

- 69 Kato, Y. *et al.* (1998) Bmk1/Erk5 is required for cell proliferation induced by epidermal growth factor. *Nature* 395, 713–716
- 70 Fukuhara, S. *et al.* (2000) Signaling from G protein-coupled receptors to ERK5/Big MAPK 1 involves G $\alpha_q$  and G $\alpha_{12/13}$  families of heterotrimeric G proteins. Evidence for the existence of a novel Ras and Rho-independent pathway. *J. Biol. Chem.* 275, 21730–21736
- 71 Mao, Z. *et al.* (1999) Neuronal activity-dependent cell survival mediated by transcription factor MEF2. *Science* 286, 785–790
- 72 Lu, J. *et al.* (2000) Regulation of skeletal myogenesis by association of the MEF2 transcription factor with class II histone deacetylases. *Mol. Cell* 6, 233–244
- 73 van Biesen, T. *et al.* (1996) Mitogenic signaling via G protein-coupled receptors. *Endocrinol. Rev.* 17, 698–714
- 74 Bockaert, J. and Pin, J.P. (1999) Molecular tinkering of G protein-coupled receptors: an evolutionary success. *EMBO J.* 18, 1723–1729
- 75 Pawson, T. and Scott, J.D. (1997) Signaling through scaffold, anchoring, and adaptor proteins. *Science* 278, 2075–2080
- 76 Hall, R.A. *et al.* (1999) Heptahelical receptor signaling: beyond the G protein paradigm. *J. Cell Biol.* 145, 927–932

# Proteomics – post-genomic cartography to understand gene function

Soren Naaby-Hansen, Michael D. Waterfield and Rainer Cramer

The completion of the genomic sequences of numerous organisms from human and mouse to *Caenorhabditis elegans* and many microorganisms, and the definition of their genes provides a database to interpret cellular protein-expression patterns and relate them to protein function. Proteomics technologies that are dependent on mass spectrometry and involve two-dimensional gel electrophoresis are providing the main window into the world of differential protein-expression analysis. In this article, the limitations and expectations of this research field are examined and the future of the analytical needs of proteomics is explored.

Biomedical research in the post-genomic era will be carried out against the backdrop of an information explosion about the ~30 000 genes whose sequence have been completed for mouse and humans. It is clear that high-throughput transcriptional profiling with chip technology will easily identify when, where and how much RNA is transcribed. However, there is usually no clear correlation between RNA transcription and protein expression<sup>1,2</sup> and transcriptomics reveals no data regarding the activity of the product, leaving a huge information gulf about the regulation of protein expression, structure and function, which must be filled to allow fast-track exploitation of genomics.

Proteomics is the large-scale study of gene expression at the protein level, which will ultimately provide direct measurement of protein expression levels and insight into the activity state of all relevant proteins. In an experimental context the term proteomics is derived from proteome, which by analogy to genome is defined as the entire protein complement expressed by a cell or tissue type. Study of the proteome dates back to the late 1970s. The key elements of classical proteomics are the separation of proteins in a sample using two-dimensional gel electrophoresis (2-DE) and their subsequent quantitation and identification. Protein expression

levels, however, might not fully reflect their functional state, and protein expression profiling *per se* provides no data regarding subcellular site of action or interaction partners unless additional experimental steps are included. The scope of modern proteomics has therefore broadened considerably (Box 1).

## The tools of proteomics

The dynamic range of protein expression and modification makes the identification of the entire proteome a far bigger and more complex challenge than the sequencing of the genome. Although gene sequencing and expression analysis can be performed with high throughput and in an automated manner, major technical problems need to be resolved and new techniques developed before proteomics can become a

