JPP Journal of Pharmacy and Pharmacology

JPP 2007, 59: 609–628

Received August 17, 2006
Accepted February 5, 2007
DOI 10.1211/jpp.59.5.0001
ISSN 0022-3573

Review Article

The use of proteomics to identify novel therapeutic targets for the treatment of disease

Fleur L. Moseley, Katrina A. Bicknell, Michael S. Marber and Gavin Brooks

Abstract

The completion of the Human Genome Project has revealed a multitude of potential avenues for the identification of therapeutic targets. Extensive sequence information enables the identification of novel genes but does not facilitate a thorough understanding of how changes in gene expression control the molecular mechanisms underlying the development and regulation of a cell or the progression of disease. Proteomics encompasses the study of proteins expressed by a population of cells, and evaluates changes in protein expression, post-translational modifications, protein interactions, protein structure and splice variants, all of which are imperative for a complete understanding of protein function within the cell. From the outset, proteomics has been used to compare the protein profiles of cells in healthy and diseased states and as such can be used to identify proteins associated with disease development and progression. These candidate proteins might provide novel targets for new therapeutic agents or aid the development of assays for disease biomarkers. This review provides an overview of the current proteomic techniques available and focuses on their application in the search for novel therapeutic targets for the treatment of disease.

Introduction

What is proteomics?

The term "proteome" was initially used to describe the set of PROTEins expressed by a genOME, cell or tissue, with multicellular organisms containing a number of different subproteomes (Wasinger et al 1995; Wilkins et al 1996a, b). The concept of mapping the human proteome was conceived over 20 years ago but rapid development in the technology of molecular biology led to a focus in genomic approaches and DNA sequencing (Anderson & Anderson 1982; Banks et al 2000). The completion of the Human Genome Project in 2003 has offered scientific researchers exceptional opportunities to improve our understanding of human cell biology and has provided a multitude of potential avenues for pursuing the identification of therapeutic drug targets (International Human Genome Sequencing Consortium 2004). However, this huge sequence database is only a starting point to identify the molecular mechanisms involved in the functional regulation of a cell or tissue during both health and disease (Brewis 1999).

Proteomics has been enhanced by the earlier achievements of genomics, which has supplied a database of possible gene products that can provide the basis for subsequent proteomic studies (Tyers & Mann 2003). Proteomic techniques are able to evaluate changes in protein expression, post-translational modifications, 3D protein conformation, alternative splicing, protein–protein interactions and protein localization, all of which are important for determining the function of a protein within the cell (Banks et al 2000). Unlike the genome, the proteome of a cell is dynamic and depends on the tissue, cell type and the developmental status or health of the cell. Cellular and environmental factors such as pH, hypoxia and drug administration may also affect the proteome of a cell, in particular during disease (Banks et al 2000). Thus, genomic approaches alone are insufficient to investigate the causative mechanisms underlying disease. Protein abundance is regulated not only through changes in gene expression but also can be influenced by delayed translation at the mRNA level and changes in protein degradation. It is important to remember that studies carried out on mRNA do not always correlate to protein levels within the cell (Pandey & Mann 2000). In addition, protein abundance does not necessarily compare with protein activity within a

School of Pharmacy, The University of Reading, Whiteknights, Reading, Berkshire, RG6 6AP, UK

Fleur L. Moseley, Katrina A. Bicknell, Gavin Brooks

Division of Cardiology, Kings College London, The Rayne Institute, St Thomas' Hospital, London, UK

Michael S. Marber

Correspondence: G. Brooks, School of Pharmacy, The University of Reading, Whiteknights, Reading, Berkshire, RG6 6AP, UK.
E-mail: g.brooks@reading.ac.uk

Funding: Funded through the BBSRC (grant number BB/C509458/1).

cell; thus, proteomic techniques that study protein function are powerful tools to improve our understanding of cell signalling in healthy cells and during disease (Burbaum & Tobal 2002).

The current version of the human gene catalogue contains 22 287 gene loci (19 438 known genes and 2188 predicted genes) with a total of 34 214 transcripts (International Human Genome Sequencing Consortium 2004). The number of proteins encoded by the genome is unknown but it is thought that a single gene could encode as many as 10–20 different protein species. Post-translational modification of proteins also increases the diversity of protein function. Proteins can be modified in several ways including alterations in protein structure, such as the formation of disulphide bridges or proteolytic cleavages. Additionally, a biochemical functional group can be attached resulting, for example, in protein phosphorylation, acetylation or methylation or the addition of lipid (such as palmitoylation), protein (for example ubiquitination) or carbohydrate groups (glycosylation). Phosphorylation and dephosphorylation are fundamental for signal transduction and are central in regulating numerous biochemical pathways within the cell from enzyme activity to receptor conformation. Moreover, acetylation of histone proteins regulates the local unwinding of DNA facilitating the access of transcription factor complexes enabling gene transcription. Therefore, it is changes in post-translational modifications of proteins that might be more important in regulating protein function as opposed to changes in protein abundance. A thorough understanding of the role of post-translational modifications of proteins is likely to reveal the mechanisms underlying cell signalling events, cellular compartmentalization and intracellular relocalization of proteins, changes in protein activity and stability and protein–protein interactions. The huge diversity of these molecules enables proteins to regulate the majority of biochemical processes within cells, tissues and entire organisms. Thus, proteins form the biggest group of drug targets and are also used as drugs themselves, confirming the importance of proteomic techniques for the identification of future therapeutic approaches.

Proteomic techniques have a multitude of applications (summarized in Figure 1), which can be used to further our understanding of cell function in healthy and disease states. These studies are likely to reveal new enzymes, signalling molecules and pathways which might be potential clinical targets (Edwards 2000). Proteomics also aims to improve our understanding of cellular biochemistry by further unravelling the complex network of protein–protein associations present in the cell that are crucial for the regulation of the majority of cellular processes. Biological units that are disrupted or play a key role in disease processes could be targeted or mimicked by new therapeutic drugs.

Proteomics and drug discovery

From its onset, proteomics has been used to study the protein profiles of cells in disease states. Identifying changes in gene expression, protein production or protein modification that correspond with disease can focus the search for therapeutic targets, but they do not reveal any information regarding the biochemical mechanism through which an alteration in a gene product mediates such an effect. Therefore, to increase our understanding of disease development, proteomic approaches need to investigate changes in protein function. The fact that a single gene can give rise to several protein products also increases the complexity of identifying a disease-related protein. Therefore, a thorough proteomic study that initially identifies a disease-associated alteration in a specific protein and concludes by determining the cellular function of that protein (including interactions with other signalling molecules) should improve the efficiency of drug discovery and development.

Proteomic approaches are also valuable for aspects of drug development beyond the initial identification of a drug target; for example, proteomics can be implemented to monitor therapeutic and toxic effects of newly developed compounds by specifically evaluating changes in the proteome of cell lines, isolated tissues or bodily fluids following treatment. Additionally, proteomic techniques can be used to assess the interaction between the drug and therapeutic target protein, which can subsequently enable modification of the drug to improve drug affinity, efficiency and efficacy. The specificity of a potential drug can also be evaluated using proteomics by using a method that isolates and identifies all proteins with an

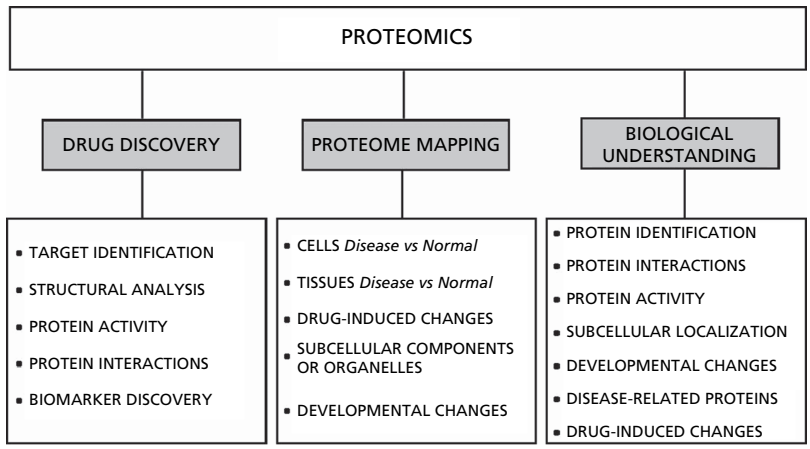

Figure 1 A summary of the potential applications of proteomics.

affinity for a compound. Such information can be used to chemically alter the drug in an effort to improve drug selectivity and reduce non-specific binding.

Proteomics and personalized medicine

In a similar manner to the identification of disease-associated proteins, proteomics can be used to understand the effects that a particular drug has on a patient. Cellular proteomes can be compared before and after drug treatment to distinguish between which proteins or signalling pathways are influenced by a specific dosage. This may lead to the identification and improved understanding of the potential clinical side effects of a particular drug. The rapid development of protein array technology potentially brings us a step closer to the prospect of personalized drug regimens. Like pharmacogenetics, individual molecular profiling, or pharmacoproteomics, might reveal patients that exhibit identical disease phenotypes but that have developed due to alterations in different biochemical pathways; thus, such individuals might be more receptive to particular treatments. Regimens like these might minimize the problem of side effects. If routine screening of protein activity could be achieved this might also enable patient-specific drug doses to be administered, which in turn might reduce side effects and improve drug selectivity.

Proteomics clearly has great potential for the progress of several aspects of drug development but before such studies can be carried out consideration must be given to identify the most suitable technical approach that will provide the information required. The core elements of the classical approach to proteomic research are schematically summarized in Figure 2, highlighting the different techniques that can be applied to each step of the study. It is important to consider these options carefully; for example, the procedure must be sensitive enough to detect low abundance proteins (unlike genes, proteins cannot be amplified) in conjunction with protein activity as well as identifying both protein–protein and protein–drug interactions within the cell, and all these criteria must be addressed with routine techniques that are quick to perform and easily implemented (Adam et al 2001; Burbaum & Tobal 2002). With this in mind, the remainder of this review summarizes the current proteomic technology available to assist in the search for novel therapeutic targets and concludes with examples of how such technology is being used in practice.

Current proteomic techniques

Sample preparation

The most significant step in any proteomic approach is sample preparation. As the proteome is highly dynamic it is imperative that the samples to be analysed correspond to the question that has been posed. Careful consideration is required to ensure that any changes in the proteome observed are physiologically significant and are not due to artefacts of the preparation technique, particularly protein degradation. It is also important to contemplate the relative abundance of the proteins of interest. Sample enrichment or cellular fractionation methods may be required to increase the concentration of proteins to a level that can be detected or to remove high abundance proteins that might mask the changes in a specific protein. For example, novel targets might be present on the cell surface; thus, careful sample preparation, potentially in conjunction with membrane isolation, might be required to identify these hydrophobic proteins (Tyers & Mann 2003). Biological and technical variability is also a concern in proteomic studies. The collection of multiple samples limits the variability associated with sample preparation and allows protein samples to be separated and analysed in parallel thus minimizing technical variation.

The sample preparation method used can depend on a number of factors ranging from the initial protein source to the proteomic technique employed. Solubilization of proteins from a tissue might require an initial homogenization step in solubilization buffer containing more stringent detergents compared with a cell pellet. It also is worth considering that the composition of a tissue is far more heterogeneous than a homogeneous cell line and is likely to contain a number of different cell types and, thus, a wider range of proteins. Commercially available extraction kits are available that utilize a combination of solubilization buffers, which contain different detergents and a number of different centrifugation steps. These kits can be used to enrich proteins from various cellular compartments, for example, nuclear versus cytoplasmic proteins or hydrophobic membrane proteins, and can be used to provide a more focused starting point for a proteomic study. However, if comparisons are to be made between samples it is important to ensure that any enrichment process produces reliable and reproducible results.

Immunoprecipitation also can be used to enrich a protein or proteins of interest. Proteins are isolated from a complex mixture using an analyte-specific reagent, such as an antibody which is attached to a matrix. This method not only isolates the proteins that directly interact with the analyte-specific reagent but also the partner proteins that interact with them. The protein of interest also can be fused to a high affinity tag, introduced to a complex protein sample and subsequently isolated using the appropriate capture reagent (Ramachandran et al 2005). Both methods will isolate a number of proteins, making it difficult to determine whether a protein is interacting directly or indirectly with the protein of interest (Ramachandran et al 2005). However, this technique remains valuable as a tool to enrich the proteins of a specific signalling complex to levels that can be either visualized by gel-electrophoresis or analysed directly by mass spectrometry.

Additionally, it is important to ensure that the solubilization buffer used for sample preparation is compatible with the desired proteomic application. For example, to map the proteome of a cell or tissue by 2D electrophoresis, non-ionic detergents must be used to enable separation in the first dimension. However, if separation by 1D electrophoresis is sufficient for the study proposed, ionic detergents can be employed, improving the solubilization of hydrophobic proteins significantly. Therefore, it is crucial to plan any proteomic study from start to finish, ensuring that the experimental design is compatible with all steps of the project, and that the data obtained is reliable and relevant to the scientific problem originally posed.

