The protein biochip content problem

The protein biochip industry has a content problem that is not only slowing its development as a whole but possibly the development of the entire proteomics industry.

By Dr Steven Bodovitz

In science, for a problem to have its own name, it needs to be well defined and to present a significant challenge. The protein biochip content problem meets both of these criteria. The problem is defined by a lack of high-quality capture agents that can be immobilised on the surface of a biochip to bind, detect and quantify protein levels from complex mixtures. Unlike DNA microarrays, in which complementary DNA capture sequences can be predicted by Watson-Crick base-pairing, protein capture agents have significantly more complex interactions with their ligands. Interactions involve charge, hydrogen bonds and/or weak hydrophobic forces. Moreover, interactions can depend on proper three-dimensional conformations, post-translational modifications and/or multiple co-factors. Because of this complexity, each capture agent must be optimised empirically, presenting a significant challenge for large-scale production.

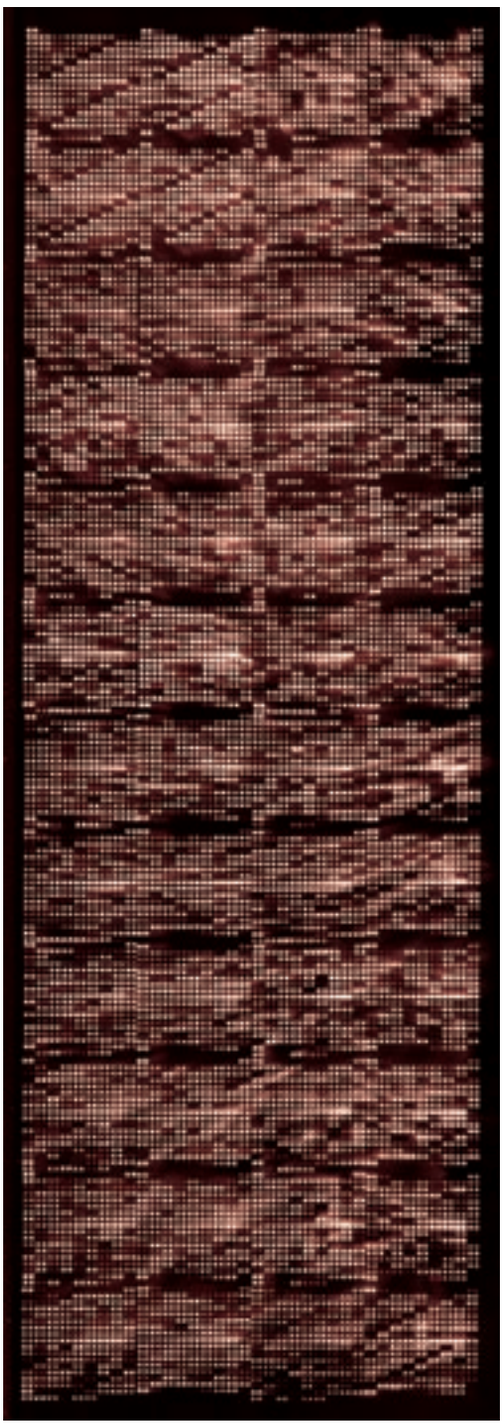
The content problem has slowed the development of the protein biochip industry and limited the breadth of product offerings. The majority of the market size in 2002, approximately $100 million[1], was due to detection-based technologies that do not rely on capture agents (see below). The predicted growth of the market to more than $400 million in 2007, however, depends on new product launches. To achieve this growth, protein biochip developers will need to employ one or more of four strategies for solving the content problem. The first and most straightforward is to improve the quality, speed and cost of capture agent production. The second is to make better use of limited content by focusing on pathways, druggable targets and surrogate markers of efficacy and toxicity. The third is to improve detection methods to compensate for shortcomings in affinity and selectivity. The fourth, and most complementary, is to generate content from synthetic peptides and recombinant proteins to study protein interactions.

The gold standard for capture agent affinity is the monoclonal antibody. It is the most common capture agent for protein biochips, but production is time-consuming and labour-intensive, which greatly limits the supply. Production requires immunisation of animals and generation of hybridoma cell lines. The most promising approach to faster production is recombinant methodology based on phage display and combinatorial libraries. The selection process can be directed against almost any target and can be completed within weeks. The initial affinities typically demonstrate K_D values in the nanomolar range, but maturation strategies can improve the affinity to K_D values in the picomolar range[2]. By selecting for recombinant antibodies that bind to targets under the same conditions, the antibodies can be designed to perform together on the surface of a biochip. Furthermore, by selecting antibodies with different affinities against the same target, the dynamic range on the biochip can be increased by several orders of magnitude.

