

# Proteomics techniques and their application to hematology

Ileana M. Cristea, Simon J. Gaskell, and Anthony D. Whetton

The recent sequencing of a number of genomes has raised the level of opportunities for studies on proteins. This area of research has been described with the all-embracing term, proteomics. In proteomics, the use of mass spectrometric techniques enables genomic databases to be used to establish the identity of proteins with relatively little data, com-

pared to the era before genome sequencing. The use of related analytical techniques also offers the opportunity to gain information on regulation, via posttranslational modification, and potential new diagnostic and prognostic indicators. Relative quantification of proteins and peptides in cellular and extracellular material remains a challenge for proteomics

and mass spectrometry. This review presents an analysis of the present and future impact of these proteomic technologies with emphasis on relative quantification for hematologic research giving an appraisal of their potential benefits. (Blood. 2004;103:3624-3634)

© 2004 by The American Society of Hematology

## Proteomics: a definition

The fundamental role of proteins in supporting life was recognized in the early stages of biologic research. The name “protein,” derived from the Greek term, *proteios*, meaning “the first rank,” was used for the first time by Berzelius in 1838 to illustrate the importance of these molecules. The multivariate functions of specific cells, from movement to mitosis, are regulated by (very) approximately 9000 specialized protein types per nucleated cell. Posttranslational modifications (PTMs, such as sulfation, phosphorylation, farnesylation, hydroxylation, methylation, glycosylation)<sup>1</sup> create microheterogeneity within a specific protein population, which adds to this complexity.

The array of proteins found within the cell, their interactions and modifications hold the key to understanding biologic systems. This is encapsulated in the term “proteome.” It can be defined as the protein population of a cell, characterized in terms of localization, PTM, interactions, and turnover, at any given time. The proteome is fundamentally dynamic and has an inherent complexity that surpasses that of the genome or the mRNA complement found within a cell (transcriptome; Figure 1). The development of DNA microarray technology allowed the analysis of gene expression at the mRNA level. Therefore, transcriptomics provides information about the degree of gene activity in individual tissues and the relationship to cell function, development stage, response to external stimuli, and disease.<sup>2,3</sup> Although transcriptome data reflect the genome’s objectives for protein synthesis, they do not provide information about the finalization of those objectives. Proteome analysis provides a view of biologic processes at their level of occurrence, thereby offering a better understanding of the physiologic and pathologic states of an organism, and becoming an important step in the development and validation of diagnostics and therapeutics. Studies of the correlation between transcriptome

and proteome data have illustrated both satisfactory as well as poor correlations between the mRNA and protein concentration and turnover.<sup>4-6</sup> The description of processes underlying hematopoietic cell development, leukemogenesis, and the functional activity of mature cells will be enhanced by application of proteomic approaches for protein characterization, study of protein-protein interaction, and, in particular, relative quantification. Here, the most effective and readily available techniques for relative sample analysis are described and their application to hematology explored.

## Separation procedures and gel electrophoresis-based quantification for proteomics

Most techniques currently used in proteomics use a variety of fractionation and separation steps prior to analysis by mass spectrometry (MS). Sample preparation and fractionation (see Dreger<sup>7</sup> and Huber et al<sup>8</sup>) are beyond the limits of this review, but their importance at different stages of a proteome study is not to be underestimated and some examples of commonly used techniques are introduced throughout this review. Sample fractionation can be performed according to compound type (see “PTM of proteins” and “SELDI-ToF MS and the search for disease markers”) or subcellular organelle localization (this section). Separation steps can be used at the protein level, as well as at the peptide level. Typical experiments include affinity separation methods, 1- or 2-dimensional gel electrophoresis, or 1- or 2-dimensional chromatographic separation.

Two-dimensional polyacrylamide gel electrophoresis (2-DE) was initially described by O’Farrell in 1975 and has evolved

From the Michael Barber Centre for Mass Spectrometry, Department of Chemistry, University of Manchester Institute of Science and Technology (UMIST), Manchester, United Kingdom; Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, Rockefeller University, New York, NY; Leukaemia Research Fund Proteomics Facility, Department of Biomolecular Sciences, UMIST, Manchester, United Kingdom; and Paterson Institute for Cancer Research, Christie Hospital, Withington, Manchester, United Kingdom.