Two-dimensional electrophoresis (2DE)

The central technology for separating cellular proteins is 2DE (summarized in Figure 3). Proteins are initially separated by

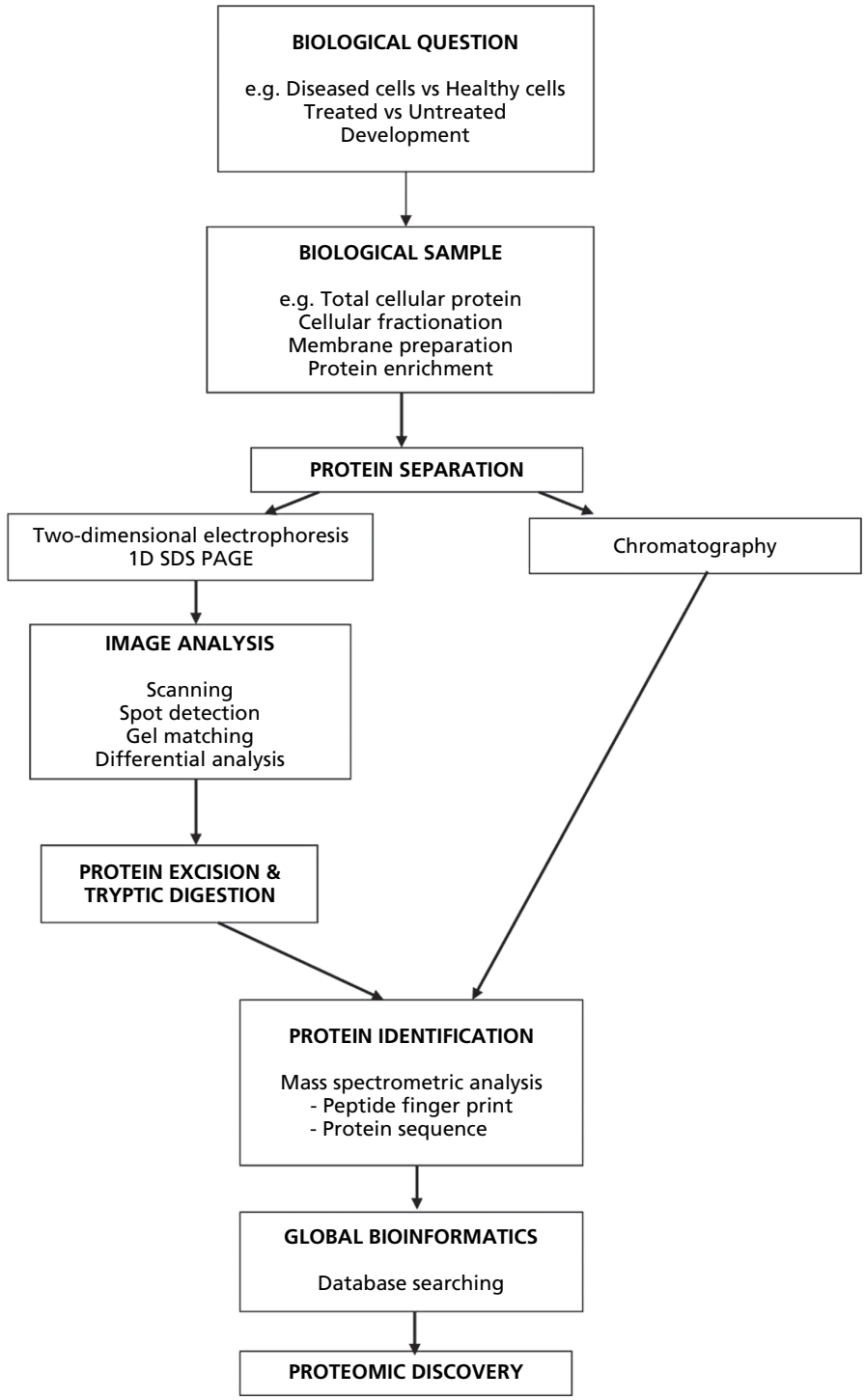

Figure 2 The standard approach to proteomics.

charge in the first dimension and subsequently by molecular weight in the second dimension on a polyacrylamide gel (Banks et al 2000). Traditionally, the analysis of two or more proteomes to identify differentially expressed proteins begins with the separation of each protein sample on individual 2D gels. This is followed by staining and comparison and quantification using computer-aided image analysis programs (Mahon & Dupree 2001; Cutler et al 2003; Hagenstein & Sewald 2006). Although 2DE is used regularly for proteomic studies, there are a number of limitations associated with the technique. Historically, 2DE is considered highly variable but the development of commercially available immobilized pH gradient (IPG) strips has significantly improved the reproducibility of the first dimension (Gorg et al 1988). These polyacrylamide gel strips

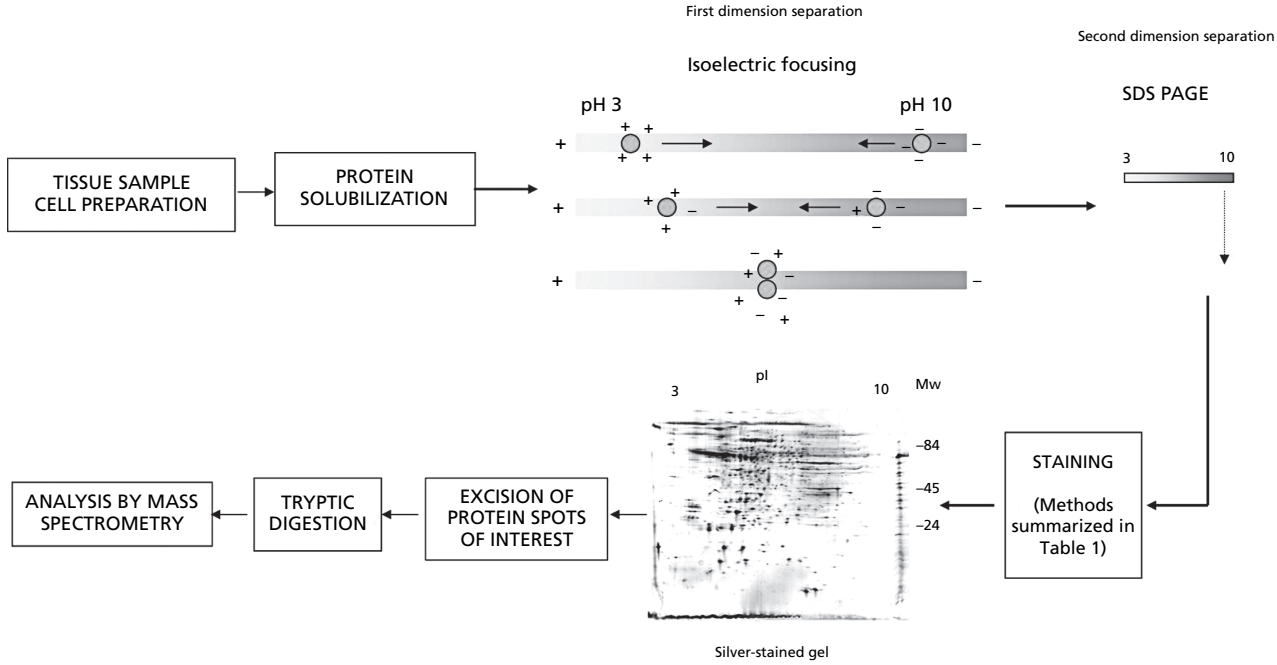

Figure 3 The conventional approach to two-dimensional electrophoresis. Proteins are initially isolated and solubilized from the tissue or cells of interest. The proteins are then separated in the first dimension according to their isoelectric point, which is the pH at which the protein is electrically neutral. Following isoelectric focusing the proteins are separated in the second dimension by standard SDS-PAGE, which separates the proteins by molecular weight. Gels are then stained (a silver-stained gel is shown as an example) and protein spots of interest are excised, digested with trypsin and analysed by mass spectrometry to obtain either a peptide fingerprint or de-novo peptide sequence data.

Table 1 A comparison of the sensitivity and linear dynamic range of the current staining methods for 2DE

Staining method	Sensitivity (ng)	Linear dynamic range (orders of magnitude)
Coomassie R-250	> 50	1
Coomassie G-250	6–8	1
Silver	< 1	1
SYPRO Ruby	< 0.5–1	3
Deep Purple	< 0.5–1	4
Radiolabelling	0.1	4–5
CyeDye	0.125	5

contain an immobilized pH gradient that can range from pH 3 to 11, enabling proteins to migrate to their isoelectric point (the pH at which the overall charge is neutral). A wide range of pH can be used as an initial screening tool and preliminary studies can be followed up with more focused experiments using narrower pH ranges, allowing a greater degree of separation within the proteins of interest. In addition, pre-cast gels are commercially available thus limiting variability in the second dimension.

Protein spots can be visualized using a range of staining methods with varying degrees of sensitivity and linear dynamic range (Table 1). Until recently, gel staining has focussed on Coomassie G-250 and silver staining that can be used to detect protein spots containing 6–8 ng and 1 ng, respectively. Silver staining clearly provides increased sensitivity but has a poor linear dynamic quantification range, making it impossible to quantify small changes in protein expression. The development of fluorescent stains, namely SYPRO-Ruby and Deep Purple, has provided a staining method of similar sensitivity to silver staining (~1 ng) but with significant improvements in the linear dynamic quantification range (approximately 3 orders of magnitude compared with one order of magnitude for silver staining), enabling changes in protein abundance associated with disease or drug treatment to be more accurately evaluated (Mackintosh et al 2003; White et al 2004).

Two dimensional fluorescence difference gel electrophoresis (2D DIGE) involves labelling protein samples with fluorescent dyes before separation by 2DE. Two samples can be labelled with different CyDye DIGE Fluors (that are spectrally distinct) in conjunction with an internal control, which are all separated on a single 2D gel thus eliminating inter-gel variability. The internal standard is a combination of equal quantities of both protein samples of interest. Including this standard enables accurate quantification of changes in protein abundance between the two experimental preparations by comparing changes in the ratio of relative abundance of a matched protein spot (Unlu et al 1997; Gharbi et al 2002). The use of CyDyes also dramatically improves the levels of sensitivity, detecting as little as 125 pg protein, and does not interfere with the process of protein identification by mass spectrometry (Marouga et al 2005).

Post-translational modification of proteins can also be identified using 2DE. Originally these studies involved either

radiolabelling, for example the incorporation of ^{32}P to investigate protein phosphorylation, or 2D Western blotting using antibodies raised against specific post-translational modifications, which requires the production of duplicate gels enabling the Western blot to be matched to a standard 2D gel, thus permitting the identification of any interesting protein. Biotinylation of surface proteins before sample preparation and 2D separation has also proved a valuable technique to investigate and identify surface membrane bound proteins. However, these techniques have recently been superseded by fluorescence-based detection methods which are able to detect proteins containing phosphate groups (Pro-Q Diamond), glycoproteins (Pro-Q Emerald) and transmembrane proteins (Pro-Q Amber) in gel without the need for radioisotopes, antibodies or Western blotting.

2DE is still widely utilized today as a starting point to investigate changes in protein expression during development, disease or drug treatment and is routinely followed by mass spectrometric analysis to identify any proteins of interest. However, there are some limitations with this technique, namely constraints on protein coverage with only the more abundant proteins being detected. Less abundant proteins can be studied but require sample enrichment or cell fractionation to increase protein levels before separation (Gygi et al 2000). Zoom gels use narrow range IPG strips to obtain a more detailed overview of the proteins over a limited pI range and can potentially further segregate proteins of varying abundances that might originally resolve as a single protein spot using the standard wide range.

An additional limitation of 2DE is the systematic exclusion of highly hydrophobic proteins that are often membrane bound (Santoni et al 2000). Membrane proteins have been said to represent nearly 50% of important drug targets; thus, 2DE might omit a series of potentially interesting therapeutic targets (Drews 2000). This exclusion can be reduced by using a combination of detergents during sample preparation but cannot be eliminated entirely. 2DE is also unable to resolve proteins that are either too large to enter the gel or too small or electrically charged to be retained in the gel (Gygi et al 1999; Harry et al 2000; Adam et al 2001).

Liquid phase isoelectric focusing fractionates protein samples in free solution by isoelectric point. Protein fractions from individual runs can be pooled, potentially increasing the levels of low abundance species. Samples can then be further separated by 2DE, providing an enriched protein sample from the pI of interest. However, liquid phase isoelectric focusing can also be used to bypass 2DE and specific pI fractions can be separated on 1D SDS-polyacrylamide gel electrophoresis (PAGE), thus avoiding the exclusion of hydrophobic and high molecular weight proteins that is observed with 2D gels. This method can be used to maintain the native conformation of proteins, but several proteins might resolve in a single 1D band making specific protein identification difficult. However, this technology has additional benefits as it can also be combined with chromatographic separation methods or direct analysis by mass spectrometry thereby completely avoiding any gel-based protein separation (Wang et al 2003; Harper et al 2004; Moritz et al 2004; Xiao et al 2004).

2DE provides a unique method to separate and visualize complex protein mixtures, enabling changes in protein expression to be investigated. When used in combination with cellular fractionation methods, 2DE is a powerful method that can be used to investigate changes in protein compartmentalization within the cell. The technique requires a high degree of sensitivity and improved technical efficiency is required if we are truly to obtain precise proteome snap-shots of cellular responses to physiological changes in external environment, disease or drug treatment (Tyers & Mann 2003). Rapid advances in proteomic technology, in particular the development of fluorescence-based detection methods, have increased the capacity of this technique to quantify changes in protein expression and investigate post-translational modifications. As mentioned above, the technique does have disadvantages including the systematic exclusion of certain proteins, importantly hydrophobic membrane bound proteins, but it does provide a valid and convincing starting point to identify interesting proteins for further characterization.