Multiple companies are developing recombinant approaches for antibody generation. Cambridge

Antibody Technology Group plc (Cambridgeshire, UK) has created *in vitro* libraries containing more than 100 billion distinct antibodies. The company has four ongoing collaborations with protein biochip developers and is planning to set up a dedicated protein biochip programme. Affibody AB (Bromma, Sweden) has developed Affibody® affinity molecules that are based on a proprietary scaffold. Affibody molecules can be double-labelled with a donor and an acceptor fluorophore, enabling fluorescence resonance energy transfer and quantitative detection of unlabelled proteins. BioInvent International AB (Lund, Sweden) has a proprietary library based on scFv antibody fragments. These variable regions of the antibodies are highly diverse due to random shuffling of all six individual Complementarity Determining Regions obtained from healthy normal donor antibodies. The constant regions, however, are identical which facilitates uniform binding, stability, antibody-surface interactions and detection on the surface of a biochip. Domantis Limited (Cambridge, UK) uses domain antibodies (dAbs), which consist of the smallest binding unit of human antibodies, the V_H or V_L domain. The unique size and shape of dAbs enable them to bind to a wider range of targets than conventional antibodies, such as deep cavities that conventional antibodies cannot access. All of these companies have the potential to use their technologies to solve the content problem, but are currently focused on developing therapeutics. These companies, including Cambridge Antibody Technology, will likely need partnerships and collaborations to transform their technologies into protein biochip content.

In addition to antibody-based methods, several other companies are using innovative approaches to generate entirely new classes of capture agents. Archemix Corporation (Cambridge, Massachusetts, USA) has developed the RiboReporter™, which is an engineered ribozyme (RNA enzyme) that binds to a molecular target. Upon binding, the enzymatic function is activated, and a signal is generated, thus obviating the need for a separate label. SomaLogic, Inc (Boulder, Colorado, USA) is developing a similar technology based on photoaptamers, which are variants of aptamers (structured nucleic acids that bind to their targets in a manner analogous to antibodies). Photoaptamers can be covalently crosslinked to their protein targets following photoactivation, thereby allowing stringent washing. Bound proteins are detected using a universal protein stain that does not cross-react with nucleic acids, also eliminating the need to label the target proteins.



Fluorescent Image of the Protometrix, Inc Yeast ProtoArray™. Microscope slides were spotted with 5,000 different proteins cloned from yeast. Each protein on the array was detected using a Cy5-labelled antibody directed against an attached epitope tag. Copyright 2003 Protometrix, Inc

Aspira Biosystems, Inc (South San Francisco, California, USA) is developing ProteinPrint™ technology, which involves polymerisation of proprietary monomers around a linear peptide corresponding to a portion of the protein of interest. The polymer is cross-linked and forms a glove around the peptide. Due to the almost limitless number of monomers, this technology has the

potential to have the largest binding diversity of any of the capture agent methods. Phylos, Inc (Lexington, Massachusetts, USA) is developing TRINECTIN™ capture agents which are based on a fibronectin scaffold. The company reports stability up to 110°C, which would enable the use of a broad range of protocols for studying protein expression, but the commercialisation has been slower than expected. All of these approaches are promising but have yet to produce binders with consistently high affinity.

If one or more of these methods can produce high-affinity binders in large numbers, the next challenge will be to limit cross-reactivity, especially in multiplexed sandwich-type assays, in which a primary capture agent is paired with a secondary detection agent to increase specificity. There is a risk of cross-reactivity between the capture and detection antibodies and between the capture agents and other ligands present in the sample. Molecular Staging (New Haven, Connecticut, USA) may have identified the upper limit for multiplexed antibody sandwich assays. The company published a study using a protein biochip to quantify 75 different cytokines, but could not perform all 75 measurements on one biochip due to cross-reactivity, and was forced to divide the biochip in two, with 37 and 38 separate pairs of antibodies[3]. Other types of capture agents may not have the same cross-reactivity, but this remains to be seen. To solve the content problem, capture agents need to have high affinity and selectivity.

The second strategy for addressing the content problem is to divide and conquer the proteome and produce high-value, focused content. The most valuable classes of proteins, in terms of commercial potential, are druggable targets, such as kinases, proteases and receptors. In addition, researchers are interested in pathway analysis and post-translational modifications to better understand protein function and/or activation. The goal is to identify drug targets that can be modulated to produce biologically significant changes with minimal side-effects – the ideal target is rate-limiting for a specific pathway. Furthermore, biotechnology and pharmaceutical companies are actively pursuing better ways to predict the efficacy and toxicity of therapeutics. The goal is to have a predictive set of surrogate protein markers that can easily be screened – companies could save billions of dollars in development costs and provide safer, more effective treatments to patients.