Submitted October 6, 2003; accepted December 4, 2003. Prepublished online as *Blood* First Edition Paper, January 15, 2004; DOI 10.1182/

blood-2003-09-3295.

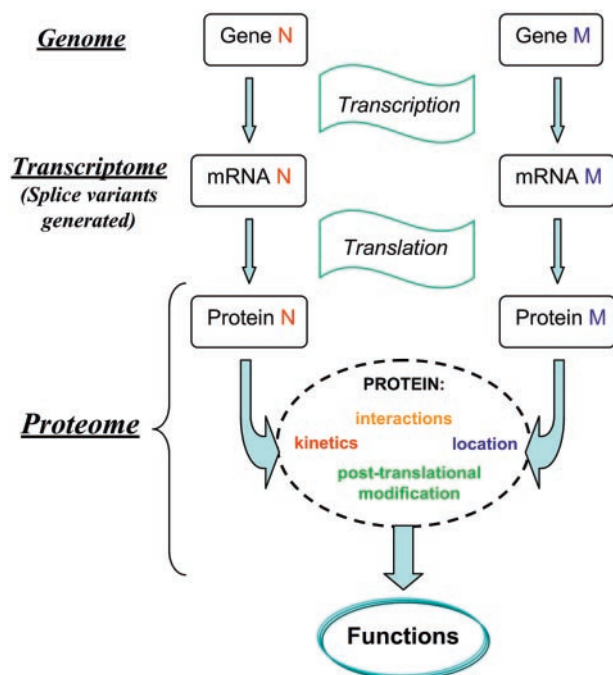
Supported by grants from Leukaemia Research Fund United Kingdom.

The online version of the article contains a data supplement.

An Inside *Blood* analysis of this article appears in the front of this issue.

**Reprints:** Anthony Whetton, Leukaemia Research Fund Proteomics Facility, Department of Biomolecular Sciences, UMIST, Manchester, M60 1QD, United Kingdom; e-mail: tony.whetton@umist.ac.uk.

© 2004 by The American Society of Hematology



**Figure 1. Progression from genome to proteome.** When a gene is expressed, the coding DNA strand is transcribed into an mRNA, which is edited by intron excision and the joining of exons. At the transcriptome level, the study of mRNA expression by a genome at a given time is routinely performed using microarray analysis.<sup>89</sup> Proteins are synthesized and may undergo cotranslational and posttranslational modification processes that are often involved in the formation of the functionally active structure of the protein. A given mRNA sequence can give rise to more than one protein. This figure illustrates aspects of this process. Although some of the observed properties of an organism can be correlated with the activity of a single gene, most of the time these are determined by the joint action of many gene products. It is only at the proteome level that the gene exerts its "function." The operation and the functions of a living cell are usually the result of proteome dynamics. Protein-protein interactions are often responsible for the regulation of cellular metabolism (enzymes), maintenance of architectural features (structural proteins), and transfer and processing of information (signal and regulatory proteins). Because proteome analysis provides a view of the biologic processes at their level of occurrence, proteomics offers a better understanding than genomics of cell cycle, cell death, development stage, cell function, and cellular responses to external stimuli and disease. Proteomics has become an important step in the development and validation of diagnostics and therapeutics.<sup>90</sup>