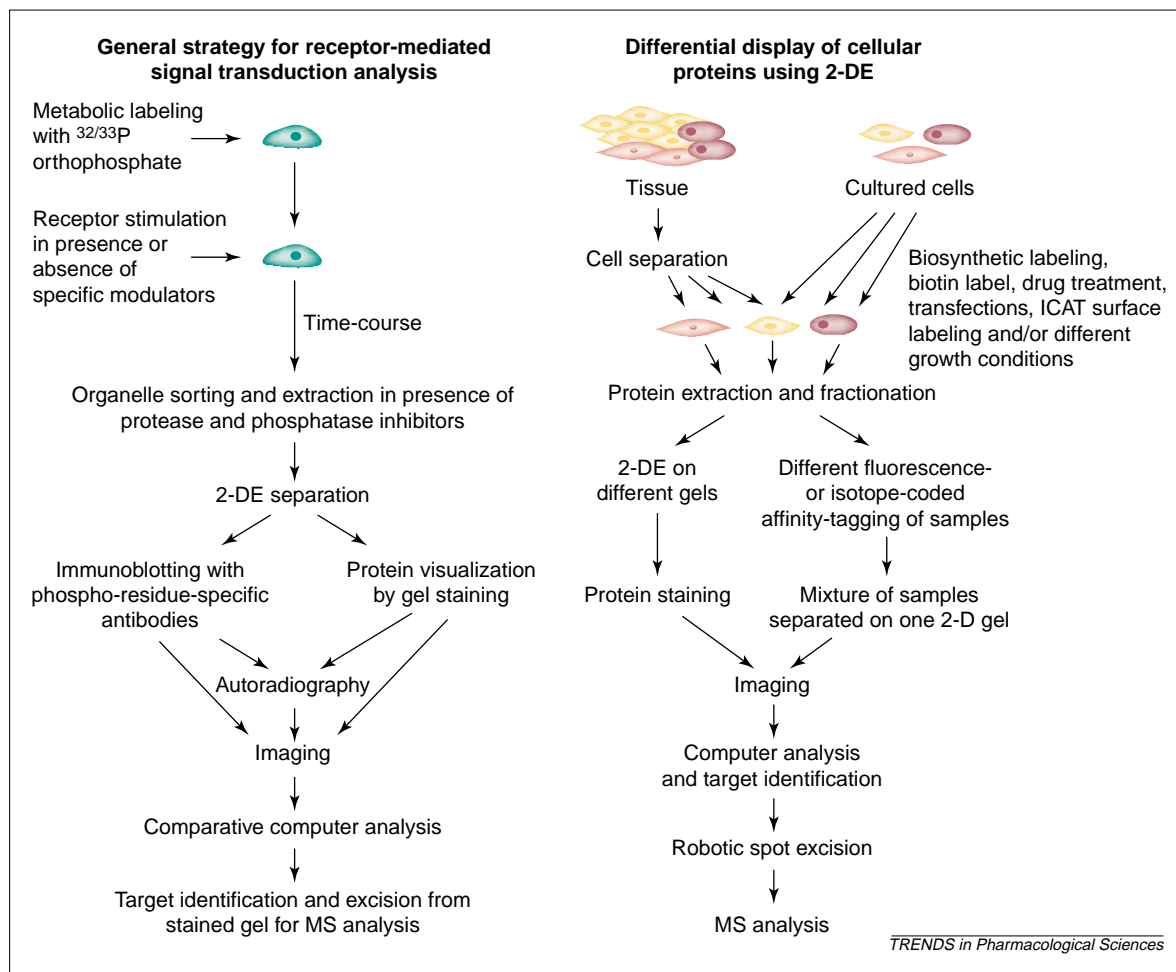
### Box 1. Modern proteomics

Post-genomic proteomics includes areas such as: (1) determination of protein function; (2) characterization of post-translational modifications; (3) structural analysis; (4) analysis of the regulation of protein activity; (5) studies of protein interactions and complex formation; (6) analysis of protein trafficking and sequestration in subcellular compartments; (7) protein expression analysis; (8) analysis of signaling and metabolic pathways; (9) drug mode-of-action; and (10) toxicity studies.

Bioinformatic integration of data from such conditional studies combined with data obtained by transcriptomic analysis will hopefully lead to a more comprehensive understanding of gene function and regulation, and to new approaches in the diagnosis, prevention and treatment of diseases.

Soren Naaby-Hansen  
Michael D. Waterfield\*  
Rainer Cramer  
Ludwig Institute for  
Cancer Research and  
Dept of Biochemistry and  
Molecular Biology, Royal  
Free and University  
College London Medical  
School, 91 Riding  
House Street, London,  
UK W1W 7BS.  
\*e-mail:  
mike@ludwig.ucl.ac.uk

Fig. 1. General strategies for the analysis of signal transduction pathways and protein expression levels using two-dimensional gel electrophoresis (2-DE). Abbreviation: ICAT, isotope-coded affinity tag; MS, mass spectrometry.



similar large-scale, highly automated affair. The sensitivity of proteomics suffers from the lack of an amplification method, analogous to the polymerase chain reaction method, to reveal and quantify the presence of low-abundance proteins. Shortcomings in throughput are due to the absence of technologies that can deliver fast and parallel quantitative analysis of complex protein distributions in an automated fashion. Industry is attempting to fill this gulf, but academia remains unable to process, for example, more than 20–30 two-dimensional (2-D) gels per week. However, collaborations between academia and industry can enhance the productivity in basic science<sup>3</sup>.

#### Two-dimensional gel electrophoresis

The 25-year-old, labor-intensive 2-DE technique is the only established method available to most laboratories to resolve large numbers of proteins in a quantitative fashion. Currently, 2-DE is the method of choice for quantitative differential display of large numbers of proteins and is being used to monitor external activation of proteins that participate in signaling pathways in addition to their perturbation by disease and drug action (Fig. 1). The resolving power of 2-DE makes it a useful technique for the identification and subsequent purification of post-translationally modified proteins.

However, several factors limit the utility of this technique for large-scale analysis of crude cellular and tissue extracts. Recent proteome data analysis<sup>4</sup> suggests that only a fraction of the genes is switched on in a given cell type at any given time. If a given cell expresses between 5000 and 10 000 genes at a given time, between 15 000 and 30 000 distinct cellular proteins must be expected as a result of mRNA splicing and post-translational modifications. Proteomic analysis of complex samples is further complicated because the dynamic range of protein abundance in cells might be as high as eight orders of magnitude<sup>5</sup> and even higher for proteins in biological fluids<sup>6</sup>.

A standard 18 × 18 cm 2-D gel [isoelectric focusing (IEF) followed by polyacrylamide gel electrophoresis (PAGE)] of 150 μg cell lysate will resolve between 1500 and 2000 proteins, dependent on sample preparation and detection strategy. Consequently, routine 2-DE analysis of complex mixtures such as cell and tissue extracts is biased towards long-lived abundant proteins. Low-abundance proteins (which might be the most interesting functionally) tend to be masked by house-keeping proteins present at several orders of magnitude higher levels and are invisible because of sensitivity detection limits. In addition, standard 2-D gels do not reflect a true representation of hydrophobic, highly insoluble, very basic, as well as