Many companies are using the divide and conquer approach. Corning Incorporated Life Sciences (Acton, Massachusetts, USA) is developing tech-

Table 1: Commercial strategies to solve the content problem

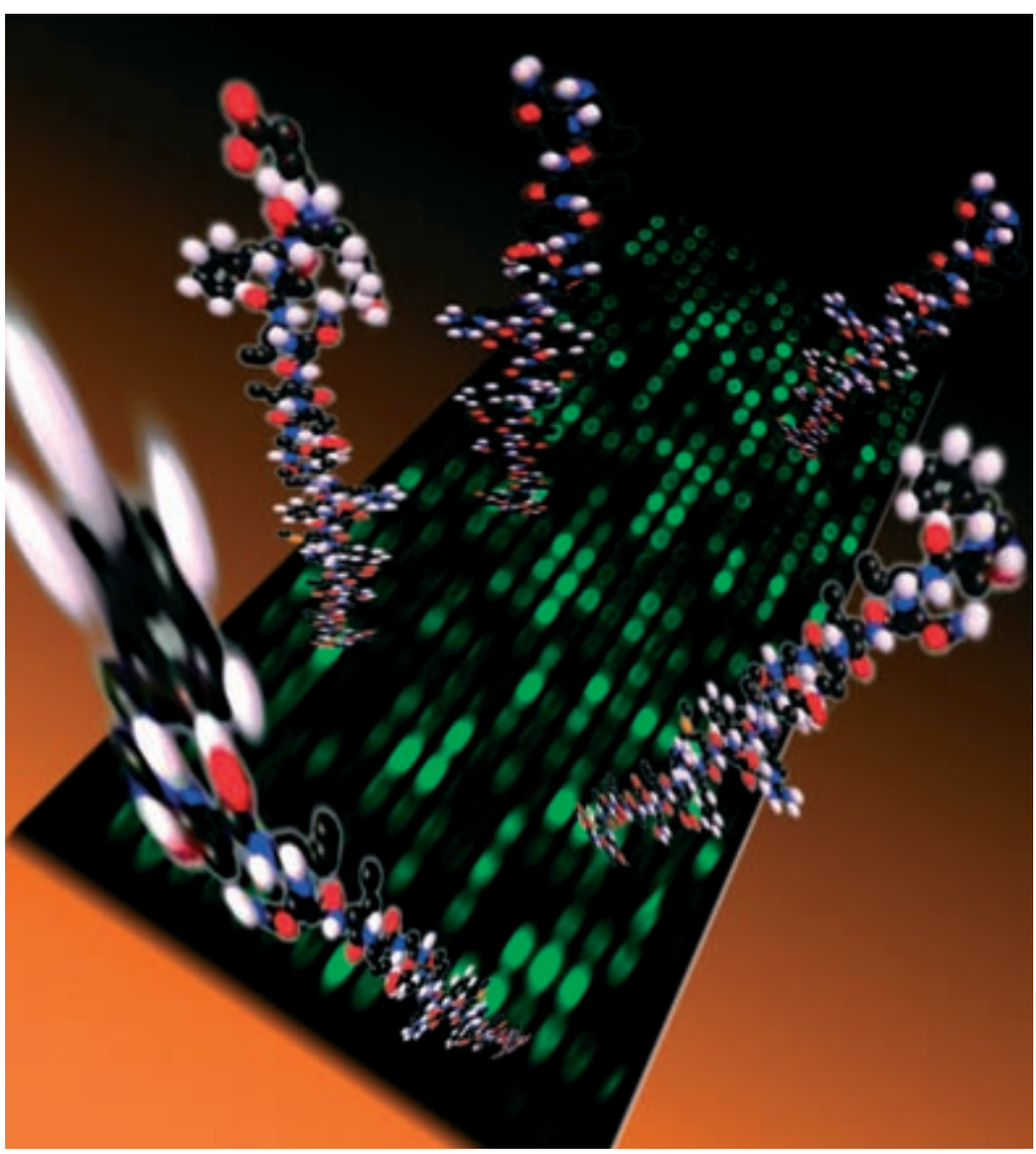
IMPROVED CAPTURE AGENTS	DIVIDING AND CONQUERING THE PROTEOME	BETTER DETECTION METHODS	SYNTHETIC PEPTIDES AND RECOMBINANT PROTEINS
Archemix Corporation	Corning Incorporated Life Sciences	Biacore International AB	Discerna Ltd
Aspira Biosystems, Inc	Hypromatrix, Inc	BioForce Nanosciences, Inc	Jerini AG
Affibody AB	Molecular Staging	Ciphergen Biosystems, Inc	NextGen Sciences Ltd
BioInvent International AB	ProteinOne Inc	HTS Biosystems, Inc	Pepscan Systems BV
Cambridge Antibody Technology Group plc	SomaLogic, Inc	Nanotype GmbH	Procognia Ltd
Domantis Limited	TeleChem International, Inc	Protiveris Inc	Protagen AG
SomaLogic, Inc	Zyomyx, Inc		Protometrix, Inc
			VBC-GENOMICS Bioscience Research GmbH