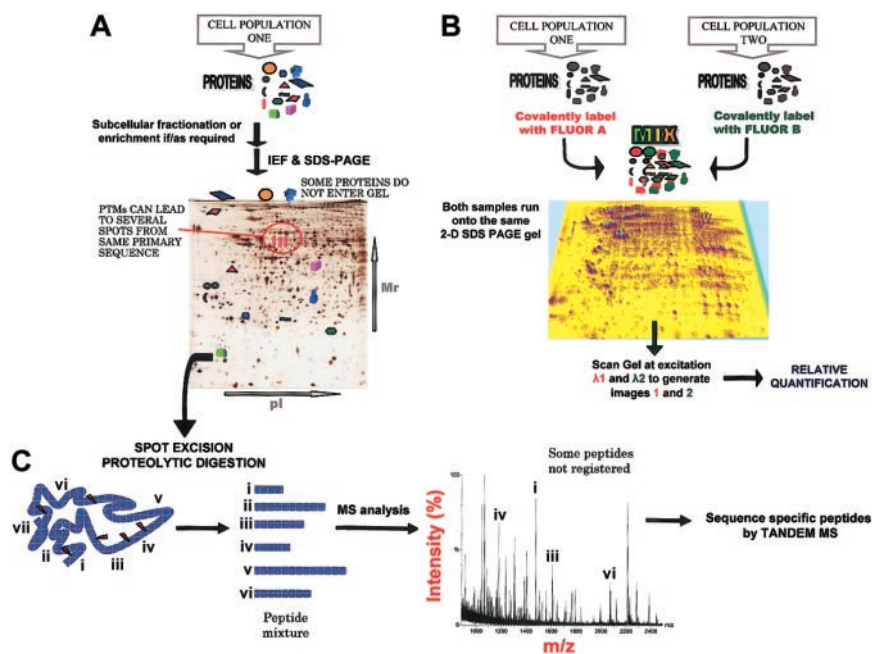
markedly as one of the core technologies for the analysis of complex protein mixtures extracted from biologic samples. The proteins are separated in 2 steps according to 2 independent properties (isoelectric point [pI] and molecular weight [MW]). Görg and coworkers<sup>9</sup> initiated the use of immobilized pH gradient (IPG) strips for the first dimension of 2-DE (isoelectric focusing), increasing the reproducibility and resolution of the separation (Figure 2A).

Visualization methods for protein detection following 1-DE or 2-DE separation represent a critical step in quantitative proteome analysis.<sup>10</sup> Methods available vary in limit of detection, dynamic range, and compatibility with analysis by MS (Table 1). Although autoradiography, following prelabeling of the cells with <sup>35</sup>S or <sup>32</sup>P, is the most sensitive detection method, with a wide dynamic range, this technique is restricted to living cells and is not suitable for tissue samples. Detection of proteins using fluorescence labeling, such as SYPRO Ruby,<sup>11</sup> has gained increased popularity and offers a wide linear dynamic range, detection of nanogram amounts of protein, end-point staining times (staining times can be varied without overdeveloping) resulting in higher reproducibility, and has the advantage of not requiring viable cells capable of taking up radiolabeled chemicals. Equivalently, the use of succinimidyl

esters of the fluorescent cyanide (Cy) dyes to label  $\alpha$  amines of peptides and the  $\epsilon$  amino groups of lysine residues on proteins prior to 2-DE is a welcome advance offering relatively good sensitivity. Fluorescent Cy3, Cy5, or Cy2 dye-prelabeled protein samples can all be run on the same 2-DE gel, allowing for intragel relative quantification of proteins from 2 or 3 preparations, a technique known as difference in-gel electrophoresis or DIGE<sup>12</sup> (Figure 2B). We have used this approach with chronic myeloid leukemia CD34<sup>+</sup> cells using  $1 \times 10^6$  cells/sample, visualizing approximately 1000 protein spots per 2-DE gel. This cell population (> 95% CD34<sup>+</sup>) gave intersample variation in spot patterns (< 5% of total spot number) in 6 samples analyzed. Changes of as little as 30% in protein expression levels could be reproducibly detected (S. Griffiths, C. A. Evans, and A.D.W., unpublished observations, August 2003), offering a sensitive means of detecting proteomic changes. The DIGE approach provides an instrumentally undemanding alternative to methods that use differentially stable isotope labeling and can be coupled to on-line liquid chromatography-tandem MS for the detection of drug or growth factor effects on specific proteins in cells.<sup>13</sup> A new cyanide dye derivative that labels cysteine residues has now been produced and offers enhanced sensitivity<sup>12</sup> (Table 1); the effect of variable cysteine content of different protein types must, however, be noted.

