# Probing the proteome – protein arrays and their applications

**Rhian E. Hayward***, **Grant Cameron** and **Roland Z. Kozlowski**, Sense Proteomic Ltd, Babraham Hall, Babraham, Cambridge, UK CB2 4AT, *tel: +44 1223 492080, fax: +44 1223 492081, e-mail: rhian.hayward@senseproteomic.com

A precedent has been set for highly parallel experimentation by the use of DNA microarrays for drug target identification[1]. To follow suit, the creation of a similar tool for proteomics is rapidly moving forward. As the drug discovery field awaits the first commercially available protein, peptide or antibody array, the IBC's *Protein Microarray Technology* meeting (27–28 September 2001, Berlin, Germany) suggested that such a product is likely to be available before 2002. Protein array developers from industry and academia reviewed the growing collections of proteins for arraying, their immobilization, functionality and format, focusing on construction, sensitivity of detection, scale and applications.

Encouraging progress has been made since the last meeting just six months ago (IBC's *Protein Microarray Technology*, 21–23 March 2001, San Diego, CA, USA), particularly in the creation and handling of the biological content of the arrays (proteins, antibodies and peptides) and the application of arrays in case studies. Other necessities remain under-subscribed and there were calls for, in particular, more efforts in bioinformatics and computing power to process proteome-size arrays and their applications.

## Antibody tools

The relative stabilities of immunoglobulin or immunoglobulin-based molecules when immobilized, their availability and our extensive knowledge of these molecules have led to the creation of antibody arrays. These capture-molecule arrays will be popular both as research tools and eventually as diagnostic tools to identify, for example, tumour marker proteins in cancerous tissue or surrogate markers of disease in patient serum, and subsequently as a tool for drug target discovery.

Antibodies and single-chain antibody phage-display libraries provide the means for substantial throughput and comprehensive *in vitro* selection experiments. Not surprisingly, at least 200 laboratories worldwide currently use phage display according to Ruud de Wilt (Diversys, Cambridge, UK). Diversys has applied phage display to the selection of specific antibodies from large naive antibody libraries. Additionally, Diversys has developed scFv-fragment, Fab-fragment and single-domain Superantibody™ libraries – recombinant antibodies encompassing superantigen characteristics of expression, folding and binding – all of which bind both protein A and protein L (microbial protein antigens) for purification and immobilization. In an example experiment, 27,000 different antibodies were screened in duplicate against 15 different antigens in two days.

An alternative technology that is currently under development is the high-throughput generation of monoclonal antibodies to components of anonymous protein mixtures described by Steve Pennington (University of Liverpool, Liverpool, UK). The 'inverse screening' of cloned hybridoma cells with the original complex protein mixture used to create them can identify cells that are producing mono-specific immunoglobulins. A shotgun approach to generating antibodies as a rapid means of gathering capture-molecules for diagnostic purposes could present a challenge to phage-display methods, and addresses the serious issue of the frequency with which commercially available antibodies are found to cross-react with unrelated proteins.

Markus Templin [Natural and Medical Science Institute (NMI), University of Tübingen, Tübingen, Germany] commented on how the characterization of interacting antigens is a bottleneck in antibody profiling. As well as generating the Human Combinatorial Antibody Library (http://www.hucal.com) via a patented technology marketed through Morphosys (Munich, Germany), the NMI is developing an array-based method to estimate the kinetic properties of a set of recombinant antibodies and rank the binding of antigens by their apparent affinities. These measurements must consider the size of an antibody spot with reference to the ambient analyte assay[2], which was highlighted by Thomas Joos (NMI) for consideration when miniaturizing both the probe (the spot) and the target. Morphosys is capturing antigens and measuring the relative binding affinity to Fab fragments from crude cell extracts using the surface plasmon resonance (SPR) technology from Biacore (Uppsala, Sweden).

William Robinson (Stanford University School of Medicine, Palo Alto, CA, USA) described how an array of 500 antigens (proteins and peptides), which act as targets of the autoimmune response in arthritis and other autoimmune diseases, has been created and screened by drawing heavily on arraying technology

1359-6446/01/$ – see front matter  PII: S1359-6446(01)02102-X 

first developed for DNA microarrays. Correlations between the presence of auto-antibodies in patient serum and observed pathology provided a convincing example of the application of antibody arrays. The profiling of tumour cell-surface receptors using antibody arrays to capture such cells was also reported by Ulrik Bjerl Nielsen (Massachusetts Institute of Technology, Cambridge, MA, USA). Phosphorylation of growth factor receptors over time could be detected in this 'micro-sandwich assay', which is essentially an ELISA in array format. This cannot be extensively multiplexed at present, that is, expanded to include a variety of functional assays, but it does provide a prototype tool for tumour profiling.

## The power of peptides

Protein arrays constructed from short synthetic peptides have several advantages that include having complete control over the sequences arrayed and combinatorial power when peptides are randomized. Stefan Schmidt described how GPC Biotech (Munich, Germany) is focusing on a kinase-substrate array to profile human kinase specificity. The company's approach is to generate fusion peptides comprising the experimental protein fused with a scaffold protein, such as green fluorescent protein (GFP) or thioredoxin, and an affinity tag for immobilization. Kinase substrates (350 identified, to date) are arrayed on membrane which, although it could affect functionality, can be interrogated for simple binding events using fluorescently labelled kinases.