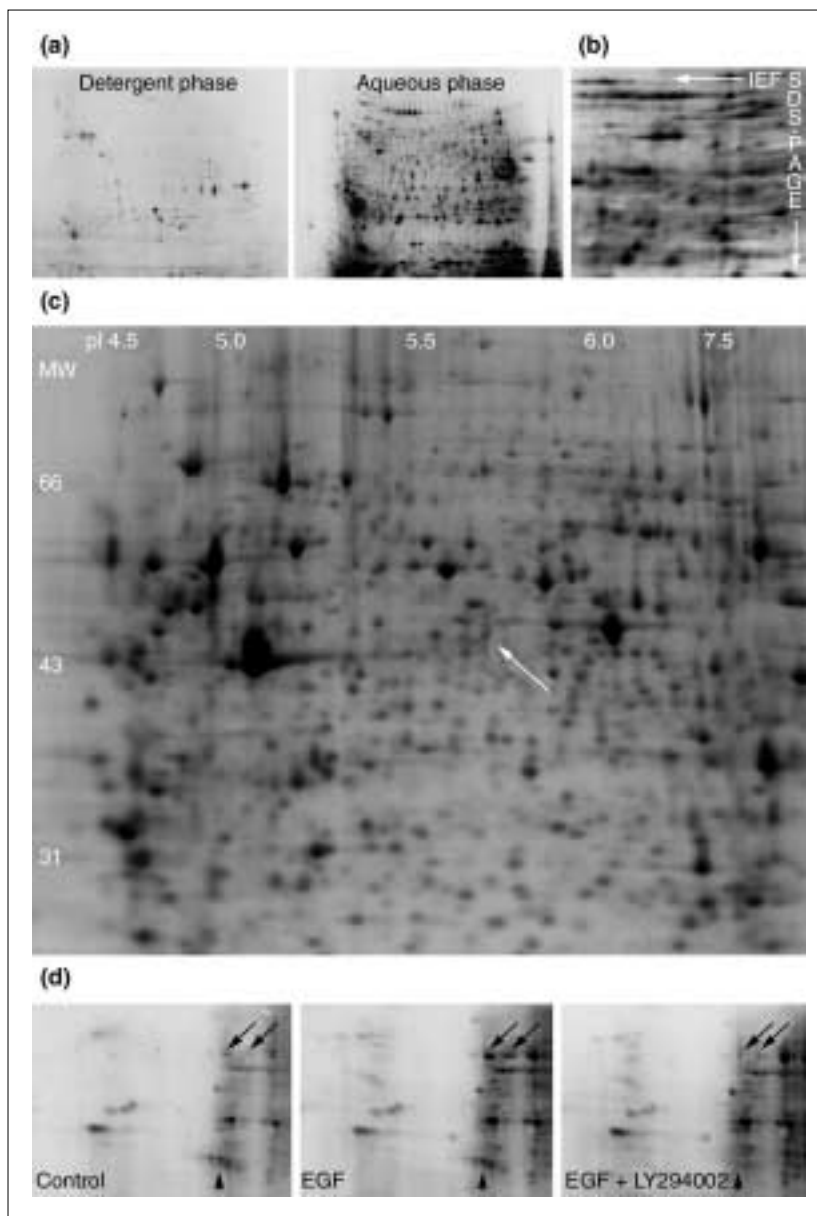


Fig. 2. Analysis of cell lysates using two-dimensional gel electrophoresis (2-DE). (a) Triton X-114 phase partitioning of proteins from epithelial cells. Proteins in the two phases were labeled on primary amines with *N*-hydroxysuccinimidyl derivatives of different fluorescent cyanine dyes (Cy3<sup>Amax550nm/Emax570nm</sup> and Cy5<sup>Amax650nm/Emax670nm</sup>, respectively, Amersham Pharmacia Biotech). The two fluorescently tagged samples were then mixed and separated on the same 2-D gel. The separated samples were differentially displayed following gel scanning (ImageMaster, Amersham Pharmacia Biotech). A few proteins were present in both phases, whereas the vast majority of proteins was separated in either the hydrophobic or the hydrophilic fraction. (b) The Rotofor Cell™ (Bio-Rad) liquid isoelectric focusing (IEF) system was employed to fractionate non-ionic detergent, urea-extracted proteins from epithelial cells according to net charge. Fractions containing proteins with pIs between 4 and 5 were pooled, dialyzed and concentrated. A sample of 3.6 mg was loaded onto a homecast 6-mm wide, narrow pH-range immobilized pH gradient strip (pH 4–5). Second-dimensional polyacrylamide gel electrophoresis (SDS-PAGE) was performed employing a 2-cm stacking gel. The area of resulting silver-stained gel is shown. (c) 2-DE of fibroblast proteins detected by the post-electrophoretic, non-covalent fluorescence staining methods of Oxford GlycoSciences. The fluorescence stain is not commercially available, but was used as a result of our collaborative programme with Oxford GlycoSciences. The protein indicated by the arrow was submitted for mass spectrometry (MS) analysis, following robotic spot excision. MS data obtained from this protein are given in Figs 4 and 5. (d) Autoradiographic detection of acidic <sup>33</sup>P-labeled phosphoproteins from an epithelial cell line (left panel). First-dimensional separation was performed using IEF with carrier ampholytes in tube gels, a technique we find best suited for analysis of labile post-translational modifications, such as phosphorylation. Additional proteins were phosphorylated following epidermal growth factor (EGF) receptor stimulation (middle panel). EGF stimulation in the presence of the phosphoinositide 3-kinase (PI3K) inhibitor LY294002 enabled analysis of EGF-induced phosphorylation downstream of PI3K (arrows). The phosphoprotein cluster indicated by the arrowhead was present in unstimulated cells, but disappeared following inhibition of PI3K, which suggests that phosphorylation of these proteins is maintained via the PI3K pathway under normal growth conditions. This is a useful complementary approach in drug target validation.

very small and very large, proteins. Although several steps in the 2-DE procedure are ill suited for automation, recent improvements in 2-DE technology, such as sample preparation procedures, gel design and loading procedures, protein detection methods and image analysis systems, together with semi-automated staining, spot picking and digestion procedures, aim to address these limitations.

#### Pre-fractionation

Sample pre-fractionation techniques aim to reduce the diversity and complexity of protein mixtures, thus increasing the number and concentration of distinct subsets of proteins that can be resolved in a series of complementary narrow pH-range 2-D gels (Fig. 2). The need for such steps has become evident<sup>2</sup>. Sequential extraction of proteins from a cell or tissue can be used to pre-fractionate proteins based on their relative solubility in a series of buffers<sup>7</sup>.

Pre-fractionation can also be based on different physio-chemical properties such as net charge, mobility, size, hydrophobicity<sup>8</sup> and affinity<sup>6</sup>. Subcellular fractionation is another useful approach that aims to enrich functionally related proteins based on their colocalization within the cell<sup>9–11</sup>, which is equally useful for the study of protein translocations. In general, pre-fractionation procedures for 2-DE analysis should be kept as simple as possible, target different molecular properties and employ downstream compatible solutions to minimize sample loss and the chance of degradation and/or introduction of artefactual protein modifications.

#### Fluorescent dyes

Improved methods of protein detection including the use of fluorescent dyes, which offer a broader dynamic and linear quantitative range of detection than silver staining (which is non-linear<sup>12,13</sup>), have also enhanced the number of proteins that can be detected following 2-DE. Difference gel electrophoresis (DIGE), in which two or three samples labeled with different fluorescence probes can be run in a single 2-DE gel, offers the opportunity of direct comparison of different samples, thus avoiding gel-to-gel variations inherent to comparative gel analysis<sup>14</sup>.

## Box 2. Mass spectrometry measures the mass of ions with high accuracy

The quantity measured by mass spectrometry is the mass of ions produced, separated and detected by the various ionization techniques, mass analyzers and detectors, respectively. Mass spectral data interpretation is therefore essential. Depending on the accuracy of the mass measurement and the elemental composition, important structural information can be gained. For example, the detection of a singly charged negative organic ion with a mass of 78.96 Da unambiguously identifies it as  $\text{PO}_3^-$  allowing for C, H, O, N and P as possible elements and a mass measurement accuracy of 100 ppm (50 ppm when S is included). Hence, this ion signal can identify phosphopeptides when ion fragmentation is induced before mass analysis. A recent report by Clauser *et al.*<sup>a</sup> illustrates the role of accurate mass measurement.