Chromatography

Non-gel based methods provide an alternative strategy for proteomic studies that might potentially alleviate some of the disadvantages associated with 2DE (Liu et al 2002). Liquid chromatography coupled to mass spectrometry is a powerful technique for the analysis of peptides and proteins. Complex mixtures containing hundreds of proteins at a range of concentrations can be efficiently separated and analysed directly by tandem mass spectrometry (MS/MS) (Mann et al 2001). Liquid chromatography-MS/MS can also be combined with 1D or 2D electrophoresis, immunoprecipitation, or other protein purification techniques to isolate a subset of interesting proteins or enrich low abundance proteins to provide more directed protein identification.

Another significant improvement in classical proteome analysis is the development of multidimensional protein identification technology (MudPIT), which uses two chromatography steps one after another and can be directly coupled with the ion source of a mass spectrometer (Figure 4) (Washburn et al 2001). The digested protein complex is initially separated by a strong cationic exchange column that separates proteins with respect to their isoelectric point (Hagenstein & Sewald 2006). Proteins are eluted with increasing salt concentrations directly into a reverse-phase column which separates the proteins based on their hydrophobicity (Washburn et al 2001). The peptides then are analysed immediately by MS/MS and protein identification achieved by sequence database search (Link et al 1999; Wolters et al 2001). This technique can be used on crude samples and provides more complete coverage of the proteome of interest but does not provide any evaluation of protein abundance or activity (Burbaum & Tobal 2002).

Isotope-coded affinity tag (ICAT)

ICAT allows the relative abundance of a peptide in two samples to be identified. Samples are labelled with ICAT reagents that comprise a thiol-specific reactive group to covalently bind cysteine residues, an isotopically light or heavy linker (e.g. deuterium, ^{13}C or ^{15}N) and an affinity tag (e.g. biotin) (Adam et al 2002a; Burbaum & Tobal 2002). The presence of a heavy or light linker provides probes that have

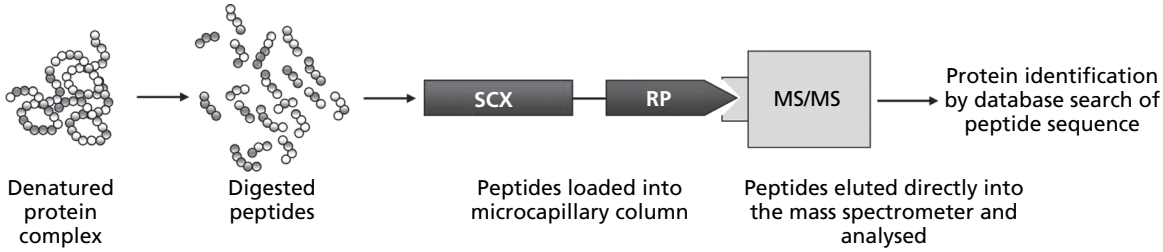

Figure 4 Multidimensional protein identification technology (MudPIT). Complex peptide mixtures are loaded onto a biphasic microcapillary column packed with strong cation exchange (SCX) and reverse-phase (RP) materials. Once loaded the column is inserted into the instrument. The peptides are separated through the columns and are eluted directly into the tandem mass spectrometer due to a voltage supply at the interface with the column. Peptide sequence data is obtained and used to achieve protein identification.

different molecular weights but are chemically identical (Gygi et al 1999; Hagenstein & Sewald 2006). Thus, two individually labelled protein samples can be combined, digested, purified by avidin affinity chromatography and subjected to multidimensional liquid chromatography and mass spectrometry (Hagenstein & Sewald 2006). Relative protein abundance can be quantified based on the intensities of the heavy and light signal (Adam et al 2002a; Burbaum & Tobal 2002). Peptides of interest can then be selected for further mass spectrometric analysis enabling protein sequence and identification to be obtained (Gygi et al 1999, 2002; Smolka et al 2002). This technology (summarized in Figure 5) provides a non-gel based method for analysing the content of complex protein mixtures and allows changes in protein abundance associated with disease or drug treatment to be evaluated. ICAT evades some of the problems of gel-based proteomic methods, providing improved access to low abundance and membrane-associated proteins (Adam et al 2002a). The major drawback of this methodology is that only cysteine-containing peptides and proteins can be analysed, thus imposing a strict limitation on the amount of quantification data that can be obtained (Smolka et al 2002; Hagenstein & Sewald 2006).

Recently, a novel method adapted from ICAT has been developed, termed isotope-coded protein label (ICPL), that is capable of high-throughput, quantitative proteome profiling on a global scale (Schmidt et al 2005). ICPL is based on tagging all free amino groups of intact proteins and can be used to label two different cell lysates with heavy and light forms before protein or peptide separation techniques (Hagenstein & Sewald 2006). This technology was shown to produce highly accurate and reproducible quantification of proteins and demonstrated high sequence coverage, thus improving the opportunity for detecting post-translational modifications and protein isoforms (Schmidt et al 2005; Sarioglu et al 2006). ICPL might prove to be a more useful technique to assess changes in protein abundance and, in turn, identify proteins implicated in the development of disease or potential therapeutic targets.

Additional methods have been developed to quantify changes in protein expression. Isobaric tags for relative and absolute quantification (iTRAQ) is a technology that employs a 4-plex set of isobaric tags (that have the same atomic mass but different arrangements) that specifically tag primary amines (N-terminus or lysine residues), thereby labelling all the peptides in a digest mixture (Ross et al 2004; Hagenstein & Sewald 2006; Wu et al 2006). Characteristic low mass reporter ions are produced for each isobaric tag upon fragmentation during MS/MS, thus enabling the comparison and quantification of up to four samples in a single experiment (Ross et al 2004; Hagenstein & Sewald 2006). iTRAQ is still a relatively new technique but is gaining in popularity as an alternative to ICAT as the technique offers more extensive proteome coverage, and improved sensitivity, and tracking of post-translational modifications is possible (Hagenstein & Sewald 2006; Wu et al 2006).

Protein microarrays

Microarray technology has proved a powerful tool for genomic research, enabling changes in gene expression to be monitored, polymorphism detection and genotyping on a genomic scale (Schena et al 1998; Lockhart & Winzeler 2000). Array-based proteomic platforms offer the potential for highly multiplexed and sensitive analysis of the proteome (Kingsmore 2006). The combination of array-based technology and phospho-specific antibodies permits the phosphorylation status of entire protein networks to be deciphered. These techniques only require small amounts of material and can be used as a high throughput tool that potentially can be used for molecular profiling, providing diagnostic and therapeutic benefits (Gulmann et al 2006). Protein microarrays can be created by a number of different methods. Tissue microarrays are arrays containing up to hundreds of tissue cores from multiple donor blocks, thus providing a method for simultaneously analysing numerous tissue samples with a specified detection mechanism, e.g. an antibody (Kononen et al 1998). Once constructed, a tissue microarray can be considered as any tissue containing wax block and can be sectioned as such, thus providing an avenue for high throughput immunohistochemical studies (Gulmann et al 2006). However, these arrays still demonstrate the same limitations that are generally associated with immunohistochemistry, such as subjectivity in the interpretation of staining and poor ability to quantify protein expression (Gulmann et al 2006).

Forward-phase protein microarrays display multiple "capture" reagents, for example a series of antibodies, on spatially defined locations (Nielsen & Geierstanger 2004). The microarrays can subsequently be probed by a complex pool of proteins, unbound proteins removed and any remaining protein interactions evaluated. Antibody microarrays are the most

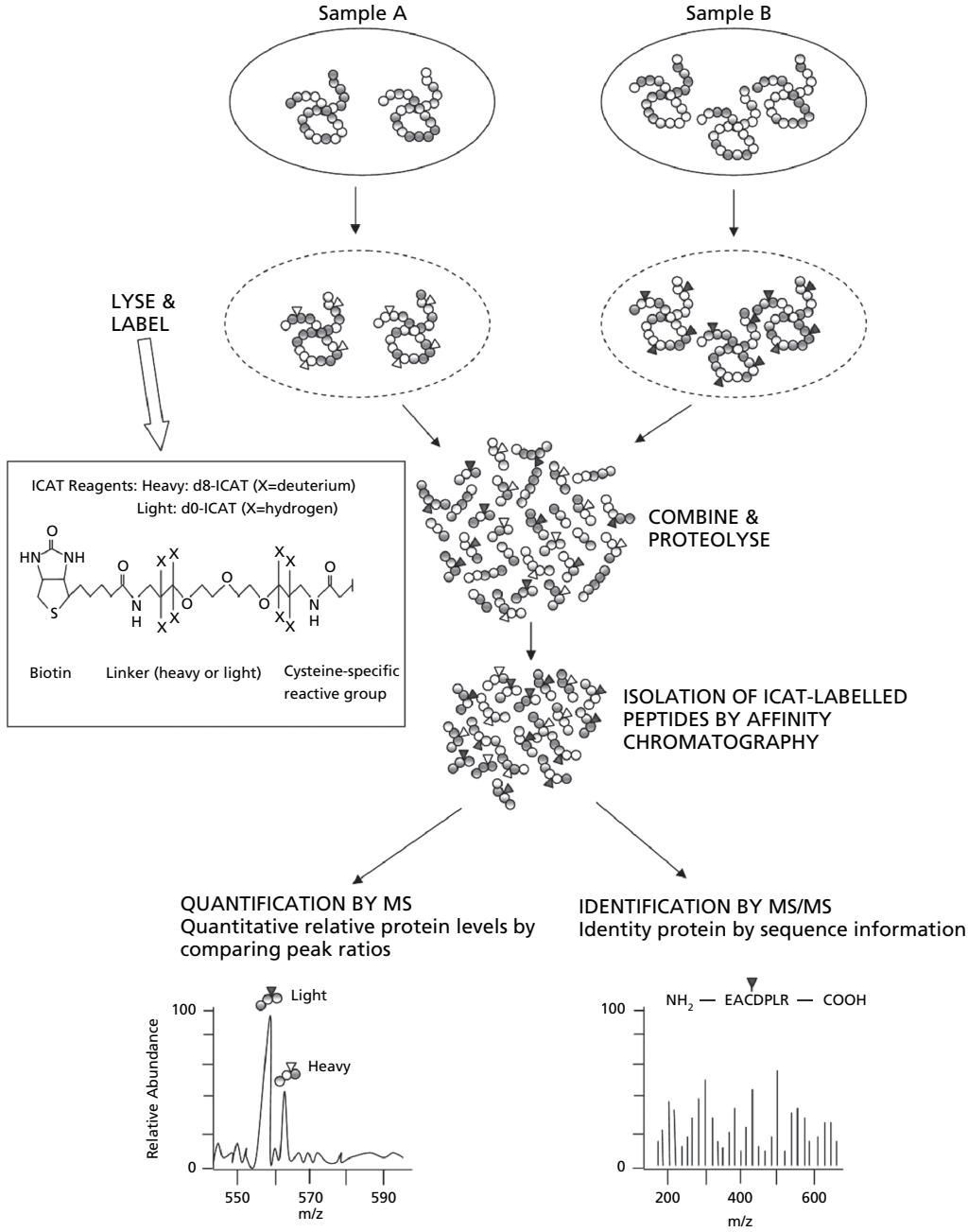

Figure 5 Schematic representation of the ICAT strategy for quantifying differential protein expression. See text for details.

common format of forward-phase microarray, but nucleic acids (Choi et al 2005), phage lysates (Zhu & Snyder 2003) and peptides (MacBeath & Schreiber 2000) have also been utilized (Gulmann et al 2006).

Evaluation of protein adherence can be achieved either by direct-labelling, in which all the proteins of a mixture are covalently labelled before incubation with the microarray, or by dual antibody sandwich assays where bound proteins are detected by a cocktail of antibodies (Haab 2003). Direct-labelling enables two protein mixtures to be labelled with different fluorophores, such as CyDyes, and co-incubated on the array permitting the evaluation of changes in protein expression. However, this method does present certain limitations, in particular that the labelling process may affect the natural binding capacities of the proteins of interest (Gulmann et al 2006). The sandwich assay offers improved sensitivity via signal amplification using biotin-conjugated antibodies, but also demonstrates significant limitations. For example, the production of two antibodies against two distinct epitopes on a single protein is difficult and the simultaneous incubation of multiple antibodies increases the risk of cross-reactivity and non-specific binding (Nielsen & Geierstanger 2004). The affinity of an antibody for its antigen also varies considerably; therefore, the amount of protein applied in a complex

mixture might be outside the concentration range required for antigen–antibody binding. Therefore, studies using forward-phase protein microarrays should incorporate serial dilutions of the analyte of interest to encompass the range of binding affinities (Gulmann et al 2006). Despite these drawbacks forward-phase protein arrays have been widely utilized and offer a potential high throughput mechanism for identifying biomarkers associated with disease, in particular cancer. This technology can be used to evaluate not only individual disease associated markers but also a combination of markers, which has been shown to be superior in the diagnosis of malignancy (Louhimo et al 2002).

Reverse-phase protein arrays have been developed also and display multiple analytes that are subsequently probed with antibodies (Paweletz et al 2001). Cellular lysates are spotted onto the array and as such several individual cell lysates, each exhibiting a complete proteome, can be examined in synchrony. This technology enables multiple samples to be directly compared as each sample is exposed to the same experimental conditions (Gulmann et al 2006). Spotting complete proteomes onto the chip does not address the problems of low abundance proteins. However, the arrays offer the advantage of being able to use amplification systems generally used for immunohistochemical studies; for example, the use of an additional staining step with a biotinylated secondary antibody to increase the signal. Drawbacks of this technique are the necessity for truly specific antibodies and experimental standardization (Nielsen & Geierstanger 2004). Currently, it is not possible to include a common reproducible reference sample on individual arrays that would enable direct comparisons between arrays and separate studies (Gulmann et al 2006).