nology for embedding G-protein coupled receptors in lipid bilayers. The main application is screening compounds for binding and selectivity. Hypromatrix, Inc (Worchester, Massachusetts, USA) has three membrane arrays on the market with focused content: signal transduction, apoptosis and the cell cycle. ProteinOne Inc (College Park, Maryland, USA) has an array of proteins spotted on to membranes. The company plans to provide three focused biochips: transcription, cancer and universal cell cycle. SomaLogic is developing focused biochips for the measurement of proteins related to angiogenesis and inflammation. The company currently has 75 aptamers from which it intends to launch the angiogenesis and inflammation arrays and is commencing work on oncology arrays. TeleChem International, Inc (Sunnyvale, Calfornia, USA) sells arrayers and surfaces for producing protein biochips, and plans to sell pre-fabricated biochips with different themes, such as autoimmune disease, neurology, cancer, apoptosis and cytokines. Zyomyx, Inc (Hayward, California, USA) immobilises antibodies and other capture agents on a proprietary surface that minimises non-specific binding. In early 2003, the company launched a protein biochip that detects and quantifies 30 human cytokines. Molecular Staging has developed protein biochips based on sandwich assays (see above), and has formed a collaboration with Eli Lilly & Co (Indianapolis, Indiana, USA) to test 1,700 blood samples from a phase III clinical trial of Lilly's new sepsis drug. Molecular Staging will use its biochips to measure 130 different blood proteins – predominantly cytokines, interleukins, cell death-related factors, TNF alpha and erythropoietins. The platforms from all of these companies mitigate the content problem by making the most of available content.

The third strategy is to improve the detection methods. Protein biochip platforms based on detection – surface plasmon resonance (SPR) from Biacore International AB (Neuchatel, Switzerland) and Surface-Enhanced Laser Desorption/Ionization (SELDI) mass spectrometry from Ciphergen Biosystems, Inc (Fremont, California, USA) – currently dominate the market, but primarily use surface chemistry rather than capture agents to bind proteins. (Biacore, however, does have a collaboration in place to explore the use of capture agents and Ciphergen's platform can be

used with antibodies.) The challenges in adapting these platforms for use with capture agents will be increasing throughput and reducing cost. To be cost-effective discovery and screening tools, protein biochips need to be fast enough and cheap enough to enable researchers to generate statistically significant results. Biacore and Ciphergen have both made progress by increasing the densities of their biochips to increase throughput. In addition, HTS Biosystems, Inc (Hopkinton, Massachusetts, USA) is developing lower-cost, higher-throughput SPR technology for protein biochips, but has yet to commercialise a product. SPR and mass spectrometry may become viable solutions to the content problem, but more development is needed.

Several companies are developing novel technologies specifically for improving detection and quantification with capture agents, while keeping costs down and providing at least moderate throughput. The main advantage of these technologies, and of SPR and mass spectrometry, is that they provide characterisation of the binding interaction between the capture agent and ligand, whereas fluorescence, the most common detection method for protein biochips, only provides an endpoint measurement. BioForce Nanosciences, Inc (Ames, Iowa, USA) is developing atomic force micsrocopy (AFM) for reading biochips. AFM measures changes in height at the molecular level, and thus provides additional information about the binding of ligands. Nanotype GmbH (Gräfelfing, Germany) is developing technology that sandwiches two biochips together. The bottom biochip con-



tains the primary capture agents, and the top biochip contains the secondary detection capture agents as well as labelled force sensor complexes. The force sensors allow researchers to set a threshold for detection and establish mechanical stringency. Moreover, by physically constraining the secondary capture agent, cross-reactivity is greatly reduced. Protiveris Inc (Rockville, Maryland, USA) uses microcantilevers attached to capture agents to detect proteins. The microcantilevers bend when exposed to a solution containing the target molecule. The bending is measured by a change in the angle of reflected laser light, thus producing a binding curve rather than an endpoint measurement. These methods may compensate for suboptimal capture agents by providing quality control information as part of the detection.