Despite these advances, the use of 2-DE still remains an object of debate concerning its value in proteomics research. Detractors argue it is cumbersome, time-consuming, and lacking in automation, whereas to others it remains an efficient way to separate complex protein mixtures. Issues of reproducibility, sensitivity, and protein losses during extraction for MS analysis have precluded its wider acceptance.<sup>9</sup> The undisputed need for replicate studies (at least 3 replicate studies per sample) for the assessment of reproducibility of protein pattern and relative quantity adds to the time necessary for the assessment and interpretation of data. Nonetheless, software available for relative 2-DE spot pattern analysis is sufficiently sophisticated to allow relative quantification from DIGE or other experiments (Figure 2 and Table 1, respectively). Using such software, 2-DE gels can be used to identify significant differences between patient samples, such as chronic lymphocytic leukemia patients with mutated and unmutated immunoglobulin heavy-chain loci, respectively.<sup>14</sup> Use of 2-DE can separate the thousands of plasma or serum proteins, including posttranslational variants of some proteins. The requirement for higher resolution separation can be achieved using several 2-DE gels with overlapping narrow pH gradients that reduce the presence of multiple proteins per spot, such as differentially posttranslationally modified proteins.<sup>15</sup> The use of narrow IPGs also allows for an increased loading capacity, and therefore improves the detection of low-abundance proteins. A concerted effort to map the human plasma proteome (including the 2-DE approach) is now underway in the belief that further proteins of diagnostic and therapeutic value will be identified.<sup>16</sup>

Further application of gel-based technologies in hematology can be seen in a variety of different studies.<sup>17,18</sup> Prelabeling cells with <sup>32</sup>P has been used to recognize proteins phosphorylated as a consequence of cytokine action.<sup>19</sup> No assumptions are made about the phosphorylated proteins prior to identification. For example, 25 new substrates for mitogen-activated protein (MAP) kinase have been recognized using an approach based on 2-DE.<sup>20</sup> Recently, a fluorescent dye that reversibly associates with phosphorylated proteins within gels has been characterized.<sup>21</sup> Preliminary experiments in our laboratory (C. A. Evans, R. Unwin, unpublished observations, July 2003) indicate that it does have value as an initial screen for agonist-stimulated phosphorylation events in



**Figure 2. The 2-DE gel-based methods in proteomics.** (A) Most techniques currently used in proteomics involve the separation of the vast number of proteins present in a cell or tissue at a given time prior to analysis by MS and recognition and characterization using bioinformatics techniques. The protein separation can be performed at the protein or peptide level. One widespread methodology is 2-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (2-D SDS-PAGE), which separates proteins according to their isoelectric points and molecular weights. One advantage of this approach is the ability to separate differentially posttranslationally modified forms of the same protein. A key disadvantage is that complete coverage of the proteome cannot be achieved, because some proteins will not enter the gel. Relative quantification of proteins between samples was in the past achieved by intergel comparisons following staining (Table 1). (B) The recently deployed technique of fluorescence 2-dimensional difference gel electrophoresis (2-D DIGE) for the relative quantification of proteins from up to 3 cell states on the same gel offers a greater reproducibility than previous approaches. Prior to the 2-D SDS-PAGE separation, the samples are covalently labeled with succinimidyl esters of different cyanide dyes (Cy2, Cy3, and Cy5). The *N*-hydroxy-succinimidyl esters undergo nucleophilic substitution reaction with the lysine  $\epsilon$  amine groups to give an amide. The samples are then mixed together and separated on the same 2-D SDS-PAGE gel. The gel is scanned at the different excitation frequencies, the images merged, and the difference in protein abundance calculated. However, to maintain the solubility of the proteins during the electrophoretic separation, 2-D DIGE<sup>12</sup> requires that only 1% to 2% of the lysine residues be derivatized (the derivatization increases the hydrophobicity of the proteins). Recently, a new set of dyes intended to fluorescently label all cysteine residues within a protein have been made available and these offer greater sensitivity.<sup>91</sup> (C) Following separation, the protein spot is usually excised, subjected to in-gel enzymatic proteolysis, and analyzed by MS, usually MALDI-MS. The MALDI peptide fingerprint is often sufficient for confident protein recognition. For further confirmation, amino acid sequence information, or recognition of low-abundance proteins, tandem mass spectrometry (MS/MS) analysis is performed, generally using electrospray (ES) ionization. A prevalent method that separates proteins at the peptide level is liquid chromatography (LC), usually used in conjunction with MS/MS (LC-MS/MS).