Protein microarrays provide a powerful technology for the high throughput identification of signalling proteins and disease-associated biomarkers. In the future, antibody microarrays sensitive to the activation state of proteins might provide additional insight into cellular signalling and potentially identify key contributors to the development of disease and thus new therapeutic targets (Nielsen & Geierstanger 2004). A range of protein arrays are commercially available but not all offer full-length, folded proteins. Therefore any interactions observed require validation in a physiological environment where proteins demonstrate the correct 3D tertiary structure. Additionally, there are issues of sample standardization, storage and annotation which remain confounding factors that currently prevent the introduction of such technology into a clinical setting (Tyers & Mann 2003).

Another proteomic technology involved in quantitative analysis of protein mixtures is surface-enhanced laser desorption ionization time-of-flight (SELDI-TOF), which is a protein chip array-based chromatographic retention technology that enables direct mass spectrometric analysis of species retained on the array (Petricoin & Liotta 2004). This methodology facilitates high throughput analysis of proteins in complex biological samples such as body fluids (Kozak et al 2003; Petricoin & Liotta 2004; Schaub et al 2004; Guerreiro et al 2006). The chips display different chromatographic surfaces which are designed to retain a subset of proteins based on specific protein characteristics, namely affinity, charge, hydrophobicity and metal-binding capabilities (Guerreiro et al 2006). Additionally, chips can be coated with biological entities such as antibodies, proteins, peptides, nucleic acids, or other small molecules (Burbaum & Tobal 2002). Only a subset of the proteins bind to the surface of the chip and the unbound proteins are washed away, producing enrichment of the potentially interesting proteins (Petricoin & Liotta 2004). The adherent proteins are subsequently treated with acid so that they become ionized, dried onto the chip and can be analysed by mass spectrometry.

The advantage of this technology is the ease of use and the high throughput capabilities; additionally, post-translational modifications can be evaluated and only a relatively small amount of starting material is required (Engwegen et al 2006). Moreover, the surface-based enrichment allows low abundant and low molecular weight species to be captured and enriched (Tirumalai et al 2003; Villanueva et al 2004; Brouwers et al 2005). These proteins are often masked or eliminated with other proteomic techniques. Until recently, analysis of these protein chips was limited to protein characterization in terms of mass-to-charge ratio. However, platforms for direct "on-chip" protein sequencing have now become available, facilitating unambiguous protein identification (Caputo et al 2003; Engwegen et al 2006). Such technology holds the potential for rapid individual protein profiling and high throughput clinical screening of patients for biomarkers of disease.

Structural proteomics

Solving the structure of a protein is also required for a complete understanding of protein function, protein interactions and the formation and organization of these complexes within the cell. To achieve this, a range of structural techniques are required, including X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and computational approaches (Tyers & Mann 2003; Liu & Hsu 2005). Elucidation of protein structure allows further prediction of catalytic mechanisms, protein–protein associations and protein–nucleic acid interactions, potentially providing additional insight into the function of a given protein (Banks et al 2000). It also facilitates a more rational approach to drug design based on protein structure. Solving the structure of a protein might provide information on the nature of protein interactions, such as between a substrate and enzyme, ligand and receptor or general protein–protein interactions. This information could improve our understanding of the mechanisms involved in regulating the activity of a protein, aiding the development of more efficient and selective drugs (Greer et al 1994; O'Brien et al 2005).

X-ray crystallography is a powerful technique that is able to determine the structure of a protein. Good quality crystals of purified protein are required to collect high-resolution data for structural determination. The process of crystallization can be time consuming and expensive as a wide range of conditions must be investigated, namely changes in pH, temperature, salt, protein concentrations and the presence of cofactors, but the technique does ultimately rely on the quality of the initial protein sample being studied (Liu & Hsu 2005). The efficiency of the crystallization process has been dramatically improved by the use of automated machines that enable thousands of experiments to be carried out in parallel. Following crystallization, the protein crystal is bombarded by

an intense X-ray beam and the three-dimensional structure determined by the pattern of diffraction caused as the X-rays encounter the electrons of the atoms present in the protein (Abola et al 2000). Advances in this technology have led to an increased number of protein structures being solved, and thus potential drug targets becoming available (Abola et al 2000; Blundell et al 2002).

Nuclear magnetic resonance (NMR) also plays a significant role in the determination of protein structure (Christendat et al 2000). In contrast to X-ray crystallography, NMR does not require the formation of protein crystals to determine structure, thus it can be carried out in a relatively short amount of time. NMR analysis is carried out in aqueous solution, which allows proteins to be analysed in a more physiological environment and is more readily used to determine the structure of low molecular weight proteins, generally under 25 kDa. A concentrated protein sample (approximately 1 mM) is used to record a set of multidimensional NMR experiments, which then are processed to provide a series of spectra that are subjected to detailed and lengthy analysis to determine the 3D protein structure (Liu & Hsu 2005). There have been a number of developments aimed to: increase the signal sensitivity (thus reducing the amount of protein required); enable identification of proteins with a large molecular weight; and automate aspects of the methodology (Pervushin et al 1997; Service 1998). However, such technology still is not routinely available.

The technological advances in high-throughput X-ray crystallography over the past decade have facilitated a substantial increase in protein structure data. Unfortunately, there remains only limited structural data for multi-protein complexes and membrane proteins as deciphering the structures of these species is challenging due to the large amounts of solubilized proteins required for analysis (Liu & Hsu 2005). NMR is also a powerful tool for drug development as it is a sensitive method that can be used to detect the molecular interaction between a protein and other small molecules. NMR can be crucial in examining the interaction of a potential drug with its target, giving additional information regarding the therapeutic mechanism of action (Jung & Lee 2004).

Computational approaches in structural proteomics can also be beneficial since they are able to suggest the structures and biological functions of previously uncharacterized proteins based on structural homology to other proteins of known structure (Liu & Hsu 2005). Therefore, computational methods are able to utilize structural information obtained by X-ray crystallography and NMR to provide further novel data, thus enhancing the potential of these techniques. Currently, the techniques available for examining protein structure have benefits and limitations. Thus, to obtain structural information for proteins on a global scale these techniques must be combined and integrated to improve output (Jung & Lee 2004).

Structural proteomics provides unrivalled information that cannot be obtained by other proteomic techniques. Elucidating the structure of a protein can provide precise information regarding ligand binding sites and, as such, further our understanding of protein function and the regulation of protein activity. Computational analysis combined with structural proteomic techniques enables the function of a protein to be determined by identifying key regions within the structure that are homologous to domains of other previously characterized proteins. Protein structures also are dynamic and change rapidly over time and it is important for us to understand how differences in protein structure lead to changes in protein function, specificity and protein–ligand interactions. It is important also to identify which regions within a protein are flexible because rigid regions generally are selected as drug targets as flexibility can interfere with binding (Maggio & Ramnarayan 2001).

The rapid generation of structural data is likely to enhance significantly drug discovery by revealing potential therapeutic targets and streamlining the process of identification and optimization of new drugs (O'Brien et al 2005; Tari et al 2005). A protein structure can be used to screen large databases of real or virtual small molecules (potential therapeutic agents) to identify suitable drugs that will interact with a protein and manipulate its function. Understanding the structure of interesting proteins and closely related proteins will enable the optimization of novel chemicals with enhanced binding to the desired target protein, thus improving the potency and selectivity of a potential drug and reducing side effects (Blundell 1996; Maggio & Ramnarayan 2001). The development of highly selective drug molecules is important not only therapeutically but also to use experimentally to identify the mechanism by which proteins perform their biological functions and, particularly, how these alter with disease.

Activity-based proteomics (chemical proteomics)

Activity-based proteomics bridges the classical approach to proteomics and functional applications. The technique utilizes small chemically reactive probes, termed activity-based probes (ABP), which are directed against the active site of an enzyme and can be used to examine the functional state of enzymes in a complex proteome (Adam et al 2002a; Burbaum & Tobal 2002). These probes comprise of at least one reporter group (that enables visualization or purification) coupled via a linker to a recognition and a reactive group forming a covalent bond to the protein of interest (Figure 6) (Adam et al 2002b; Hagenstein & Sewald 2006). ABPs are designed against particular classes of enzymes and interact specifically with the active form, enabling changes in protein activity independently from changes in abundance to be evaluated (Adam et al 2001). Thus, ABPs are highly specific and provide a mechanism to directly evaluate enzyme activity of even low abundance proteins within a complex proteome (Berger et al 2004). ABPs can be used also as a means to tag, enrich and isolate distinct sets of proteins based on their enzymatic activity and have been used to profile the activity of a range of enzymes including cysteine, serine, threonine and metalloproteases and protein tyrosine phosphatases (reviewed in Berger et al (2004), Hagenstein & Sewald (2006)) (Adam et al 2002a; Elrick et al 2006).

Activity-based protein profiling offers a unique opportunity to try and further elucidate the activity of a protein or subset of proteins in a proteome. Such technology might expand our understanding of the complex protein networks present in the cell and the role of these proteins in disease, potentially allowing the identification of specific disease biomarkers. ABPs can be used also in-situ and in-vivo to localize enzyme

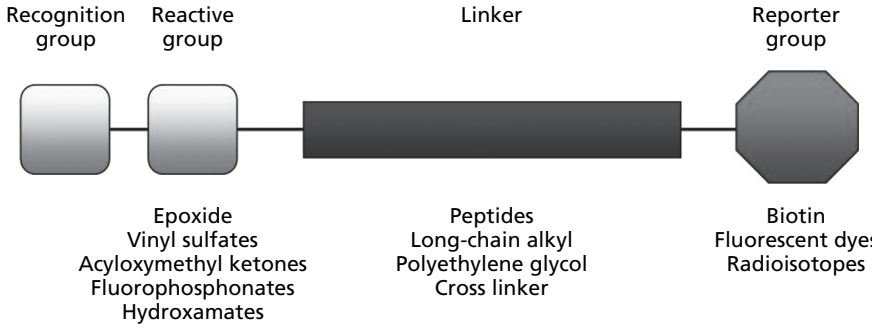

Figure 6 General structure of an activity-based probe (ABP) highlighting the probe's four main components. The key elements of an ABP are the reactive and recognition groups that form a covalent bond with an enzyme or protein in an activity-based manner. The linker region introduces space between the reactive and reporter groups to avoid steric hindrance or interference and is designed to ensure selectivity of binding. The reporter group can be a number of different tags that can be used for either visualization or purification of the protein (Hagenstein & Sewald 2006).

activity within cells and whole animals. Importantly, ABPs enable enzymes to be tagged directly, facilitating the isolation of potential novel therapeutic targets, and, in combination with an indirect competition profile, can provide information regarding selectivity and potency of a drug lead (Berger et al 2004).

Yeast two-hybrid system

The yeast two-hybrid system is one example of a technique routinely used to screen for protein–protein interactions. One protein (or set of proteins) is genetically fused to a DNA binding domain, while the other is fused to a transcription activation domain. Both hybrids are then expressed in a cell containing the reporter gene(s) which is only transcribed if the two proteins interact (Figure 7) (Phizicky & Fields 1995; Ito et al 2001; Burbaum & Tobal 2002). A variety of reporter systems have been incorporated into the yeast two-hybrid system, resulting in the production of enzymes to support growth or induce colour changes in specific substrates (Ramachandran et al 2005). This technology has been used to identify huge numbers of protein interactions in a high throughput manner, but it does also produce a number of false positives as the technique relies on the proteins of interest interacting when expressed as fusion proteins and subsequently all prospective interactions must be validated (Ito et al 2001; Burbaum & Tobal 2002; Fields 2005). False negatives are also a limitation of this technique as the methodology relies on the proteins of interest interacting in yeast cells, specifically within the nucleus, which might not be appropriate for some mammalian proteins (Ramachandran et al 2005). Mammalian two-hybrid systems have been developed to address this issue, which utilize a range of reporter mechanisms such as the stimulation of functional enzymes or the production of ubiquitin or fluorescent proteins (Stagljar et al 1998; Michnick et al 2000; Wehrman et al 2002).