The fourth strategy is to leverage genomics databases and technology to select peptide sequences for synthesis and to express and purify recombinant proteins. The result is the generation of protein biochips that are used to study interactions, providing information about pathways, about protein, compound and capture agent affinity and selectivity and about allergies and autoimmune diseases. Protometrix, Inc (Branford, Connecticut, USA) has technology for expressing and purifying proteins in high-throughput and immobilising them in their active conformations. The company's first product, the yeast ProtoArray, contains almost 5,000 Saccharomyces cerevisiae proteins and has been validated by a major publication[4]. The company plans to launch the yeast ProtoArray in mid-2003, and is planning additional protein biochips with focused content from humans and other species, but has not disclosed details. Discerna Ltd (Cambridge, UK) synthesises proteins *in vitro* by cell-free transcription and translation from PCR-generated DNA and directly immobilises them on to suitable surfaces via incorporated 3' tags. Jerini AG (Berlin, Germany) and Pepscan Systems BV (Lelystad, The Netherlands) both use high-output peptide synthesis to generate peptide arrays. The primary application of Jerini's biochip is to determine the substrate specificity of kinases, phosphatases and proteases; the primary application of Pepscan's biochip is antibody characterisation. NextGen Sciences Ltd (Cambridgeshire, UK) is developing an integrated suite of technologies for automating protein expression and purification. Procognia Ltd (Berkshire, UK) expresses proteins from a cDNA library into which a control for proper protein folding has been engineered. Protagen AG (Dortmund, Germany) has a set of more than 10,000 unique, non-redundant, recombinant human proteins (from fetal brain library hEx1) that are routinely expressed and purified in high throughput. The proteins can be immobilised directly for use on the biochip, or used as antigens to generate antibodies or other capture agents. In addition, the company plans to develop protein biochips to test for autoimmune diseases such as multiple sclerosis and rheumatoid arthritis. VBC-GENOMICS Bioscience Research GmbH (Vienna, Austria) has developed a biochip with proteins immobilised on the surface to test for allergies. The efficacy of the biochip has been proven in clinical trials, and the company plans to launch a commercial version by the end of 2003. All of these protein interaction platforms are expected to propel the growth of the protein biochip market, even if the lack of high-quality capture agents persists.

The content problem is not only slowing the development of protein biochips, but it may be slowing the development of the entire proteomics industry. Proteomics researchers are generating a flood of potential biomarkers and drug targets, but are facing a growing bottleneck in screening[5]. Just because there is a change between disease and normal samples does not immediately confer any biological significance. There are normal fluctuations in protein expression, distribution, post-translational modifications, interactions and functions. Validation is critical, yet the high-output technologies for proteomics discovery – 2-D gel electrophoresis, chromatography and mass spectrometry – give way to low-throughput, low-output technologies for validation. Protein biochips have the potential to widen this bottleneck by providing medium-output and high-throughput validation of potential biomarkers and drug targets, but only if the content problem is solved. DDW

Dr Steven Bodovitz *is Principal Consultant of Select Biosciences (Sudbury, UK) where he has written Protein Biochips 2003 and Proteomics 2003, and has presented his analyses at numerous international conferences on those topics. Prior to Select Biosciences, Dr Bodovitz was a co-founder of BioInsights (Cupertino, CA, USA) where he wrote numerous reports on drug discovery, including Biochips: From Hype to Reality, Protein Biochips: On the Threshold of Success and Profiting from Biological Complexity: The Emergence of Systems Biology. In addition he has worked with numerous biotechnology and biopharmaceutical companies to assess their market opportunities and optimise their business strategies. Dr Bodovitz received his PhD in neuroscience at Northwestern Universtiy.*

References

1 Select Biosciences, Protein Biochips 2003, April 2003.
2 Schier et al (1996). J. Mol. Biol. 263, 551-567; Gao et al (1999). Proc. Natl. Acad. Sci. U S A, 96, 6025-6030.
3 Schweitzer et al (2002). Nature Biotechnology, 20, 359-365.
4 Zhu et al (2001). Global analysis of protein activities using proteome biochips, Science, 293, 2101-2105.
5 Select Biosciences, Proteomics 2003, July 2003.