### Reference

- a Clauser, K.R. *et al.* (1999) Role of accurate mass measurement ( $\pm 10$  ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal. Chem.* 71, 2871–2882

### Immobilized pH gradient strips

The process of solubilization and subsequent separation of hydrophobic proteins in immobilized pH gradient (IPG) strips have been improved by combining more effective reducing agents such as the uncharged agent tributyl phosphine<sup>15</sup> with more powerful chaotropes such as thiourea<sup>16</sup> and surfactants such as linear sulfobetaine-type zwitterionic detergents containing a carboxyamido group to improve urea tolerance<sup>17</sup>. In-gel rehydration, in which the dehydrated IPG strip is re-swollen directly with the protein sample, have also improved resolution of hydrophobic proteins by reducing their aggregation at the loading point. Hydrophobic proteins might also be

lost as a result of hydrophobic interactions between the proteins and the basic acrylamido derivatives of the IPG matrix<sup>18</sup>. The protein recovery can be improved by the use of thiourea, sulfobetaines and a small amount of non-ionic detergent in the sample buffer and by addition of carrier ampholytes (CA) to the sample before first-dimensional IPG electrophoresis. One of the most promising 2-DE membrane protein analyses using IPG strips was recently performed by Molloy and colleagues<sup>19</sup>, who identified 21 out of 26 (80%) of the predicted *Escherichia coli* integral outer membrane proteins (OMPs) that are annotated in SWISS-PROT release 37, and should separate within the range of pH 4–7 with a molecular mass of 10–80 kDa. The authors combined OMPs isolation following carbonate incubation, 2-DE and mass spectrometry (MS) analysis.

Until recently, first-dimensional electrophoretic separation of basic proteins was only possible using non-equilibrium pH-gradient electrophoresis (NEPHGE)<sup>20</sup>. NEPHGE is technically demanding and variations in protein migration between experiments are frequent, even for the experienced operator. However, the recent development of basic IPGs up to pH 12 will facilitate easy, reproducible analysis of alkaline proteins as soon as the problem with reverse electroosmotic flow (anode-directed water flow) has been fully solved<sup>21</sup>.

2-DE analysis of high molecular weight (HMW) proteins (>150 kDa) is best performed using carrier ampholyte isoelectric focusing (CA-IEF) in tube gels for the first-dimensional separation or by 'in-gel rehydration' of IPG strips using low voltages. Both of these approaches facilitate the migration of HMW proteins during IEF. The use of tube gels enhances the recovery of HMW proteins into the second-dimension low-density gradient SDS-PAGE gels. Alternatively, the first-dimension isoelectric focusing can be performed in agarose. However, 2-DE is unsuitable for analysis of very HMW proteins, which should be separated using one-dimensional SDS-PAGE, preferably following sample pre-fractionation to enrich the targeted proteins. Improved second-dimensional separation of very low molecular weight proteins can be achieved if the Tris-glycine buffer, which is usually used, is substituted by the Tris-tricine-SDS buffer system in combination with high-percentage-gradient gels.

### Biological mass spectrometry

Recent advances in MS (Box 2) have been the pivotal element in the establishment of proteomics research projects. Electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) in combination with new instrumental designs such as high-voltage delayed extraction ion sources and hybrid quadrupole-orthogonal-acceleration-time-of-flight (QTOF) mass analyzers have revolutionized mass spectrometric protein analysis. Before these new developments in MS, protein characterization

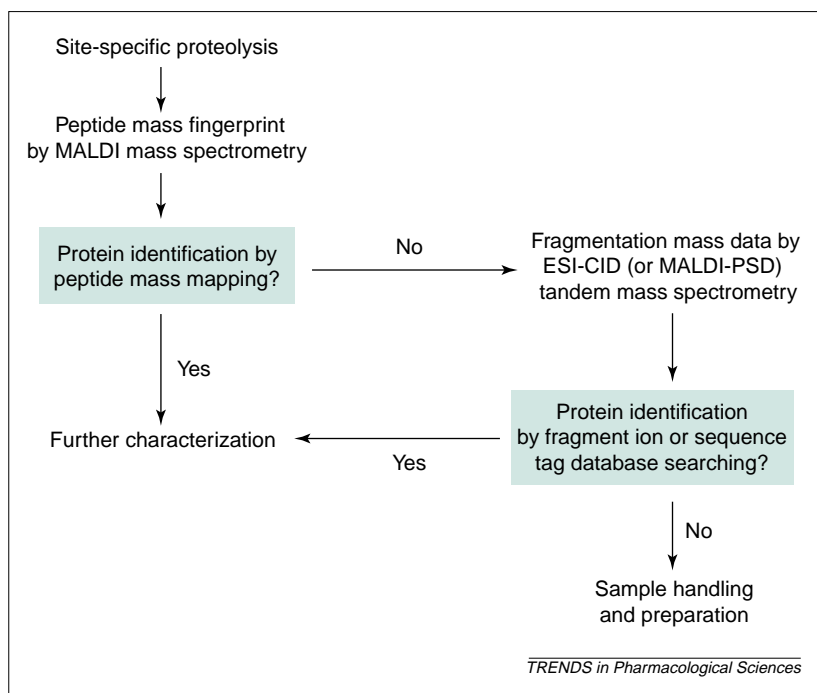


Fig. 3. Typical workflow for protein identification by mass spectrometry (see text for further description). Abbreviations: CID, collision-induced dissociation; ESI, electrospray ionization; MALDI, matrix-assisted laser desorption/ionization; PSD, post-source decay.