Currently there is a plethora of proteomic techniques available for research. Each technology demonstrates strengths and weaknesses that need to be carefully considered when contemplating potential experimental methods, but all hold great promise to reveal extensive information regarding the function of proteins within the cell and their role in disease development. Proteomics has been utilized in many areas of research including studies to identify new diagnostic and prognostic markers for cancer, infectious diseases and

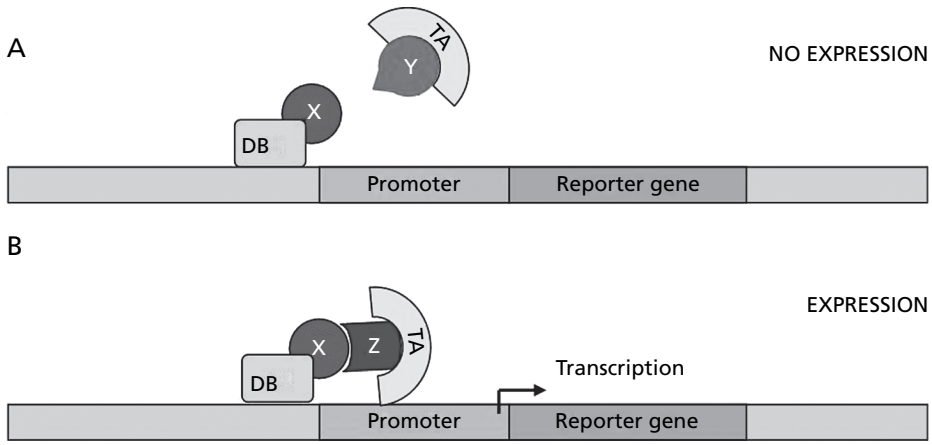

Figure 7 The principle of the yeast two-hybrid system. Chimeric proteins containing either a DNA-binding domain (DB) or a transcription activation domain (TA) are transfected into a host cell. If the two proteins do not interact (A) then there is no expression of the reporter gene. However, if the proteins of interest do interact (B) then the DB and TA are brought into proximity and can activate transcription of the reporter gene.

cardiovascular conditions, as well as the identification of novel therapeutic opportunities and investigations to improve our understanding of the underlying mechanisms of disease. Proteomics has been particularly exploited in several aspects of cancer research, for example breast (current literature review by Bertucci et al (2006) and Hudelist et al (2006)), ovarian (as reviewed by Boyce & Kohn (2005) and Tchabo et al (2005)) and prostatic cancer (summarized by Ornstein & Tyson (2006) and Semmes et al (2006)). Studies have been carried out in an attempt to improve our understanding of the aberrant cellular mechanisms involved in carcinoma development, the occurrence of chemotherapy-resistant tumours, and the identification of diagnostic and prognostic markers to further clarify cancer progression. The remainder of this review will highlight some areas of basic and clinical research that have used proteomic techniques to identify and examine potential novel therapeutic avenues for the diagnosis, prognosis and treatment of disease.

Therapeutic applications of proteomics

Advances in diagnostic and prognostic techniques generally centre on the identification of changes in protein expression associated with a condition and its progression, thus providing potential novel biomarkers that can be specifically developed to provide a clinical assay for disease. Ideally a clinically effective biomarker should be measurable in an accessible body fluid, namely serum, urine or saliva, which can be easily obtained from a patient (Petricoin & Liotta 2004). However, more readily experimental research has initially focussed on changes in protein expression in tissue biopsies isolated from individuals that are already displaying some health concerns.

Cancer

Two-dimensional electrophoresis has been used to demonstrate the pattern of protein expression of different bladder tumours (Ostergaard et al 1997). Changes in protein expression enabled six squamous-cell (epithelial cells) carcinomas to be identified from 150 transitional-cell (lining of the bladder) tumours, suggesting that this technology might facilitate accurate clinical diagnosis. The study identified psoriasin, keratins 10 and 14, psoriasis-associated fatty acid-protein (PA-FABP), galectin 7 and stratifin as being associated with the degree of cellular differentiation and psoriasin was suggested as a potential non-invasive urinary marker (Ostergaard et al 1997; Banks et al 2000). Stulik et al (1997) also used 2DE to identify the low molecular weight protein, calgranulin, as a protein expressed at increased levels in colorectal cancer and ulcerative colitis (Stulik et al 1997; Jungblut et al 1999). This protein appears to be highly specific for cancerous and precancerous cells, but its role in carcinogenesis is unclear and thus its potential to become a future biomarker remains unclear.

The protein profiles of breast carcinomas of various grades have also been studied using 2DE. Increased levels of several proteins, including proliferating cell nuclear antigen (PCNA), have been observed in invasive carcinomas (Franzen et al 1996a, b; Alaiya et al 2003). Similar studies have been carried out to define differences between benign, borderline and malignant ovarian tumours. Malignant tumours were shown to possess increased levels of many proteins including the cell cycle regulated proteins, oncoprotein 18/stathmin and PCNA (Bergman et al 2000; Alaiya et al 2003). In contrast, levels of tropomyosin 2 and lamin C decreased in comparison with levels observed in benign tumours. Analysis of a combination of nine different protein markers permitted accurate identification of benign, borderline and malignant tumours, suggesting that the simultaneous analysis of multiple polypeptides might prove a more efficient diagnostic method (Alaiya et al 1997).

Currently, the identification and examination of disease markers is based on individual proteins and this can lead to inconclusive diagnoses. For example, evaluation of prostate serum antigen (PSA) concentration is used to screen for prostate cancer, but levels of PSA can also be increased in benign conditions of the prostate (Kavallaris & Marshall 2005). A comparison of serum proteins of prostatic cancer patients with those of patients with benign prostate hyperplasia has revealed that the ratio of low molecular weight free PSA to total PSA in the serum can be used to accurately distinguish the two conditions (Charrier et al 2001; Marshall & Williams 2002). SELDI has also been used by several groups to evaluate differences in protein expression between prostatic carcinoma and benign prostatic hyperplasia. Specific seminal plasma and serum protein profiles have been identified for each condition, enabling accurate non-invasive diagnoses of cancer versus non-cancer patients (Xiao et al 2001; Cazares et al 2002; Qu et al 2002). Studies are currently underway to validate this approach for clinical diagnostics with particular relevance to inter-laboratory reproducibility (Grizzle et al 2003; Semmes et al 2005).

Cancer cells can become resistant to chemotherapy but currently the causative mechanisms responsible for drug resistance remain elusive. Proteomics can be employed to improve our understanding of drug resistance by comparing the protein profiles of resistant and non-resistant cancers and might reveal alternative therapeutic strategies. For example, 2DE and mass spectrometry have been used to identify proteins potentially important for drug resistant forms of acute lymphoblastic leukaemia (ALL). *Vinca* alkaloids, such as vincristine and vinblastine, are widely used to treat childhood and adult cancers including ALL, and the development of resistance to chemotherapy remains a clinical problem. Verrills et al (2003) identified 10 proteins that were altered with drug resistance, including cytoskeletal proteins, chaperone proteins and proteins involved in RNA processing, thus providing a number of novel insights into the mechanisms behind chemotherapy resistance and identifying possible targets for the improved treatment of relapsed disease.

Proteomics can also be employed to improve our understanding of the cellular mechanisms behind chemotherapeutic agents, for example, bohemine. Bohemine is a synthetic cyclin-dependent kinase (CDK) inhibitor and has been examined in a lymphoblastic leukaemia cell line (Kovarova et al 2000). Changes in protein expression induced by bohemine treatment have been identified by 2DE and liquid phase isoelectric focusing in combination with reverse-phase chromatography followed by mass spectrometric identification (Kovarova et al 2000; Skalnikova et al 2005). Proteins involved in glycolysis,

protein biosynthesis and reorganization of the cytoskeleton were found to be down-regulated by bohemine. This suggests that the effects of this anti-cancer drug might not only be dependent on CDK inhibition but involve other cellular protein targets and pathways (Kovarova et al 2000). Four proteins have also been shown to increase in abundance in response to bohemine, and these might represent candidate biomarkers for the positive response of cancer cells to treatment and thus a good prognosis for remission (Skalnikova et al 2005).

Infectious diseases

Proteomic techniques have been used to study a number of infectious diseases. Two-dimensional electrophoresis has been employed to identify proteins secreted into culture medium by *Mycobacterium tuberculosis* K-strain (a member of the Beijing family of bacteria prevalent in Korea) in comparison with strains H37Rv (a laboratory-adapted strain) and CDC1551 (a less prevalent strain originating in the USA) (Valway et al 1998; Bahk et al 2004; Malik & Godfrey-Faussett 2005). The Beijing strain is the most widely studied family of *M. tuberculosis* and is the dominant strain in several locations in Asia, often displaying drug resistance due to a number of different underlying mechanisms (Malik & Godfrey-Faussett 2005). Proteomic studies identified eight proteins as being more abundant in the K-strain culture medium and three further proteins were identified as being expressed at increased levels in K- and CDC1551 strains compared with the laboratory isolate. Following molecular cloning and affinity purification, two proteins, rRv3369 and rRv3874, were used to diagnose healthy and infected sera and demonstrated good sensitivity and specificity. These proteins might provide a promising avenue for developments in the serodiagnosis of tuberculosis and in the future permit the detection of preclinical infection, enabling earlier treatment and thus reduced transmission rates (Bahk et al 2004; Kavallaris & Marshall 2005).

The sera of severe acute respiratory syndrome (SARS) patients have also been evaluated using proteomic techniques (Ren et al 2004). Two-dimensional electrophoresis and tandem mass spectrometry were used to demonstrate an increase in truncated forms of α1-antitrypsin in SARS patients, and this observation subsequently was used to distinguish between infected sera and control subjects using 2D-Western blotting. α1-Antitrypsin is known to play a key role in the protection of lung function. Therefore, the presence of high levels of truncated forms of this protein might be indicative of protein degradation that might be crucial for the development of SARS. This proteomic study has not only identified a promising novel biomarker for SARS but has also provided potential mechanistic data for understanding the pathogenesis of this disease (Ren et al 2004).

Proteomic approaches have been used to identify key macromolecular interactions between human immunodeficiency virus (HIV) and host cellular proteins. Currently, anti-HIV therapy has focused on inhibiting viral enzymes and has resulted in patients being treated with a cocktail of reverse transcriptase and protease inhibitors. This drug regime is able to prolong the lifespan of the patient but currently is not curative (Tang et al 2002). Improvements in our understanding of viral–host interactions might reveal potential cellular proteins that can be targeted by drugs. Co-immunoprecipitation studies have identified an interaction between the HIV protein, Gag, and the host protein, HP68 (Zimmerman et al 2002). Gag protein is involved in driving the assembly of the HIV-1 capsid and HP68 has also been implicated in the assembly step, thus confirming that this interaction might be important for virion replication (Zimmerman et al 2002). Therefore, HP68 and any network of interacting proteins should be considered promising therapeutic targets that might inhibit HIV assembly and limit the spread of infection (Tang et al 2002). Additionally, yeast two-hybrid analysis has identified an interaction between Gag and the cellular protein, Tsg101 (Garrus et al 2001; VerPlank et al 2001). The interaction between Gag and cellular proteins also facilitate the last step of HIV-1 budding and release (Garrus et al 2001; Martin-Serrano et al 2001). Tsg101 (tumour susceptibility gene 101) is involved in the cellular trafficking of membrane-associated proteins and binding to Gag appears to be via the N-terminus of this protein (Demirov et al 2002). Inhibition of this interaction was shown to inhibit HIV-1 budding and might also provide a potential avenue for drug development to inhibit the release of viral particles (Martin-Serrano et al 2001; Goila-Gaur et al 2003).

Pocernich et al (2005) carried out a study using 2DE proteome mapping and mass spectrometry to identify proteins involved in the rapid progression of dementia that is associated with the HIV-infection of astrocytes. The HIV protein, Tat, has been implicated in the neuropathogenesis of HIV infection, demonstrating intracellular protection from oxidative-stress related toxins (Chauhan et al 2003). Proteomics was used to identify alterations in protein abundance associated with Tat expression and oxidative stress in human astrocytes in an attempt to identify proteins protected from oxidation. Several proteins showed changes in abundance and were subsequently identified. In that study, protein identification of these proteins did not directly reveal novel therapeutic targets but provided insight into the potential mechanism by which astrocytes act as a reservoir for the HIV virus leading to the development of dementia (Janssen et al 1992; Pocernich et al 2005). Improvements in our understanding of disease progression might ultimately reveal new avenues for therapeutic intervention.

Cardiovascular disease

The identification of biomarkers of cardiovascular disease, namely atherosclerosis, myocardial infarction, heart failure and stroke, might provide invaluable tools in the diagnosis, prognosis and treatment of these prevalent conditions (Vivanco et al 2005; Fu & Van Eyk 2006). Currently, the presence of the structural cardiac proteins, troponin I and troponin T in patient serum are used as markers of cardiac injury (necrosis) and are routinely used to diagnose acute myocardial infarction (McDonough & Van Eyk 2004; Babuin & Jaffe 2005). The presence of elevated levels of serum creatine kinase is also widely and effectively used to diagnose myocardial infarction (Rajappa & Sharma 2005). However, to date there are no markers available for early events such as the initial ischaemia. Moreover, chronic heart failure is monitored by the presence of elevated levels of B-type natriuretic peptide in

the blood which can also be affected by several other clinical factors (Jarai et al 2005; Wu 2005). There is also a lack of specific markers available for vascular diseases, such as atherosclerosis. Recently, C-reactive proteins, interleukin (IL)-6, CD40 ligand, IL-10 and IL-8 have been associated with atherosclerosis but, unfortunately, do not demonstrate clinically acceptable levels of sensitivity and specificity (Ridker et al 2001; Vivanco et al 2005).

Abnormal levels of low-density (LDL) and high-density lipoprotein (HDL) are known to associate with heart disease. Lipoproteins are spherical particles comprising a hydrophobic core surrounded by a coat of apolipoproteins. A comprehensive study of these proteins might offer an avenue for the identification of biomarkers for cardiovascular disease, as well as improving our understanding of the mechanisms behind these diseases. Currently, five major apolipoproteins have been identified but these are known to exist in multiple forms and are associated with a variety of proteins (Fu & Van Eyk 2006). For example, serum amyloid A (an apolipoprotein) has been found in HDL and LDL, and the serum amyloid A/LDL complex appears to reflect intravascular inflammation and might provide a sensitive marker for the prediction of prognosis in patients with coronary artery disease (Ogasawara et al 2004).