Another peptide company that is reportedly close to marketing a peptide microarray product is Jerini AG (Berlin, Germany). Jens Schneider-Mergener described how the company synthesizes peptides and protein domains of <30 amino acids directly onto cellulose, polypropylene or glass and has applied this technology to enzyme profiling, including kinases, and mapping protein–protein interactions. Protein engineering by this method could be exploited to transform peptides of interest into drug-like structures and thus could make a significant contribution to drug discovery. In a similar approach, Jerini has derived fully synthetic protein domains by joining peptides together. Although this is unlikely to lead to the replication of conformation-dependent functions and interactions with other molecules, linear epitopes (i.e. those with no tertiary structure) can be recreated for analysis.

## The challenge of protein expression

After an extensive protein expression project, Dolores Cahill (Max-Planck Institute, Göttingen, Germany) has distilled a largely non-redundant collection of ~15,000 hexahistidine-tagged proteins from a human foetal brain cDNA library. Fusion proteins have been arrayed onto glass and membrane substrates for interrogation and, although these proteins are likely to be denatured on the surface, data were presented that demonstrated successful screening using antibodies.

This project raises several basic issues concerning attempts to express a proteome:

- Working with cDNA libraries, as noted by others during the meeting, requires the removal of significant redundancy;
- Several expression hosts are needed to express diverse proteins; and
- Large protein-expression projects are time consuming in the order of years using conventional strategies.

The difficulty of protein expression during the Max-Planck project has been minimized using robotic methods and it cannot be underestimated how valuable collections of large recombinant proteins will be in developing protein array technology further.

One of the most challenging areas of protein microarraying is the immobilization of full-length proteins and the functional interrogation of these molecules. Sense Proteomic Ltd (Cambridge, UK) is generating collections of recombinant proteins with a view to developing high-throughput pharmacological assays. However, rather than using conventional cloning techniques the proteins are gathered by tagging an entire cDNA library in one tube, with no need for previous sequence information, using their proprietary technique known as COVET™. Full-length proteins are expressed, and then immobilized in a site-specific manner via a terminal affinity tag. Data were presented by Roland Kozlowski (Sense Proteomic) showing that these proteins remain functional in an array format. Protein–DNA, protein–protein, protein–small molecule/drug and protein enzymatic modification on the Sense Proteomic arrays were demonstrated. A major application of these arrays of functional proteins will be their integration into the drug discovery process. The identification of drug targets will be possible using arrays of proteins that are whole, soluble and functional.

## Technology development

For the arraying of functional proteins, Steffen Nock (Zyomyx, Hayward, CA, USA) presented a modular substrate surface that differs significantly from the flat surfaces used by others. The architecture of the Zyomyx surface comprises pillars that are engineered into a silicon wafer and coated with a monolayer, such as biotin-derivatized poly(L-lysine)-g-poly(ethylene glycol), which resists non-specific protein adsorption. Packaged alone, or surrounded by flow cells for a microfluidics environment, these biochips require ~100 amols of protein for saturation. Nock called for more functional proteins to test out the Zyomyx biochip. Data shown confirmed that the oriented attachment of capture molecules leads to far more efficient binding of antigen than attachment via random surface residues.

Andrei Mirzabekov (Engelhardt Institute of Molecular Biology, Moscow, Russia)

described how the retention of protein function after immobilization could be reliably achieved with a stabilizing three-dimensional polyacrylamide-based gel. Using pads of hydrogel surrounded by a hydrophobic surface, DNA, immunoglobulins and other proteins can be immobilized either by diffusing molecules into the gel pad or by incorporating molecules into the pad during photopolymerization of the gel. Interestingly, it was reported that antibodies can be stored within the gel for weeks or more and, although there were no data on the storage of more labile proteins in this matrix, it is promising that a hydrogel-like surface (now commercialized by Packard BioScience, Meriden, CT, USA) might begin to address the practicalities of the storage and transport of protein microarrays.

## Summary

A proteome is an extremely complex mixture containing many expressed versions of the same genes resulting from post-translational modification. Hanno Langen from Roche (Basel, Switzerland) commented that sampling of the protein complement of a cell should be considered on an organelle by organelle basis to simplify mixtures and avoid missing rare proteins. Similarly, Ian Humphrey-Smith (Glaucus BV, Utrecht, The Netherlands) added that attempts to array proteins to cover an entire proteome would require substantial computing power. A notable absence from this meeting were reports of the use of bioinformatics tools and there was a call for the employment and expansion of proteomics-compatible software.

Whole open reading frames are difficult to express in heterologous systems. Therefore, collections of proteins for functional interrogation are still rare. It is likely that the first protein microarray or array products on the market will be peptide-based because of the relative ease of synthesis of peptides. Antibody arrays might follow, possibly as diagnostic tools, but only after extensive quality assurance to exceed the standards set by the enzyme-linked immunosorbent assay (ELISA). Arrays of functional proteins, glycans (Glycominds, Maccabim, Israel) and whole cells (Genescan Europe AG, Freiburg, Germany) will require further optimization. Additionally, it appears that companies working to improve surface chemistries might need to collaborate with those generating protein or peptide collections to create protein arrays. Expertise for both, unlike the DNA microarray field, is unlikely to be found entirely within one group or company. Self (consumer)-fabrication of protein arrays is another unlikely possibility because of the difficulties encountered with protein expression.

## References

1 Jain, K.K. (2000) Applications of biochip and microarray systems in pharmacogenomics. *Pharmacogenomics* 1, 289–307

2 Ekins, R.P. (1989) Multi-analyte immunoassay. *J. Pharm. Biomed. Anal.* 7, 155–168