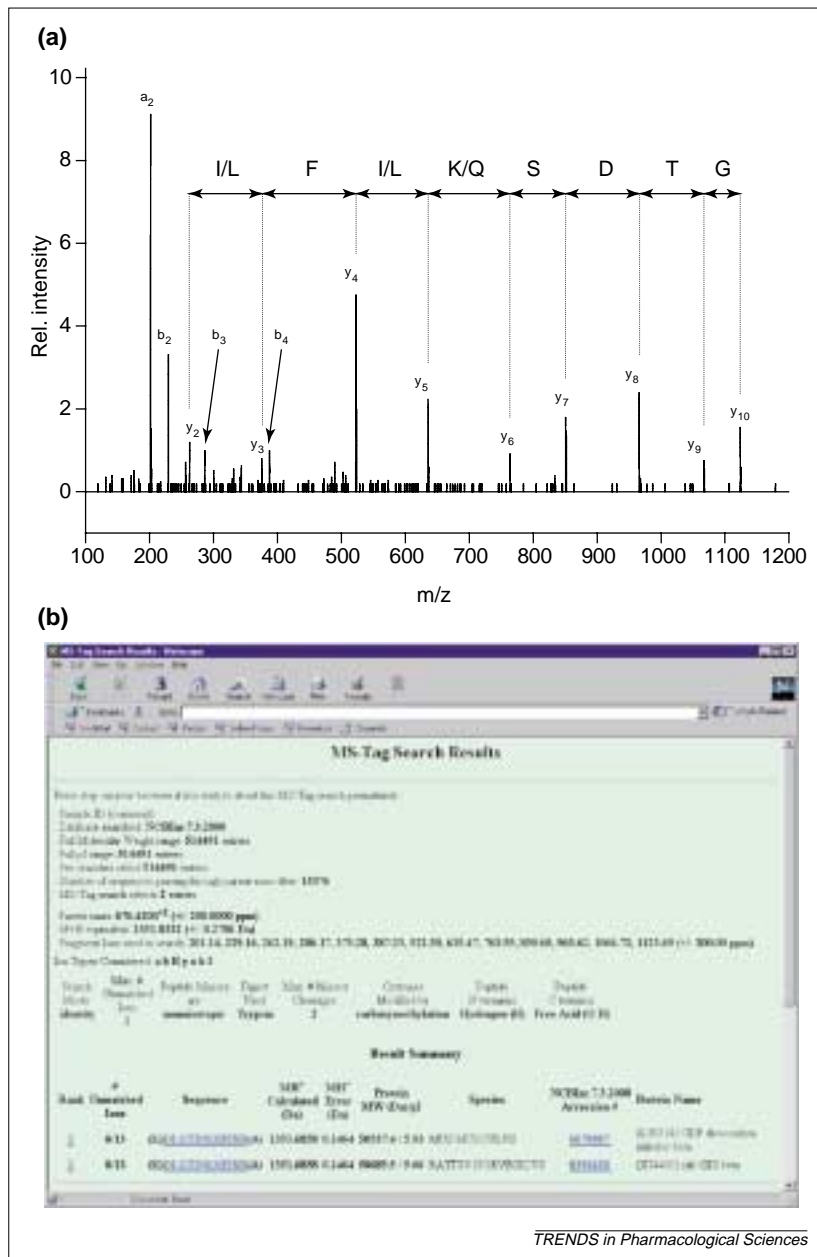


Fig. 5. Protein identification by sequence tag database searching of the gel spot sample marked in Fig. 2c. (a) ESI CID MS/MS (electrospray ionization collision-induced dissociation tandem mass spectrometry) fragmentation mass spectrum of a doubly charged peptide ion at  $m/z$  676.42 Da from the same tryptic in-gel digest as in Fig. 4. (b) Database search results using the 13 monoisotopic masses of the most intense fragment ion signals in (a). Mass spectral annotation of fragment ions considers DLGTDSQIFISR as the parent peptide ion. It follows the nomenclature proposed by Roepstorff and Fohlman<sup>63</sup>. Mass differences within the  $y$ -ion series enable the nomination of amino acids with possible substitution of leucine (L) for isoleucine (I) and glutamine (Q) for lysine (K). This way a sequence tag can be created. However, the ProteinProspector MS-Tag search routine was used for database searching using fragment ion masses only, without the knowledge about the fragment ion type. The search results summary consists of two proteins only, the GDP dissociation inhibitor beta homologs from mouse and rat, although mass accuracy was set at 200 ppm for both parent ion mass and fragment ion masses and one unmatched fragment ion and two missed cleavages was allowed for (compare with the peptide mass mapping search results and parameters in Fig. 2).

superior to MALDI PSD because of its greater MS/MS sensitivity and less complex fragment ion spectra and is therefore the preferred method for MS/MS (Ref. 27). ESI can be performed by direct infusion using a syringe pump or nano-vials. Alternatively, high-pressure liquid chromatography (HPLC) can be

interfaced to ESI ion sources to clean up the sample as well as separate and concentrate peptides. Because ESI and MALDI often yield different sets of peptides, ESI MS is initially performed to identify suitable ions for ESI MS/MS. Nonetheless, the peptide mass information obtained from the earlier MALDI peptide mass mapping is often helpful for ESI MS/MS.

Frequently, ESI MS/MS is successful on peptide ions identified by MALDI MS, but non-detectable in the ESI MS mode. Ideally, fragmentation spectra from at least one peptide can be recorded (Fig. 5a). Similar to database searching in peptide mass mapping, fragment ion masses are compared to theoretical fragment ion masses from tryptic peptides of database proteins<sup>25</sup> (for web search engines see Box 3). Conversely, due to the high specificity of this data, protein identification is far less ambiguous (Fig. 5b) and proteins can be ranked according to the number of mismatched fragment ions.

The most important prerequisite for this identification strategy is the database presence of the protein. Otherwise *de novo* sequencing to various degrees might be necessary to screen cDNA libraries and/or degenerate oligonucleotides and clone the entire gene sequence.

#### Sample preparation for mass spectrometric analysis in proteomics studies

It has been shown that sample amounts in the zeptomole ( $10^{-21}$  mol) range can be analyzed by MS. However, for 2-DE gel spot analysis realistic sensitivity limits are around the low femtomole level. Sample origin and constitution can severely limit the sensitivity. Proteins from 2-DE gel spots, for example, are not readily accessible for proteolysis. The method of choice is still in-gel digestion, although the protein confinement in the gel spot and its volume of typically more than  $1 \text{ mm}^3$  governs essential sample preparation steps. Combined with sample volumes of  $0.5 \mu\text{l}$  used in classical MALDI MS sample spotting, sample concentrations can be so low that adhesion to tubes, tips and other surfaces results in substantial sample loss. Non-gel separation techniques such as capillary (zone) electrophoresis or the new isotope-coded affinity tag (ICAT)<sup>28</sup> method are approaches to circumvent the problems associated with gel electrophoresis. These techniques can provide the sample in an MS-compatible constitution, resulting in less discrimination of proteins, which are difficult to analyze on or from gels. Isotope labeling also enables quantification via MS and can therefore be used for the identification of differentially expressed proteins.