Currently, endomyocardial biopsy is the most reliable approach for detecting rejection following a heart transplant. Once again, 2DE and mass spectrometry have been used to compare the protein profiles of sequential biopsies taken from patients showing either rejection or no rejection following cardiac transplantation. Over 100 proteins were found to be upregulated with rejection and 13 of these were identified as cardiac specific proteins or heat shock proteins. Two of these proteins, namely alphaB-crystallin and tropomyosin, were further investigated in patient sera in an attempt to identify novel blood markers for rejection. Both proteins were found to be upregulated during all stages of rejection, thus confirming their potential as biomarkers of cardiac rejection and also highlighting the potential of proteomics in the identification of such prognostic markers (Borozdenkova et al 2004; Kavallaris & Marshall 2005).

Detailed proteomic studies have also been carried out to map the proteins present in human myocardium isolated from left ventricular tissue (Westbrook et al 2006). The investigation identified 388 different proteins and provides a unique protein map that can be used as a reference for other proteomic studies. The proteomic comparison of healthy myocardium with diseased heart should also result in the generation of new diagnostic and therapeutic markers (McGregor & Dunn 2006). A cohort of more than 200 hearts from patients undergoing heart transplants has recently been compiled in conjunction with 50 non-failing hearts (Dos Remedios et al 2003). This collection of tissue is to be studied by the leading experts in genomic and proteomic approaches, enabling different groups to carry out in depth studies on tissue from the same heart. Such a study promises to reveal a complete description of the molecular processes underlying a number of conditions that result in heart failure (Dos Remedios et al 2003). Several proteomic comparisons of healthy and diseased heart tissue have already been carried out, revealing cellular changes associated with disease. For example, the protein composition of membrane microdomains isolated from failing and non-failing human hearts has recently been compared, leading to the identification of 30 proteins that were differentially expressed in heart failure membrane microdomains (Banfi et al 2006). Identification of proteins whose expression is altered in heart failure provides information regarding the mechanism underlying disease development and might also reveal potential novel therapeutic targets.

Blood biomarkers

As previously discussed, an ideal biomarker for disease should be obtainable from an easily accessible body fluid. With this in mind several groups have attempted proteomic studies of plasma proteins that can be isolated from a routine blood sample. However, these investigations have proved difficult due to the wide range of protein abundances in blood plasma; proteins are thought to be present over a linear dynamic range of 10–12 orders of magnitude. Current proteomic techniques are only able to resolve proteins within 3–4 orders of magnitude, thus limiting the range of proteins that can be studied (Anderson & Anderson 2002). The low molecular range (<15 kDa) of the serum proteome has remained largely uncharacterized but is also thought to be a potential source of previously undiscovered biomarkers. Biological processes give rise to cascades of enzymatically-generated and proteolytically-clipped protein fragments that might associate with disease diagnosis or provide mechanistic insights into disease progression (Tirumalai et al 2003; Petricoin & Liotta 2004). A SELDI-based approach has recently been used successfully to identify 23 protein fragments/peptides uniquely present in the serum of cancer patients following radiotherapy (Menard et al 2006). This study identifies a series of proteins that warrant further investigation as potential biomarkers for exposure to ionizing radiation, as well as demonstrating that proteomic technology is developing to enable precise studies to be carried out on serum proteins (Menard et al 2006).

Additionally, understanding changes in post-translational modifications is essential to completely comprehend the changes in cellular signalling responsible for normal cell function and how aberrant processes lead to the development of disease. Glycosylation is one of the most common post-translational modifications found in blood plasma proteins and several of these glycosylated proteins are related to cardiovascular disease (Fu & Van Eyk 2006). One such example is the soluble form of CD40 ligand that has been shown to become elevated in unstable angina, acute myocardial infarction and atherosclerosis, and has been associated with an increased risk of cardiovascular complications, such as congestive heart failure (Vishnevetsky et al 2004). Therefore, proteomic studies focussing on a subset of blood proteins exhibiting a specific post-translational modification might prove profitable in the identification of biomarkers of disease, including cardiac disease, and the identification of potential therapeutic targets.

Proteomics has been used to study a vast range of diseases. It has proven beneficial in the identification of potential biomarkers for disease diagnosis and prognosis, the identification of potential novel drug targets and advances in our

understanding of normal cell function, disease development and progression, and the mechanisms of drug action and the development of resistance. Further advances in proteomic technology will enable the simultaneous evaluation of multiple biomarkers, potentially dramatically improving diagnostic capabilities and raising the possibility of patient-specific drug regimens. However, substantial developments are still required before such technology can be routinely and rapidly performed in the clinic and at an affordable cost, but such diagnostic procedures are likely to become a reality in the future (Pixton 2004).

Summary and current limitations of proteomics

Approximately 75% of the total cost of drug development is attributed to failures in development (Burbaum & Tobal 2002). Proteomics is a valuable tool that can be used to enhance the identification of potential therapeutic targets and to assist drug development by evaluating the cellular effects of novel therapeutic agents, enabling assessment of drug specificity and toxicity. Such thorough studies of drug–protein interactions and the cellular effects of candidate compounds can aid in the elimination of inadequate therapeutic agents before the expense of clinical trials.

Proteomic technology is developing rapidly and has been shown to be highly valuable in examining a range of clinical conditions. However, such technology holds even greater potential if the remaining limitations of these techniques can be overcome. The linear dynamic range of the proteins within the cell is still far greater than can be detected with today's proteomic technology. Therefore, proteomic studies to date have been unable to evaluate the role of low abundance proteins. Cellular fractionation and protein enrichment have been used to address this issue such that proteins are concentrated to a level where they can be evaluated. Improvements in the sensitivity of instrumentation should eliminate the need for protein enrichment. However, subcellular fractionation might prove beneficial, allowing targeted studies to analyse a specific subset of proteins. The proteins of specific cellular compartments (for example, mitochondria, nucleus, lysosomes, endoplasmic reticulum) can be isolated, reducing the complexity of the proteomic study but also enabling changes in protein localization and potentially cellular compartmentalization to be evaluated (Souchelnytskyi 2005). Currently, there are still problems with analysing complex mixtures without a degree of prior protein separation, either via a gel-based or a non-gel based method. If proteomics is to directly enter the clinical environment, routine analysis of complex protein samples, such as tissue samples or bodily fluids, is required and the procedure must be simple to carry out and be cost efficient.

Studies of cellular proteomes are further confounded by the fact that proteins can be found with a range of post-translational modifications and as a variety of splice variants. This increases the functional diversity of a single gene transcript but makes understanding the role of proteins within the cell additionally complicated. It is worth considering that a protein might play various roles within different regions of the cell and so a protein might not change in overall abundance but might be reorganized into different compartments of the cell thereby altering cell function. Moreover, a change in protein expression has no correlation with a change in protein activity. Activity-base probes are able to evaluate changes in protein activity thus permitting the evaluation of the functional properties of a protein. Activity profiles are important for drug discovery, enabling the assessment of drug specificity and investigating drug–protein interactions thus identifying the mechanisms behind drug toxicity.

Proteomic profiling is also important for the early detection of disease, namely the identification of biomarkers that can be used clinically for diagnostic and prognostic purposes. However, it is crucial to remember that the use of proteomics to identify potential biomarkers for disease is merely a starting point and must be followed by vigorous validation experiments to confirm their clinical utility. Validation should be made on a large cohort of patients taken from different populations and include an adequate number of controls, ensuring a proper evaluation of risk factor distribution. A successful biomarker must also surpass the predictive power of current existing risk factors (Mayr et al 2006). In addition to the identification of novel drug targets, proteomics can be used to verify the mode of action of a candidate target compound, confirming whether a drug is specifically targeting the candidate protein. Proteomics can be used to provide a complete overview of the changes in protein expression associated with drug treatment. This can ensure that potentially toxic compounds are eliminated during the early stages of drug development, which might prevent the significant financial losses associated with the failure of a clinical trial due to poor efficiency or adverse side effects.

Concluding remarks

It appears that the ideal proteomic approach for the identification of novel therapeutic targets might involve a combination of proteomic techniques, enabling the rapid generation and interpretation of data and advancing our understanding of the processes involved in regulating normal cell development and function. An amalgamation of several proteomic studies of a select number of samples can reveal extensive detail of the dynamic changes in protein abundance, cellular localization, post-translational modifications and protein activity (Adam et al 2002a). Currently, drug treatment is generally directed against a single target. Thorough proteomic studies can provide a complete picture of cellular signalling and can reveal the aberrant processes underlying disease. Information about entire signalling pathways can be used to develop drugs that target multiple cellular moieties within these cascades. This therapeutic strategy might potentially reduce drug toxicity as the use of a number of drugs at lower concentrations might be more effective at reducing side effects (Araujo et al 2005; Gulmann et al 2006).

In summary, proteomics has already proved valuable for the identification of novel candidate biomarkers for the diagnosis and prognosis of disease. Proteomics clearly is exceptionally powerful in the identification of aberrant protein expression associated with disease, revealing potential signalling cascades that can be targeted therapeutically. Such technology has also been used to improve our understanding of the cellular effects of current drug treatment and the development of drug resistance. There is great industrial interest in

developing and improving proteomic methodology due its broad applicability and its direct clinical potential. Thus, proteomic technology is advancing rapidly and will subsequently enhance several aspects of drug discovery, potentially aiding the production of more effective, more specific and more cost-effective drugs.

References

Abola, E., Kuhn, P., Earnest, T., Stevens, R. C. (2000) Automation of X-ray crystallography. *Nat. Struct. Biol*. **7** (Suppl.): 973–977

Adam, G. C., Cravatt, B. F., Sorensen, E. J. (2001) Profiling the specific reactivity of the proteome with non-directed activity-based probes. *Chem. Biol*. **8**: 81–95

Adam, G. C., Sorensen, E. J., Cravatt, B. F. (2002a) Chemical strategies for functional proteomics. *Mol. Cell. Proteomics* **1**: 781–790

Adam, G. C., Sorensen, E. J., Cravatt, B. F. (2002b) Trifunctional chemical probes for the consolidated detection and identification of enzyme activities from complex proteomes. *Mol. Cell. Proteomics* **1**: 828–835

Alaiya, A. A., Franzen, B., Fujioka, K., Moberger, B., Schedvins, K., Silfversvard, C., Linder, S., Auer, G. (1997) Phenotypic analysis of ovarian carcinoma: polypeptide expression in benign, borderline and malignant tumors. *Int. J. Cancer* **73**: 678–683

Alaiya, A. A., Roblick, U. J., Franzen, B., Bruch, H. P., Auer, G. (2003) Protein expression profiling in human lung, breast, bladder, renal, colorectal and ovarian cancers. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci*. **787**: 207–222

Anderson, N. G., Anderson, L. (1982) The Human Protein Index. *Clin. Chem*. **28**: 739–748

Anderson, N. L., Anderson, N. G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1**: 845–867

Araujo, R. P., Petricoin, E. F., Liotta, L. A. (2005) A mathematical model of combination therapy using the EGFR signaling network. *Biosystems* **80**: 57–69

Babuin, L., Jaffe, A. S. (2005) Troponin: the biomarker of choice for the detection of cardiac injury. *CMAJ* **173**: 1191–1202

Bahk, Y. Y., Kim, S. A., Kim, J. S., Euh, H. J., Bai, G. H., Cho, S. N., Kim, Y. S. (2004) Antigens secreted from Mycobacterium tuberculosis: identification by proteomics approach and test for diagnostic marker. *Proteomics* **4**: 3299–3307

Banfi, C., Brioschi, M., Wait, R., Begum, S., Gianazza, E., Fratto, P., Polvani, G., Vitali, E., Parolari, A., Mussoni, L., Tremoli, E. (2006) Proteomic analysis of membrane microdomains derived from both failing and non-failing human hearts. *Proteomics* **6**: 1976–1988

Banks, R. E., Dunn, M. J., Hochstrasser, D. F., Sanchez, J. C., Blackstock, W., Pappin, D. J., Selby, P. J. (2000) Proteomics: new perspectives, new biomedical opportunities. *Lancet* **356**: 1749–1756

Berger, A. B., Vitorino, P. M., Bogyo, M. (2004) Activity-based protein profiling: applications to biomarker discovery, in vivo imaging and drug discovery. *Am. J. Pharmacogenomics* **4**: 371–381

Bergman, A. C., Benjamin, T., Alaiya, A., Waltham, M., Sakaguchi, K., Franzen, B., Linder, S., Bergman, T., Auer, G., Appella, E., Wirth, P. J., Jornvall, H. (2000) Identification of gel-separated tumor marker proteins by mass spectrometry. *Electrophoresis* **21**: 679–686

Bertucci, F., Birnbaum, D., Goncalves, A. (2006) Proteomics of breast cancer: principles and potential clinical applications. *Mol. Cell. Proteomics* **5**: 1772–1786

Blundell, T. L. (1996) Structure-based drug design. *Nature* **384**: 23–26

Blundell, T. L., Jhoti, H., Abell, C. (2002) High-throughput crystallography for lead discovery in drug design. *Nat. Rev. Drug Discov*. **1**: 45–54

Borozdenkova, S., Westbrook, J. A., Patel, V., Wait, R., Bolad, I., Burke, M. M., Bell, A. D., Banner, N. R., Dunn, M. J., Rose, M. L. (2004) Use of proteomics to discover novel markers of cardiac allograft rejection. *J. Proteome Res*. **3**: 282–288

Boyce, E. A., Kohn, E. C. (2005) Ovarian cancer in the proteomics era: diagnosis, prognosis, and therapeutics targets. *Int. J. Gynecol. Cancer* **15** (Suppl. 3): 266–273

Brewis, I. A. (1999) Proteomics in reproductive research: the potential importance of proteomics to research in reproduction. *Hum. Reprod*. **14**: 2927–2929

