overwhelming [9]. Although this problem can be obviated to a certain extent by using multiple biochips, the benefits of miniaturization are diminished.

The analogy between DNA and protein biochips has fared better when applied to interaction protein biochips, where a limited number of proteins or peptides are immobilized on a surface to examine specific protein interactions. In this case, content is not a significant limitation because peptides can be synthesized and proteins can be produced using recombinant technology. The potential limitation is the question surrounding the physiological relevance of the detected interactions; in contrast to capture protein biochips, the interactions between immobilized protein or peptide and soluble protein are supposed to represent biochemical events that take place in a cell or biological fluid *in vivo*. All of these interactions take place on the surface of a biochip under a single set of conditions, which goes against the traditional dogma among biochemists that each biochemical reaction is unique and affected by a wide range of variables, including buffer conditions, protein conformation and co-factors. Moreover, in a cellular environment, a biochemical reaction will only take place if the reactants are co-localized in the same subcellular compartment; interacting biomolecules that come together on the surface of a biochip may or may not come together *in vivo*. By contrast, Michael Snyder's research group at Yale (http://.yale.edu/synder/) and several companies, such as Invitrogen Corporation (http://www.invitrogen.com), Protagen AG (http://www.protagen.com), Pepscan Systems BV (http://www.pepscan.nl) and Jerini AG (http://www.jerini.com), have shown compelling data with interaction protein biochips, in which the relevance of interactions was at least partially validated through other experimental methods [10]. The physiological relevance of the data thus remains a question that will probably only be resolved from extensive testing by researchers. However, the value of generating high-quality data from a protein biochip is readily apparent. No other technique for studying protein–protein interactions generates as many data points as quickly. If the goal is to cast the net of possible interactions far and wide, interaction protein biochips offer unparalleled speed.

Near-term prospects

The content problem initially slowed down the emergence of capture protein biochips, but developers have forged ahead by using available content and focusing on specific research areas. According to our recent analysis, at least 49 companies offer commercial capture protein biochip products and/or services, including both planar and bead-based platforms; scores of additional platforms are also available through collaborative agreements or are still under development [10]. The analytes measured include: adipokines, angiogenesis factors, apolipoproteins, cardiovascular disease markers, cell signaling molecules, chemokines, coagulation proteins, common allergens, cytokines, drugs of abuse, endocrine hormones, growth factors, HLA antigens matrix metalloproteinases, and serum cancer biomarkers. Examples for biomarker analysis are shown in Figure 1. Application areas for capture biochip assays in research include: allergies, Alzheimer's disease, apoptosis, autoimmunity, cancer, cardiovascular disease, connective tissue disorders, immunoglobulin isotyping, inflammation, infectious disease, metabolic diseases, oncology and sepsis. In addition to research and development tools, capture protein biochips are also currently used as diagnostic assays for medical conditions, such as allergies, autoimmune diseases, drug abuse, infectious diseases, heart failure and other coronary problems. In addition to diagnosing disease, capture protein biochips have the potential to identify biological warfare agents. The National Institutes of Health (http://www.nih.gov/) have allocated in excess of $1 billion to fund research on bioterrorism agents and development of vaccines and diagnostic tests.

Regarding those areas of the biochip industry dealing with interaction protein arrays, the number of companies with products and/or services on the market, according to our recent analysis, is only nine. Applications for interaction protein biochips include: characterizing antibodies and mapping epitopes; determining substrate specificities for kinases, proteases and phosphatases;