In general, sample preparation protocols that limit the number of preparation steps, circumvent the loss or dilution of the sample and purify and concentrate the sample will be the next major step forward to higher sensitivity. Miniaturization will also play an important role in future preparation techniques for proteomics<sup>29,30</sup>.

### Future developments in mass spectrometry for proteomics

Apart from sample preparation, instrumentation as well as data processing and analysis are fields of constant development. Two aspects in proteomics guide instrumental development: mass accuracy and sensitivity. In MALDI TOF instruments, ion formation and mass analysis are directly coupled. Hence, the initial kinetic ion energy spread from the laser desorption/ionization process severely limits the mass accuracy. Because mass analysis with QTOF or Fourier transformation (FT) MS instruments depend far less on the ion formation history, this influence is significantly reduced<sup>31,32</sup>. FT mass spectrometers with MALDI and ESI ion sources have been used for some time, but their complexity, cost and lower sensitivity are major disadvantages. However, the mass accuracy (and resolution) that is achievable with FT MS is exceptional. Recent results obtained with a MALDI source on QTOF instruments also exhibit improved mass accuracy combined with a potentially high sensitivity<sup>33</sup>. Furthermore, infrared (IR)-MALDI MS, being more affected by its desorption characteristics than UV-MALDI MS, can benefit even more from the QTOF design. Its softer ionization character making labile ion species much more accessible is its distinguished advantage and could make it the preferred MALDI technique<sup>34</sup>.

Advances in proteomics research can also be expected from improvements in peptide dissociation. Increasing the precursor ion conversion efficiency and directing dissociation into one specific type of fragmentation channel resulting in simpler fragmentation spectra are major aspects of such improvements. Several publications have shown that fixed-charge derivatization of peptides can achieve these goals<sup>35,36</sup>. Alternatively, the newly emerging electron capture dissociation (ECD) technique on FT MS instruments exhibits these characteristics combined with extremely low side-chain fragmentation<sup>37,38</sup>. However, for both techniques sensitivity limits for 2-DE gel samples showing their superiority to conventional ESI CID MS/MS still have to be established.

Furthermore, the introduction of new search algorithms will increase the confidence level of protein identification<sup>39</sup> and can overcome the importance of absolute mass accuracy as long as the relative accuracy is sufficiently maintained<sup>40</sup>. These trends will probably lead to search engines based on peak pattern recognition rather than simple shared peak counts.

Data filtering will also become increasingly important. For example, it seems that the omnipresent chemical noise is periodically well defined, and FT and frequency filtering can therefore dramatically increase the signal-to-noise ratio leading to higher sensitivity.

### Proteomic studies beyond protein identification

Although the initial goal of proteomics studies is the mapping of the proteome of biological systems

(e.g. tissues, cells and fluids, among others) there are many challenges in proteomics beyond the mere identification of proteins. For example, the implication of post-translational modifications is one of the main reasons for research groups to switch from the genomic to the proteomic level. Unfortunately, the biologically most common post-translational modifications (phosphorylation and glycosylation) often pose challenges in protein analysis as a result of their high lability, complexity and low stoichiometry. Identifying phosphorylation sites from 2-D gel spots of isoforms can be extremely difficult. Enrichment techniques for both the protein of interest and its phosphorylated peptides can dramatically improve sensitivity. Protein-specific affinity purification such as tandem affinity purification (TAP)<sup>41</sup> with subsequent immobilized metal-ion chromatography (IMAC)<sup>42</sup> separation of phosphopeptides can be superior to 2-D gel separation combined with in-gel digestion for the identification of novel phosphorylation sites.

The newly developed TAP technique aside other immunoaffinity purification techniques can also contribute to the objective of increased analytical sensitivity for the study of protein complexes. This technique can provide an important alternative to overexpression and immunoprecipitation of ordinary fusion proteins as a result of its potential for the isolation of protein complexes at their natural expression levels and the proteolytic release of the bound material under native conditions.

Isolation of protein complexes using immunoaffinity technologies can also be achieved by phage antibody display systems, which bypasses hybridoma technology and immunization, and by generation of affibodies using combinatorial chemistry, as recently described by Gunneriusson and colleagues<sup>43</sup>. The use of phage particles that express the cytosolic tail of the epidermal growth factor receptor, Shc-pY317 peptide or Grb2 fusion protein as baits has recently been successfully applied for mapping novel components of the signal transduction complexes downstream of the receptor<sup>44</sup>.

Large-scale yeast two-hybrid screens have resulted in large numbers of putative interaction partners to be identified in *Saccharomyces cerevisiae*<sup>45</sup> and in *E. coli*<sup>46</sup>. However, the yeast two-hybrid system identifies pairwise interactions rather than whole complexes and relies on protein interactions in the nucleus.

Protein array techniques similar to the methods employed for semi-quantitative mRNA solid-phase assays are being developed<sup>47</sup>. The bait on protein chips can be antibodies, known binding partners, or key scaffolding proteins, which in principle can be displayed by bacteriophages. By combination with laser capture microscopic dissection of tissue this might become a powerful tool in the study of signal transduction complex activation in oncology<sup>48</sup>.

An interesting new strategy for the identification of nucleic-acid-protein interactions employs an *in vitro* selection-amplification method. Genomic SELEX

(Ref. 49) consists of repeated rounds of binding a library of nucleic acids to the target protein, separating the bound nucleic acids from the unbound ones and amplifying the bound ones for the next round. At the end of the selection, the resulting high-affinity nucleic acid molecules are cloned and sequenced.

Affinity purification, proteolysis and other biomolecular reactions can also be achieved by immobilizing active surfaces on MALDI targets. This approach towards chip-based proteomics is used in surface-enhanced laser desorption/ionization (SELDI)<sup>50</sup> or biomolecular interaction analysis (BIA)<sup>51,52</sup> techniques. In surface plasmon resonance-biomolecular interaction analysis (SPR-BIA)<sup>51,52</sup> coupled to MALDI MS, the biomolecular reaction can be directly monitored and analyzed in a concerted approach.

In principle, tagging and affinity purification techniques in targeted proteomics research offer greater sample enrichment and recovery than gel electrophoretic techniques and will enable the elucidation of protein function and pathway-specific signaling through the characterization of post-translational modifications, binding partners and sites, and their conditional and temporal changes.