Brouwers, F. M., Petricoin, E. F., 3rd, Ksinantova, L., Breza, J., Rajapakse, V., Ross, S., Johann, D., Mannelli, M., Shulkin, B. L., Kvetnansky, R., Eisenhofer, G., Walther, M. M., Hitt, B. A., Conrads, T. P., Veenstra, T. D., Mannion, D. P., Wall, M. R., Wolfe, G. M., Fusaro, V. A., Liotta, L. A., Pacak, K. (2005) Low molecular weight proteomic information distinguishes metastatic from benign pheochromocytoma. *Endocr. Relat. Cancer* **12**: 263–272

Burbaum, J., Tobal, G. M. (2002) Proteomics in drug discovery. *Curr. Opin. Chem. Biol*. **6**: 427–433

Caputo, E., Moharram, R., Martin, B. M. (2003) Methods for on-chip protein analysis. *Anal. Biochem*. **321**: 116–124

Cazares, L. H., Adam, B. L., Ward, M. D., Nasim, S., Schellhammer, P. F., Semmes, O. J., Wright, G. L., Jr. (2002) Normal, benign, preneoplastic, and malignant prostate cells have distinct protein expression profiles resolved by surface enhanced laser desorption/ionization mass spectrometry. *Clin. Cancer Res*. **8**: 2541–2552

Charrier, J. P., Tournel, C., Michel, S., Comby, S., Jolivet-Reynaud, C., Passagot, J., Dalbon, P., Chautard, D., Jolivet, M. (2001) Differential diagnosis of prostate cancer and benign prostate hyperplasia using two-dimensional electrophoresis. *Electrophoresis* **22**: 1861–1866

Chauhan, A., Turchan, J., Pocernich, C., Bruce-Keller, A., Roth, S., Butterfield, D. A., Major, E. O., Nath, A. (2003) Intracellular human immunodeficiency virus Tat expression in astrocytes promotes astrocyte survival but induces potent neurotoxicity at distant sites via axonal transport. *J. Biol. Chem*. **278**: 13512–13519

Choi, Y. S., Pack, S. P., Yoo, Y. J. (2005) Development of a protein microarray using sequence-specific DNA binding domain on DNA chip surface. *Biochem. Biophys. Res. Commun*. **329**: 1315–1319

Christendat, D., Yee, A., Dharamsi, A., Kluger, Y., Savchenko, A., Cort, J. R., Booth, V., Mackereth, C. D., Saridakis, V., Ekiel, I., Kozlov, G., Maxwell, K. L., Wu, N., McIntosh, L. P., Gehring, K., Kennedy, M. A., Davidson, A. R., Pai, E. F., Gerstein, M., Edwards, A. M., Arrowsmith, C. H. (2000) Structural proteomics of an archaeon. *Nat. Struct. Biol*. **7**: 903–909

Cutler, P., Heald, G., White, I. R., Ruan, J. (2003) A novel approach to spot detection for two-dimensional gel electrophoresis images using pixel value collection. *Proteomics* **3**: 392–401

Demirov, D. G., Ono, A., Orenstein, J. M., Freed, E. O. (2002) Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. *Proc. Natl. Acad. Sci. U S A* **99**: 955–960

Dos Remedios, C. G., Liew, C. C., Allen, P. D., Winslow, R. L., Van Eyk, J. E., Dunn, M. J. (2003) Genomics, proteomics and bioinformatics of human heart failure. *J. Muscle Res. Cell Motil*. **24**: 251–260

Drews, J. (2000) Drug discovery: a historical perspective. *Science* **287**: 1960–1964

Edwards, A. M., Arrowsmith, C. H.; des Pallieres, B. (2000) Proteomics: New tools for a new era. *Modern Drug Discovery* **3**: 34–46

Elrick, M. M., Walgren, J. L., Mitchell, M. D., Thompson, D. C. (2006) Proteomics: recent applications and new technologies. *Basic Clin. Pharmacol. Toxicol*. **98**: 432–441

Engwegen, J. Y., Gast, M. C., Schellens, J. H., Beijnen, J. H. (2006) Clinical proteomics: searching for better tumour markers with SELDI-TOF mass spectrometry. *Trends Pharmacol. Sci*. **5**: 251–259

Fields, S. (2005) High-throughput two-hybrid analysis. The promise and the peril. *FEBS J*. **272**: 5391–5399

Franzen, B., Linder, S., Alaiya, A. A., Eriksson, E., Uruy, K., Hirano, T., Okuzawa, K., Auer, G. (1996a) Analysis of polypeptide expression in benign and malignant human breast lesions: down-regulation of cytokeratins. *Br. J. Cancer* **74**: 1632–1638

Franzen, B., Auer, G., Alaiya, A. A., Eriksson, E., Uryu, K., Hirano, T., Okuzawa, K., Kato, H., Linder, S. (1996b) Assessment of homogeneity in polypeptide expression in breast carcinomas shows widely variable expression in highly malignant tumors. *Int. J. Cancer* **69**: 408–414

Fu, Q., Van Eyk, J. E. (2006) Proteomics and heart disease: identifying biomarkers of clinical utility. *Expert Rev. Proteomics* **3**: 237–249

Garrus, J. E., von Schwedler, U. K., Pornillos, O. W., Morham, S. G., Zavitz, K. H., Wang, H. E., Wettstein, D. A., Stray, K. M., Cote, M., Rich, R. L., Myszka, D. G., Sundquist, W. I. (2001) Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **107**: 55–65

Gharbi, S., Gaffney, P., Yang, A., Zvelebil, M. J., Cramer, R., Waterfield, M. D., Timms, J. F. (2002) Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system. *Mol. Cell. Proteomics* **1**: 91–98

Goila-Gaur, R., Demirov, D. G., Orenstein, J. M., Ono, A., Freed, E. O. (2003) Defects in human immunodeficiency virus budding and endosomal sorting induced by TSG101 overexpression. *J. Virol*. **77**: 6507–6519

Gorg, A., Postel, W., Gunther, S. (1988) The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **9**: 531–546

Greer, J., Erickson, J. W., Baldwin, J. J., Varney, M. D. (1994) Application of the three-dimensional structures of protein target molecules in structure-based drug design. *J. Med. Chem*. **37**: 1035–1054

Grizzle, W. E., Adam, B. L., Bigbee, W. L., Conrads, T. P., Carroll, C., Feng, Z., Izbicka, E., Jendoubi, M., Johnsey, D., Kagan, J., Leach, R. J., McCarthy, D. B., Semmes, O. J., Srivastava, S., Srivastava, S., Thompson, I. M., Thornquist, M. D., Verma, M., Zhang, Z., Zou, Z. (2003) Serum protein expression profiling for cancer detection: validation of a SELDI-based approach for prostate cancer. *Dis. Markers* **19**: 185–195

Guerreiro, N., Gomez-Mancilla, B., Charmont, S. (2006) Optimization and evaluation of surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry for protein profiling of cerebrospinal fluid. *Proteome Sci*. **4**: 7

Gulmann, C., Sheehan, K. M., Kay, E. W., Liotta, L. A., Petricoin, E. F. (2006) Array-based proteomics: mapping of protein circuitries for diagnostics, prognostics, and therapy guidance in cancer. *J. Pathol*. **208**: 595–606

Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol*. **17**: 994–999

Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., Aebersold, R. (2000) Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. U S A* **97**: 9390–9395

Gygi, S. P., Rist, B., Griffin, T. J., Eng, J., Aebersold, R. (2002) Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags. *J. Proteome Res*. **1**: 47–54

Haab, B. B. (2003) Methods and applications of antibody microarrays in cancer research. *Proteomics* **3**: 2116–2122

Hagenstein, M. C., Sewald, N. (2006) Chemical tools for activity-based proteomics. *J. Biotechnol*. **124**: 56–73

Harper, R. G., Workman, S. R., Schuetzner, S., Timperman, A. T., Sutton, J. N. (2004) Low-molecular-weight human serum proteome using ultrafiltration, isoelectric focusing, and mass spectrometry. *Electrophoresis* **25**: 1299–1306

Harry, J. L., Wilkins, M. R., Herbert, B. R., Packer, N. H., Gooley, A. A., Williams, K. L. (2000) Proteomics: capacity versus utility. *Electrophoresis* **21**: 1071–1081

Hudelist, G., Singer, C. F., Pischinger, K. I., Kaserer, K., Manavi, M., Kubista, E., Czerwenka, K. F. (2006) Proteomic analysis in human breast cancer: identification of a characteristic protein expression profile of malignant breast epithelium. *Proteomics* **6**: 1989–2002

International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* **431**: 931–945

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. U S A* **98**: 4569–4574

Janssen, R. S., Nwanyanwu, O. C., Selik, R. M., Stehr-Green, J. K. (1992) Epidemiology of human immunodeficiency virus encephalopathy in the United States. *Neurology* **42**: 1472–1476

Jarai, R., Wojta, J., Huber, K. (2005) Circulating B-type natriuretic peptides in patients with acute coronary syndromes. Pathophysiological, prognostical and therapeutical considerations. *Thromb. Haemost*. **94**: 926–932

Jung, J. W., Lee, W. (2004) Structure-based functional discovery of proteins: structural proteomics. *J. Biochem. Mol. Biol*. **37**: 28–34

Jungblut, P. R., Zimny-Arndt, U., Zeindl-Eberhart, E., Stulik, J., Koupilova, K., Pleissner, K. P., Otto, A., Muller, E. C., Sokolowska-Kohler, W., Grabher, G., Stoffler, G. (1999) Proteomics in human disease: cancer, heart and infectious diseases. *Electrophoresis* **20**: 2100–2110

Kavallaris, M., Marshall, G. M. (2005) Proteomics and disease: opportunities and challenges. *Med. J. Aust*. **182**: 575–579

Kingsmore, S. F. (2006) Multiplexed protein measurement: technologies and applications of protein and antibody arrays. *Nat. Rev. Drug Discov*. **5**: 310–320

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

Kovarova, H., Hajduch, M., Korinkova, G., Halada, P., Krupickova, S., Gouldsworthy, A., Zhelev, N., Strnad, M. (2000) Proteomics approach in classifying the biochemical basis of the anticancer activity of the new olomoucine-derived synthetic cyclin-dependent kinase inhibitor, bohemine. *Electrophoresis* **21**: 3757–3764

Kozak, K. R., Amneus, M. W., Pusey, S. M., Su, F., Luong, M. N., Luong, S. A., Reddy, S. T., Farias-Eisner, R. (2003) Identification of biomarkers for ovarian cancer using strong anion-exchange ProteinChips: potential use in diagnosis and prognosis. *Proc. Natl. Acad. Sci. U S A* **100**: 12343–12348

Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., Yates, J. R. (1999) Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol*. **17**: 676–682

Liu, H., Lin, D., Yates, J. R. (2002) Multidimensional separations for protein/peptide analysis in the post-genomic era. *Biotechniques* **32**: 898, 900, 902 passim

Liu, H. L., Hsu, J. P. (2005) Recent developments in structural proteomics for protein structure determination. *Proteomics* **5**: 2056–2068

Lockhart, D. J., Winzeler, E. A. (2000) Genomics, gene expression and DNA arrays. *Nature* **405**: 827–836

Louhimo, J., Finne, P., Alfthan, H., Stenman, U. H., Haglund, C. (2002) Combination of HCGbeta, CA 19–9 and CEA with logistic regression improves accuracy in gastrointestinal malignancies. *Anticancer Res*. **22**: 1759–1764

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

Mackintosh, J. A., Choi, H. Y., Bae, S. H., Veal, D. A., Bell, P. J., Ferrari, B. C., Van Dyk, D. D., Verrills, N. M., Paik, Y. K., Karuso, P. (2003) A fluorescent natural product for ultra sensitive detection of proteins in one-dimensional and two-dimensional gel electrophoresis. *Proteomics* **3**: 2273–2288

Maggio, E. T., Ramnarayan, K. (2001) Recent developments in computational proteomics. *Trends Biotechnol*. **19**: 266–272

Mahon, P., Dupree, P. (2001) Quantitative and reproducible two-dimensional gel analysis using Phoretix 2D Full. *Electrophoresis* **22**: 2075–2085

Malik, A. N., Godfrey-Faussett, P. (2005) Effects of genetic variability of Mycobacterium tuberculosis strains on the presentation of disease. *Lancet Infect. Dis*. **5**: 174–183

Mann, M., Hendrickson, R. C., Pandey, A. (2001) Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem*. **70**: 437–473

Marouga, R., David, S., Hawkins, E. (2005) The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Anal. Bioanal. Chem*. **382**: 669–678

Marshall, T., Williams, K. M. (2002) Proteomics and its impact upon biomedical science. *Br. J. Biomed. Sci*. **59**: 47–64

Martin-Serrano, J., Zang, T., Bieniasz, P. D. (2001) HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. *Nat. Med*. **7**: 1313–1319

Mayr, M., Zhang, J., Greene, A. S., Gutterman, D. D., Perloff, J. K., Ping, P. (2006) Proteomic based development of biomarkers in cardiovascular disease: Mechanistic, clinical, and therapeutic insights. *Mol. Cell. Proteomics* **5**: 1853–1864

McDonough, J. L., Van Eyk, J. E. (2004) Developing the next generation of cardiac markers: disease-induced modifications of troponin I. *Prog. Cardiovasc. Dis*. **47**: 207–216

McGregor, E., Dunn, M. J. (2006) Proteomics of the heart: unraveling disease. *Circ. Res*. **98**: 309–321