hematopoietic cells, obviating the need for prelabeling and opening the possibility of performing such studies on freshly isolated primary cells. The sensitivity and specificity of the technique, however, will be key issues. The experimental simplicity of this technique contrasts with alternative strategies such as the more quantitatively rigorous approach taken by Mann's group to signal transduction proteomics.<sup>22</sup> Untreated HeLa cells, grown in medium containing [<sup>12</sup>C]-arginine, were mixed with cells treated with epidermal growth factor (EGF) after culture in [<sup>13</sup>C]-arginine-containing medium. A total of 28 proteins that selectively bind the SH2 domain of the adaptor protein Grb2 were recognized using affinity purification, followed by liquid chromatography (LC) separation and analysis by MS. The levels of the proteins from the 2 samples were compared using the inherent isotope coding of all the proteins (see "MS-based quantitative proteomics"). This metabolically stable isotope-labeling technique can, in theory, be applied to any mammalian cell line under 2 different conditions. It therefore lends itself to the study of hematopoiesis.

Using 2-DE gels and standard staining methods, myeloid cell development can be followed in terms of multiple quantitative changes in protein composition.<sup>6</sup> Similarly, determination of the change in the proteome of lymphocytes as they prepare for antibody production provides an elegant example of the application of 2-DE to a specific research issue.<sup>23</sup> Cells were shown to increase expression of metabolic enzymes and endoplasmic reticulum components sequentially, prior to engaging on the

generation of their secretory product. This study illustrates the value of proteomics methods in generating a single multifaceted picture of cellular events, providing prompts for more detailed investigations. A 2-DE proteomics approach was also used for the identification of potential markers for lymphoma classification, eosinophil activation, and recognition of isoforms of proteins involved in apoptotic suppression.<sup>24-27</sup>

These studies focused on the proteome assay of whole-cell lysates. There is a growing recognition that prefractionation of the sample offers substantial analytical advantages.<sup>7,8</sup> Removal of highly expressed proteins presents opportunities for increasing dynamic range and thereby improving the detection of proteins expressed at lower levels. Thus, an important advance, increasing the prospect of protein identification in organelles, is subcellular fractionation. Studies on enriched organelles allow the identification of proteins with no previously recognized organellar function.<sup>28,29</sup> With respect to hematopoietic cells, phagosomes have been isolated from macrophages and over 140 proteins recognized in a 2-DE/MS study.<sup>30,31</sup> Localization of proteins using techniques such as green fluorescent protein tagging and high-resolution imaging confirm that this is a robust method for assigning proteins within complexes or organelles; this approach to organelle analysis is fully reviewed by Dreger.<sup>7</sup> By extension, relative quantification of low-abundance proteins will be achieved more effectively using prefractionation techniques. Nonetheless, any such study is limited by the failure of some hydrophobic, large, or basic proteins to enter 2-DE gels.

**Table 1. Methods of protein detection for gel-based proteomics**

Visualization method	Limit of detection, ng	Dynamic range	Detection	No. of samples/gel	Comments
Colloidal Coomassie blue	8-10	20-fold Linear response restricted to high ng amounts of protein	Densitometry	1	Compatible with analysis by MS. This method, which uses trichloroacetic acid and alcohol in the staining solution, results in esterification of aspartic and glutamic side chain carboxyl groups, complicating the interpretation of the mass spectra.
Silver	2-10	8- to 10-fold Linear response restricted to low ng amounts of protein	Densitometry	1	Some protocols compatible with analysis by MS. Staining times and reaction temperatures are critical for reproducibility.
Zinc-imidazole	5-10	Linear response restricted to high ng to $\mu$ g amounts of protein	Densitometry	1	Compatible with analysis by MS. Reverse staining allows higher protein recovery. Rapid (5-15 minutes). Sample overloading decreases the capacity of distinguishing between different bands (1-DE) or spots (2-DE).
Cy3, Cy5	5-10	1000-fold Linear response on wide range	Fluorescent	2-3	Compatible with analysis by MS. Allows for both intragel and intergel relative quantification of protein spots from samples.
SYPRO Ruby	1-8	1000-fold Linear response on wide range	Fluorescent	1	Compatible with analysis by MS. Rapid. Staining time is not critical and can be varied from experiment to experiment without overdeveloping.