#### Biological insight from proteomics

The powerful combination of 2-DE, MS and bioinformatic analysis has provided valuable information in a variety of biological fields, including signal transduction analysis<sup>53</sup>, oncology<sup>54</sup>, cardiomyopathy<sup>55</sup>, drug action analysis<sup>56,57</sup>, toxicology<sup>58,59</sup> and the pathogenesis of schizophrenia<sup>60</sup>, and has enabled identification of novel components of the spliceosome<sup>61</sup>.

Phosphorylated proteins involved in signal transduction can be identified by the temporal differential display approach employing phospho-isotope labeling or immunoblotting with phospho-residue specific antibodies. The divergence and convergence of phospho-signaling pathways can be dissected by the application of inhibitors to kinases or phosphatases at the key crossroads. Figure 2d illustrates such an approach, which is depicted in Fig. 1.

Several approaches have been made to circumvent 2-DE with its inherent difficulties as the method of choice for the separation of complex protein mixtures. 2-D liquid chromatography coupled to MS has been

applied successfully in the study of the relatively small proteome of *E. coli*<sup>62</sup>. However, the protein separation capacity of this approach is lower than that of 2-DE, and exhaustive analysis of the sequence data obtained using this method demands powerful computing facilities. Improvements in sample preparation, resolution and data analysis are therefore necessary before multi-dimensional liquid chromatography becomes functional for the study of proteomes of higher and more complex organisms.

#### Concluding remarks

The post-genomic era will focus on the analysis of gene function via analysis of the proteome. Modern proteomics needs immense powers of resolution, sensitivity and speed of identification, which so far can only be reached through 2-D gels and MS in combination with advanced bioinformatic analysis. New techniques that can deliver fast and parallel quantitative analysis of complex protein distributions in an automated fashion need to be developed before proteomics can reach its full potential. Only a handful of companies have started to develop such systems. The detail in the catalogues, which Celera and others might produce by the use of high-throughput proteomics employing robotic front ends for mass spectrometers, is as yet unknown. The development of very sensitive optical reporters that can be used to visualize proteins will help detection and push differential analysis to greater utility, but in the end the whole shredded or blotted gel might be analyzed by MS to reveal proteins only detectable at present by isotopic labels. Chips, which pluck complexes from cells at high-throughput, along with new affinity purification, separation and detection methods, must be developed for fast analysis of the signaling pathways under various conditions and allow detection and analysis of low copy number proteins. The future of such chip-based analyzers linked to mass spectrometers could give true high-throughput and sensitivity, allowing massive protein screening to be routine in research and even medical diagnostics. Obviously, any proteomics system has to integrate instrumentation and protocols depending on the methods used for protein purification, separation and characterization. However, whatever methods for protein separation and purification will be used in the future, MS seems to be ideal for both high-throughput and high-sensitivity protein analyses.

#### References

- Anderson, L. and Seilhamer, J. (1997) A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* 18, 533–537
- Cygi, S.P. *et al.* (1999) Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* 19, 1720–1730
- Page, M.J. *et al.* (1999) Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mamoplasties. *Proc. Natl. Acad. Sci. U. S. A.* 96, 12589–12594
- Celis, J.E. *et al.* (1998) Human and mouse proteomic databases: novel resources in the protein universe. *FEBS Lett.* 430, 64–72
- Anderson, N.L. and Anderson, N.G. (1998) Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* 19, 1853–1861
- Corthals, G.L. *et al.* (2000) The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* 21, 1104–1115
- Ramsby, M.L. and Makowski, G.S. (1999) Differential detergent fractionation of eukaryotic cells. In *2-D Proteome Analysis Protocols* (Link, A.J., ed.), pp. 53–66, Humana Press
- Wissing, J. *et al.* (2000) Enrichment of hydrophobic proteins via Triton X-114 phase partitioning and hydroxyapatite column chromatography for mass spectrometry. *Electrophoresis* 21, 2589–2593
- Rouquie, D. *et al.* (1997) Construction of a directory of tobacco plasma membrane proteins by combined two-dimensional gel electrophoresis and protein sequencing. *Electrophoresis* 18, 654–660
- Maicher, M.T. and Tiedtke, A. (1999) Biochemical analysis of membrane proteins from an early maturation stage of phagosomes. *Electrophoresis* 20, 1011–1016
- Pasquali, C. *et al.* (1999) Subcellular fractionation, electromigration analysis and mapping of organelles. *J. Chromatogr. B Biomed. Sci. Appl.* 722, 89–102