Menard, C., Johann, D., Lowenthal, M., Muanza, T., Sproull, M., Ross, S., Gulley, J., Petricoin, E., Coleman, C. N., Whiteley, G., Liotta, L., Camphausen, K. (2006) Discovering clinical biomarkers of ionizing radiation exposure with serum proteomic analysis. *Cancer Res*. **66**: 1844–1850

Michnick, S. W., Remy, I., Campbell-Valois, F. X., Vallee-Belisle, A., Pelletier, J. N. (2000) Detection of protein-protein interactions by protein fragment complementation strategies. *Methods Enzymol*. **328**: 208–230

Moritz, R. L., Ji, H., Schutz, F., Connolly, L. M., Kapp, E. A., Speed, T. P., Simpson, R. J. (2004) A proteome strategy for fractionating proteins and peptides using continuous free-flow electrophoresis coupled off-line to reversed-phase high-performance liquid chromatography. *Anal. Chem*. **76**: 4811–4824

Nielsen, U. B., Geierstanger, B. H. (2004) Multiplexed sandwich assays in microarray format. *J. Immunol. Methods* **290**: 107–120

O'Brien, S. E., Brown, D. G., Mills, J. E., Phillips, C., Morris, G. (2005) Computational tools for the analysis and visualization of multiple protein-ligand complexes. *J. Mol. Graph. Model* **24**: 186–194

Ogasawara, K., Mashiba, S., Wada, Y., Sahara, M., Uchida, K., Aizawa, T., Kodama, T. (2004) A serum amyloid A and LDL complex as a new prognostic marker in stable coronary artery disease. *Atherosclerosis* **174**: 349–356

Ornstein, D. K., Tyson, D. R. (2006) Proteomics for the identification of new prostate cancer biomarkers. *Urol. Oncol*. **24**: 231–236

Ostergaard, M., Rasmussen, H. H., Nielsen, H. V., Vorum, H., Orntoft, T. F., Wolf, H., Celis, J. E. (1997) Proteome profiling of bladder squamous cell carcinomas: identification of markers that define their degree of differentiation. *Cancer Res*. **57**: 4111–4117

Pandey, A., Mann, M. (2000) Proteomics to study genes and genomes. *Nature* **405**: 837–846

Paweletz, C. P., Charboneau, L., Bichsel, V. E., Simone, N. L., Chen, T., Gillespie, J. W., Emmert-Buck, M. R., Roth, M. J., Petricoin, I. E., Liotta, L. A. (2001) Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* **20**: 1981–1989

Pervushin, K., Riek, R., Wider, G., Wuthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U S A* **94**: 12366–12371

Petricoin, E. F., Liotta, L. A. (2004) SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr. Opin. Biotechnol*. **15**: 24–30

Phizicky, E. M., Fields, S. (1995) Protein-protein interactions: methods for detection and analysis. *Microbiol. Rev*. **59**: 94–123

Pixton, K. L. (2004) Is it functional? Report on the first BSPR/EBI meeting on functional proteomics. *Proteomics* **4**: 3762–3764

Pocernich, C. B., Boyd-Kimball, D., Poon, H. F., Thongboonkerd, V., Lynn, B. C., Klein, J. B., Calebrese, V., Nath, A., Butterfield, D. A. (2005) Proteomics analysis of human astrocytes expressing the HIV protein Tat. *Brain Res. Mol. Brain Res*. **133**: 307–316

Qu, Y., Adam, B. L., Yasui, Y., Ward, M. D., Cazares, L. H., Schellhammer, P. F., Feng, Z., Semmes, O. J., Wright, G. L. (2002) Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients. *Clin. Chem*. **48**: 1835–1843

Rajappa, M., Sharma, A. (2005) Biomarkers of cardiac injury: an update. *Angiology* **56**: 677–691

Ramachandran, N., Larson, D. N., Stark, P. R., Hainsworth, E., LaBaer, J. (2005) Emerging tools for real-time label-free detection of interactions on functional protein microarrays. *FEBS J*. **272**: 5412–5425

Ren, Y., He, Q. Y., Fan, J., Jones, B., Zhou, Y., Xie, Y., Cheung, C. Y., Wu, A., Chiu, J. F., Peiris, J. S., Tam, P. K. (2004) The use of proteomics in the discovery of serum biomarkers from patients with severe acute respiratory syndrome. *Proteomics* **4**: 3477–3484

Ridker, P. M., Stampfer, M. J., Rifai, N. (2001) Novel risk factors for systemic atherosclerosis: a comparison of C-reactive protein, fibrinogen, homocysteine, lipoprotein(a), and standard cholesterol screening as predictors of peripheral arterial disease. *JAMA* **285**: 2481–2485

Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., Pappin, D. J. (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3**: 1154–1169

Santoni, V., Molloy, M., Rabilloud, T. (2000) Membrane proteins and proteomics: un amour impossible? *Electrophoresis* **21**: 1054–1070

Sarioglu, H., Brandner, S., Jacobsen, C., Meindl, T., Schmidt, A., Kellermann, J., Lottspeich, F., Andrae, U. (2006) Quantitative analysis of 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced proteome alterations in 5L rat hepatoma cells using isotope-coded protein labels. *Proteomics* **6**: 2407–2421

Schaub, S., Rush, D., Wilkins, J., Gibson, I. W., Weiler, T., Sangster, K., Nicolle, L., Karpinski, M., Jeffery, J., Nickerson, P. (2004)

Proteomic-based detection of urine proteins associated with acute renal allograft rejection. *J. Am. Soc. Nephrol*. **15**: 219–227

Schena, M., Heller, R. A., Theriault, T. P., Konrad, K., Lachenmeier, E., Davis, R. W. (1998) Microarrays: biotechnology's discovery platform for functional genomics. *Trends Biotechnol.* **16**: 301–306

Schmidt, A., Kellermann, J., Lottspeich, F. (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5**: 4–15

Semmes, O. J., Feng, Z., Adam, B. L., Banez, L. L., Bigbee, W. L., Campos, D., Cazares, L. H., Chan, D. W., Grizzle, W. E., Izbicka, E., Kagan, J., Malik, G., McLerran, D., Moul, J. W., Partin, A., Prasanna, P., Rosenzweig, J., Sokoll, L. J., Srivastava, S., Srivastava, S., Thompson, I., Welsh, M. J., White, N., Winget, M., Yasui, Y., Zhang, Z., Zhu, L. (2005) Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. *Clin. Chem*. **51**: 102–112

Semmes, O. J., Malik, G., Ward, M. (2006) Application of mass spectrometry to the discovery of biomarkers for detection of prostate cancer. *J. Cell Biochem*. **98**: 496–503

Service, R. F. (1998) NMR researchers look to the next generation of machines. *Science* **279**: 1127–1128

Skalnikova, H., Halada, P., Dzubak, P., Hajduch, M., Kovarova, H. (2005) Protein fingerprints of anti-cancer effects of cyclin-dependent kinase inhibition: identification of candidate biomarkers using 2-D liquid phase separation coupled to mass spectrometry. *Technol. Cancer Res. Treat*. **4**: 447–454

Smolka, M., Zhou, H., Aebersold, R. (2002) Quantitative protein profiling using two-dimensional gel electrophoresis, isotope-coded affinity tag labeling, and mass spectrometry. *Mol. Cell. Proteomics* **1**: 19–29

Souchelnytskyi, S. (2005) Bridging proteomics and systems biology: what are the roads to be traveled? *Proteomics* **5**: 4123–4137

Stagljar, I., Korostensky, C., Johnsson, N., te Heesen, S. (1998) A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proc. Natl. Acad. Sci. U S A* **95**: 5187–5192

Stulik, J., Kovarova, H., Macela, A., Bures, J., Jandik, P., Langr, F., Otto, A., Thiede, B., Jungblut, P. (1997) Overexpression of calcium-binding protein calgranulin B in colonic mucosal diseases. *Clin. Chim. Acta* **265**: 41–55

Tang, H., Peng, T., Wong-Staal, F. (2002) Novel technologies for studying virus-host interaction and discovering new drug targets for HCV and HIV. *Curr. Opin. Pharmacol*. **2**: 541–547

Tari, L. W., Rosenberg, M., Schryvers, A. B. (2005) Structural proteomics in drug discovery. *Expert Rev. Proteomics* **2**: 511–519

Tchabo, N. E., Liel, M. S., Kohn, E. C. (2005) Applying proteomics in clinical trials: assessing the potential and practical limitations in ovarian cancer. *Am. J. Pharmacogenomics* **5**: 141–148

Tirumalai, R. S., Chan, K. C., Prieto, D. A., Issaq, H. J., Conrads, T. P., Veenstra, T. D. (2003) Characterization of the low molecular weight human serum proteome. *Mol Cell Proteomics* **2**: 1096–1103

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

Unlu, M., Morgan, M. E., Minden, J. S. (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18**: 2071–2077

Valway, S. E., Sanchez, M. P., Shinnick, T. F., Orme, I., Agerton, T., Hoy, D., Jones, J. S., Westmoreland, H., Onorato, I. M. (1998) An outbreak involving extensive transmission of a virulent strain of Mycobacterium tuberculosis. *N. Engl. J. Med*. **338**: 633–639

VerPlank, L., Bouamr, F., LaGrassa, T. J., Agresta, B., Kikonyogo, A., Leis, J., Carter, C. A. (2001) Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). *Proc. Natl. Acad. Sci. U S A* **98**: 7724–7729

Verrills, N. M., Walsh, B. J., Cobon, G. S., Hains, P. G., Kavallaris, M. (2003) Proteome analysis of vinca alkaloid response and resistance in acute lymphoblastic leukemia reveals novel cytoskeletal alterations. *J. Biol. Chem*. **278**: 45082–45093

Villanueva, J., Philip, J., Entenberg, D., Chaparro, C. A., Tanwar, M. K., Holland, E. C., Tempst, P. (2004) Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal. Chem*. **76**: 1560–1570

Vishnevetsky, D., Kiyanista, V. A., Gandhi, P. J. (2004) CD40 ligand: a novel target in the fight against cardiovascular disease. *Ann. Pharmacother*. **38**: 1500–1508

Vivanco, F., Martin-Ventura, J. L., Duran, M. C., Barderas, M. G., Blanco-Colio, L., Darde, V. M., Mas, S., Meilhac, O., Michel, J. B., Tunon, J., Egido, J. (2005) Quest for novel cardiovascular biomarkers by proteomic analysis. *J. Proteome Res*. **4**: 1181–1191

Wang, M. Z., Howard, B., Campa, M. J., Patz, E. F., Fitzgerald, M. C. (2003) Analysis of human serum proteins by liquid phase isoelectric focusing and matrix-assisted laser desorption/ionization-mass spectrometry. *Proteomics* **3**: 1661–1666

Washburn, M. P., Wolters, D., Yates, J. R. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol*. **19**: 242–247

Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R., Duncan, M. W., Harris, R., Williams, K. L., Humphery-Smith, I. (1995) Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. *Electrophoresis* **16**: 1090–1094

Wehrman, T., Kleaveland, B., Her, J. H., Balint, R. F., Blau, H. M. (2002) Protein-protein interactions monitored in mammalian cells via complementation of beta-lactamase enzyme fragments. *Proc. Natl. Acad. Sci. U S A* **99**: 3469–3474

Westbrook, J. A., Wheeler, J. X., Wait, R., Welson, S. Y., Dunn, M. J. (2006) The human heart proteome: Two-dimensional maps using narrow-range immobilised pH gradients. *Electrophoresis* **27**: 1547–1555

White, I. R., Pickford, R., Wood, J., Skehel, J. M., Gangadharan, B., Cutler, P. (2004) A statistical comparison of silver and SYPRO Ruby staining for proteomic analysis. *Electrophoresis* **25**: 3048–3054

Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., Williams, K. L. (1996a) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev*. **13**: 19–50

Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., Sanchez, J. C., Yan, J. X., Gooley, A. A., Hughes, G., Humphery-Smith, I., Williams, K. L., Hochstrasser, D. F. (1996b) From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y)* **14**: 61–65

Wolters, D. A., Washburn, M. P., Yates, J. R. (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem*. **73**: 5683–5690

Wu, A. H. (2005) Markers for early detection of cardiac diseases. *Scand. J. Clin. Lab. Invest. Suppl*. **240**: 112–121

Wu, W. W., Wang, G., Baek, S. J., Shen, R. F. (2006) Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. *J. Proteome Res*. **5**: 651–658

Xiao, Z., Adam, B. L., Cazares, L. H., Clements, M. A., Davis, J. W., Schellhammer, P. F., Dalmasso, E. A., Wright, G. L., Jr. (2001) Quantitation of serum prostate-specific membrane antigen by a novel protein biochip immunoassay discriminates benign from malignant prostate disease. *Cancer Res*. **61**: 6029–6033

Xiao, Z., Conrads, T. P., Lucas, D. A., Janini, G. M., Schaefer, C. F., Buetow, K. H., Issaq, H. J., Veenstra, T. D. (2004) Direct ampholyte-free liquid-phase isoelectric peptide focusing: application to the human serum proteome. *Electrophoresis* **25**: 128–133

Zhu, H., Snyder, M. (2003) Protein chip technology. *Curr. Opin. Chem. Biol*. **7**: 55–63

Zimmerman, C., Klein, K. C., Kiser, P. K., Singh, A. R., Firestein, B. L., Riba, S. C., Lingappa, J. R. (2002) Identification of a host protein essential for assembly of immature HIV-1 capsids. *Nature* **415**: 88–92