For proteome analysis, the most important criteria to be considered when evaluating methods of protein detection are limit of detection, dynamic range, homogeneity, and compatibility with MS analysis. The properties of some of the commonly techniques are shown. Most proteomics studies require limits of detection in the low nanogram protein range, which would correspond to the femtomole levels usually detectable by MS analysis (eg, 1-20 ng, corresponding to 100 fmol of a 10-200 kDa protein). A wide linear dynamic range is essential when the aim of the study is relative quantification between samples. The homogeneity of the detection refers to the uniformity of staining between different classes of proteins. When the method targets only one amino acid, proteins that do not contain that particular residue are not detected (eg, cysteine). An ideal detection method would target all amino acid residues equally, so that the signal detected would be directly proportional to the length of the amino acid chain. The compatibility of the method of detection with MS analysis is particularly important for the recognition of the proteins. The presence of formaldehyde or glutaraldehyde in various solutions used for silver staining (fixing, silver impregnation, and developing) interferes with protein recognition because these aldehydes alkylate the  $\alpha$  and  $\epsilon$  amino groups of proteins.<sup>10</sup> Silver staining methods compatible with MS analysis were developed by removing the glutaraldehyde and decreasing the use of formaldehyde in the stain.<sup>99</sup> Although most detection methods allow only intergel comparison, the new generation of cyanide dyes that chemically modify lysine or cysteine residues offer the potential for multiplexing in gel-based experiments (see "Separation procedures and gel electrophoresis-based quantification for proteomics" and Figure 2) while not compromising downstream MS analysis. Regardless of the detection method, the principal method for relative quantification is densitometry analysis. Providing that there is a good reproducibility between the 2-DE protein patterns, densitometry is an important tool in proteome analysis because it allows the simultaneous detection and quantification of hundreds of proteins from a multitude of cell states. A number of computer-assisted image analysis software packages that allow the comparison of 2-DE protein patterns, such as PDQuest (Bio-Rad, Hercules, CA), Melanie (Bio-Rad), and Progenesis (Non-linear Dynamics, Durham, NC), are now commercially available. However, the technique suffers from a number of limitations, either inherited from 2-DE's lack of reproducibility and poor dynamic range of staining methods, or, depending on which software is used for the analysis, from its own poor sensitivity of detection. New software and methods constantly improve protein detection and image warping, but the success of the matching still depends extensively on the similarity of the protein patterns. For in-gel comparison, DIGE experiments plus the appropriate analytical software therefore offer an effective way forward in relative quantification of the hundreds of proteins found in 2-DE gel experiments.

A nonquantitative approach to identification of such proteins within biologic material is exemplified by the work of Link et al<sup>32</sup> who submitted a protein complex directly to enzymatic digestion and characterized the resulting complex mixture of peptides by 2-dimensional chromatographic (strong cation exchange-reversed phase) separation and electrospray-tandem MS (ES-MS/MS) analysis (see next section). The resulting product ion spectra were automatically correlated with predicted amino acid sequences in translated genomic databases. This method was further developed by the same group<sup>33</sup> to achieve full automation. The technique, termed multidimensional protein identification technology (Mud-Pit), involves not only the separation of peptide mixtures but direct MS analysis and database searching. This study achieved in a single experiment a much higher number of detected and recognized proteins than with a 2-DE study. This large-scale identification, loosely termed "shotgun" proteomics, has a throughput higher than 2-DE/MS approaches and can deliver high-quality data but

requires sophisticated and automated LC-MS/MS apparatus and an acceptance that few conclusions can be drawn on the amount of any protein/peptide found in the sample.