- 12 Rabilloud, T. and Charmont, S. (2000) Detection of proteins on 2D electrophoresis gels. In *Proteome Research: Two-dimensional Electrophoresis and Identification Methods* (Rabilloud, T., ed.), pp. 107–126, Springer
- 13 Patton, W.F. (2000) A thousand points of light: the application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics. *Electrophoresis* 21, 1123–1144
- 14 Unlu, M. *et al.* (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 18, 2071–2077
- 15 Herbert, B.R. *et al.* (1998) Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reducing agent. *Electrophoresis* 19, 845–851
- 16 Rabilloud, T. *et al.* (1997) Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 18, 307–316
- 17 Chevallet, M. *et al.* (1998) New zwitterionic detergents improve the analysis of membrane proteins by 2-D electrophoresis. *Electrophoresis* 19, 1901–1909
- 18 Righetti, P.G. and Bossi, A. (1997) Isoelectric focusing in immobilized pH gradients: an update. *J. Chromatogr. B Biomed. Sci. Appl.* 699, 77–89
- 19 Molloy, M.P. *et al.* (2000) Proteomic analysis of the *Escherichia coli* outer membrane. *Eur. J. Biochem.* 267, 2871–2881
- 20 Naaby-Hansen, S. *et al.* (1997) Two-dimensional gel electrophoretic analysis of vectorially labeled surface proteins of human spermatozoa. *Biol. Reprod.* 56, 771–787
- 21 Görg, A. *et al.* (2000) The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21, 1037–1053
- 22 Shevchenko, A. *et al.* (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl. Acad. Sci. U. S. A.* 93, 14440–14445
- 23 Yates, J.R. (1998) Mass spectrometry and the age of the proteome. *J. Mass Spectrom.* 33, 1–19
- 24 Gevaert, K. and Vandekerckhove, J. (2000) Protein identification methods in proteomics. *Electrophoresis* 21, 1145–1154
- 25 Yates, J.R., III (1998) Database searching using mass spectrometry data. *Electrophoresis* 19, 893–900
- 26 Spengler, B. (1997) Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules. *J. Mass Spectrom.* 32, 1019–1036
- 27 Humphrey-Smith, I. *et al.* (1997) Proteome research: complementarity and limitations with respect to the RNA and DNA worlds. *Electrophoresis* 18, 1217–1242
- 28 Gygi, S.P. *et al.* (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 17, 994–999
- 29 Staudenmann, W. *et al.* (1998) Sample handling for proteome analysis. *Electrophoresis* 19, 901–908
- 30 Gevaert, K. *et al.* (1998) A peptide concentration and purification method for protein characterization in the subpicomole range using matrix-assisted laser desorption/ionization-postsource decay (MALDI-PSD) sequencing. *Electrophoresis* 19, 909–917
- 31 Guilhaus, M. *et al.* (2000) Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrom. Rev.* 19, 65–107
- 32 Marshall, A.G. *et al.* (1998) Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom. Rev.* 17, 1–35
- 33 Shevchenko, A. *et al.* (2000) MALDI quadrupole time-of-flight mass spectrometry: a powerful tool for proteomic research. *Anal. Chem.* 72, 2132–2141
- 34 Cramer, R. and Burlingame, A.L. (2000) IR-MALDI – softer ionization in MALDI-MS for studies of labile macromolecules. In *Mass Spectrometry in Biology and Medicine* (Burlingame, A.L. *et al.*, eds), pp. 289–307, Humana Press
- 35 Adamczyk, M. *et al.* (1999) Charge derivatization of peptides to simplify their sequencing with an ion trap mass spectrometer. *Rapid Commun. Mass Spectrom.* 13, 1413–1422
- 36 Keough, T. *et al.* (1999) A method for high-sensitivity peptide sequencing using postsource decay matrix-assisted laser desorption ionization mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 96, 7131–7136
- 37 Mirgorodskaya, E. *et al.* (1999) Localization of O-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. *Anal. Chem.* 71, 4431–4436
- 38 Zubarev, R.A. *et al.* (2000) Electron capture dissociation for structural characterization of multiply charged protein cations. *Anal. Chem.* 72, 563–573
- 39 Zhang, W. and Chait, B.T. (2000) ProFound: an expert system for protein identification using mass spectrometric peptide mapping information. *Anal. Chem.* 72, 2482–2489
- 40 Egelhofer, V. *et al.* (2000) Improvements in protein identification by MALDI-TOF-MS peptide mapping. *Anal. Chem.* 72, 2741–2750
- 41 Rigaut, G. *et al.* (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17, 1030–1032
- 42 Posewitz, M.C. and Tempst, P. (1999) Immobilized gallium (III) affinity chromatography of phosphopeptides. *Anal. Chem.* 71, 2883–2892
- 43 Gunnerusson, E. *et al.* (1999) Affinity maturation of a Taq DNA polymerase specific affibody by helix shuffling. *Protein Eng.* 12, 873–878
- 44 Zozulya, S. *et al.* (1999) Mapping signal transduction pathways by phage display. *Nat. Biotechnol.* 17, 1193–1198
- 45 Uetz, P. *et al.* (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* 403, 623–627
- 46 Houry, W.A. *et al.* (1999) Identification of *in vitro* substrates of the chaperonin GroEL. *Nature* 402, 147–154
- 47 Lueking, A. *et al.* (1999) Protein microarrays for gene expression and antibody screening. *Anal. Biochem.* 270, 103–111
- 48 Liotta, L. and Petricoin, E. (2000) Molecular profiling of human cancer. *Nat. Rev. Genet.* 1, 48–56
- 49 Shtatland, T. *et al.* (2000) Interactions of *E. coli* RNA with bacteriophage MS2 coat protein: genomic SELEX. *Nucleic Acids Res.* 28, E93
- 50 Merchant, M. and Weinberger, S.R. (2000) Recent advancements in surface-enhanced laser desorption/ionization-time of flight–mass spectrometry. *Electrophoresis* 21, 1164–1177
- 51 Nelson, R.W. *et al.* (2000) Biosensor chip mass spectrometry: a chip-based proteomic approach. *Electrophoresis* 21, 1155–1163
- 52 Williams, C. and Addona, T.A. (2000) The integration of SPR biosensors with mass spectrometry: possible applications for proteome analysis. *Trends Biotechnol.* 18, 45–48
- 53 Soskic, V. *et al.* (1999) Functional proteomics analysis of signal transduction pathways of the platelet-derived growth factor  $\beta$  receptor. *Biochemistry* 38, 1757–1764
- 54 Ostergaard, M. *et al.* (1999) Psoriasis (S100A7): a putative urinary marker for the follow-up of patients with bladder squamous cell carcinomas. *Electrophoresis* 20, 349–354
- 55 Weekes, J. *et al.* (1999) Bovine dialated cardiomyopathy: proteomic analysis of an animal model of human dilated cardiomyopathy. *Electrophoresis* 20, 898–906
- 56 Steiner, S. *et al.* (2000) Proteomics to display lovastatin-induced protein and pathway regulation in rat liver. *Electrophoresis* 21, 2129–2137
- 57 Anderson, N.L. (1996) The effects of peroxisome proliferators on protein abundances in mouse liver. *Toxicol. Appl. Pharmacol.* 137, 75–89
- 58 Kanitz, M.H. (1999) Alterations in rabbit kidney protein expression following lead exposure as analyzed by two-dimensional gel electrophoresis. *Electrophoresis* 20, 2977–2985
- 59 Aicher, L. (1998) New insights into cyclosporine A nephrotoxicity by proteome analysis. *Electrophoresis* 19, 1998–2003
- 60 Edgar, P.F. *et al.* (2000) Comparative analysis of the hippocampus implicates chromosome 6q in schizophrenia. *Mol. Psychiatry* 5, 85–90
- 61 Neubauer, G. *et al.* (1998) Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat. Genet.* 20, 46–50
- 62 Opitck, G.J. *et al.* (1998) Comprehensive two-dimensional high-performance liquid chromatography for the isolation of overexpressed proteins and proteome mapping. *Anal. Biochem.* 258, 349–361
- 63 Roepstorff, P. and Fohlman, J. (1984) Proposal of a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom* 11, 601

### Chemical name

LY294002: [2-(4-morpholinyl)-8-phenyl-4H-1-benzopyran-4-one]

## Meeting Reports

TIPS is pleased to publish highlights from international conferences. Conference participants who wish to cover a meeting should contact the Editor.

TIPS, Elsevier Science London, 84 Theobald's Road, London, UK WC1X 8RR (<http://tips.trends.com>)