## Protein recognition using MS

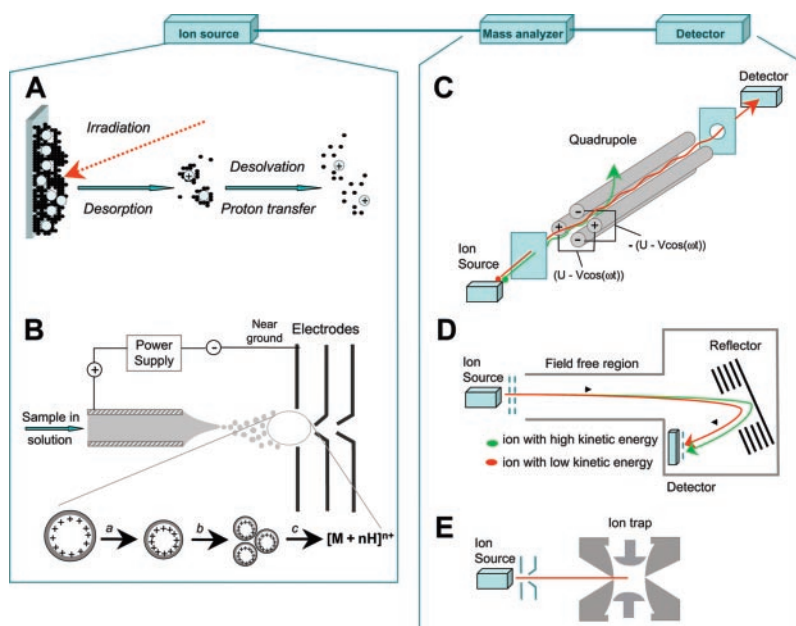
The recent advances in proteomics were mainly driven by the increasing ability of MS to detect and characterize low levels of proteins. MS-based proteomics has a growing role in biomedical research where limited sample material is available, as femtomole sensitivity is routinely achieved.<sup>34</sup> In hematology, MS has been used for protein, peptide, and DNA sequencing,<sup>6</sup> protein-folding studies, noncovalent interaction analysis, analysis of PTMs,<sup>35</sup> identification of novel proteins,<sup>36</sup> and in a variety of diagnostic and drug discovery projects.<sup>16,37-40</sup>

The main components of a mass spectrometer are an ion source, one or several mass analyzers, and a detector.<sup>34</sup> Electrospray (ES; Figure 3B) and matrix-assisted laser desorption/ionization (MALDI; Figure 3A) processes provide “soft” (meaning low-energy) ionization methods for a variety of biomolecules, including peptides, proteins, drug metabolites, oligonucleotides, and carbohydrates, to enable their measurement by MS. There are several types of commercially available mass spectrometers that combine ES or MALDI with a variety of mass analyzers. In simple terms, gas-phase ions, produced in the ion source, are introduced into the mass analyzer and differentiated according to their mass/charge ( $m/z$ ) ratio on the basis of their motion in a vacuum under the influence of electric or magnetic fields. Figure 2C illustrates a typical experiment in which the identity of a protein from a 2-DE gel can be discerned using peptide mass fingerprinting by MALDI time-of-flight (ToF) MS, or tandem MS (MS/MS).

### SELDI-ToF MS and the search for disease markers

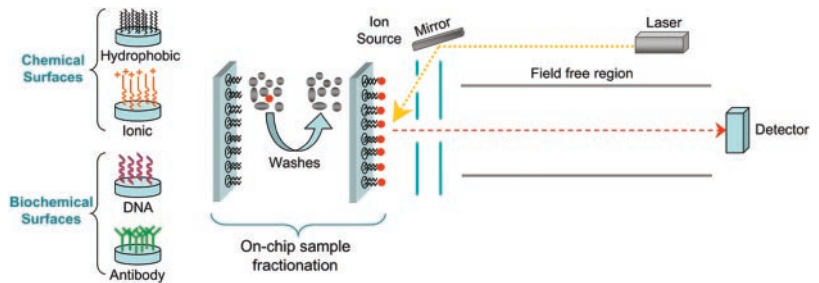
The technique (Figure 4) has at its heart a means of enriching for proteins with certain chemical characteristics that determine their

interaction with a set of specific surfaces used as a laser desorption ionization target.<sup>41</sup> As such, it is often cited as an array technology (see “Protein arrays”). Unlike other MS-based strategies (Figure 2), the SELDI approach does not generally include enzymatic digestion of proteins, so that detection of intact (generally small) proteins is involved. One advantage of this chip-based technique is that crude samples can readily and rapidly be analyzed with high throughput. The key disadvantage is that the mass spectrum obtained does not enable identification of the proteins analyzed and further work is therefore required. The obtained data are analyzed to recognize critical features that may be used as markers for diagnosis or prognosis. In a landmark study, a training set of serum SELDI-ToF spectra from 50 unaffected women and 50 patients with ovarian cancer was analyzed by an iterative algorithm that identified a proteomic pattern to discriminate healthy patients from those with malignant disease, using the minimum amount of information.<sup>42</sup> A similar study on prostate cancer and lung cancer reveals the general applicability of this technique.<sup>43,44</sup> Furthermore, there are indications that determination of a serum protein level (eg, high-density lipoproteins) can be monitored effectively and straightforwardly with SELDI-ToF MS<sup>45</sup> and therefore, indicators of hematologic disease could theoretically be measured in the same



**Figure 3. MS instrumentation used in proteomics.** The fundamental principle of MS analysis involves the conversion of the subject molecules to either cations or anions in the ion source, separation according to their mass/charge ( $m/z$ ) ratios in the mass analyzer, and subsequent detection. Several configurations of mass spectrometers that combine ES and MALDI with a variety of mass analyzers (linear quadrupole mass filter [Q], time-of-flight [ToF], quadrupole ion trap, and Fourier transform ion cyclotron resonance [FTICR] instrument) are routinely used. (A) In MALDI, the sample is embedded in a large excess of a matrix that has a strong absorption at the laser wavelength. Following the laser irradiation of the sample surface, the matrix accumulates a large amount of energy that is thought to initiate the proton transfer between the matrix and the analyte compound to form ions. The ions observed during MALDI MS are mainly singly charged. This results in simple spectra (because the  $m/z$  ratio is what is being measured) even in the case of analysis of mixtures but can be a disadvantage for peptide sequencing, which requires the achievement of peptide fragmentation, a process that preferentially occurs with multicharged peptides. The concurrent and independent development of ToF mass analyzer and its compatibility with this ionization method resulted in the rapid development of MALDI-ToF as a routine analytical mass spectrometer.<sup>92</sup> Other configurations, however, have also been successfully applied for the analysis of peptides and proteins, such as MALDI-ion trap,<sup>50</sup> MALDI-QqToF<sup>52</sup> (where q represents a quadrupole ion decomposition region), and MALDI-ToF-ToF.<sup>51</sup> This type of ionization method is also used in SELDI-ToF MS. (B) ES is a continuous nebulization process that produces ions directly from solution and facilitates interfacing of LC with MS. A sample is passed with flow rates of 1 to 10  $\mu\text{L}/\text{min}$  through a capillary held at high potential relative to ground and counter electrode. The strong electric field obtained induces charge accumulation at the liquid surface situated at the end of the capillary and leads to the formation of a mist of highly charged droplets. The ES process results in the formation of multiply charged ions. This is one of the principal advantages of this method because it allows the analysis of ions from high-molecular-mass molecules, such as proteins and peptides, using mass spectrometers of limited  $m/z$  ratio range. (C) The quadrupole analyzer is frequently termed a mass filter because it transmits only ions within a narrow  $m/z$  range. The analyzer uses the stability of the trajectory to separate these ions according to their  $m/z$  ratio. The stability of the oscillating trajectories of the ions is based on the joint application of direct current and radio-frequency voltages on 4 parallel cylindrical metal rods. (D) Unlike the scanning devices such as quadrupoles, ToF analyzers separate ions temporally rather than spatially. The ions are rapidly accelerated into a “field-free” drift region also called a “flight tube,” and their separation is achieved by measuring the difference in transit time from the ion source to the detector. The reflectron, which consists of a series of rings or grids that act as an ion mirror, compensates for the initial kinetic energy distributions of ions. (E) The ion trap analyzer captures the ions, which collide with the helium “bath” gas and start to oscillate in a predicted motion. Ion trap can be used as a “tandem-in-time” instrument as selection, fragmentation, and analysis of ions take place in the same space. Fourier transform ion cyclotron resonance mass spectrometer (FTICR; not illustrated) is also a trapping device, in this case by using strong magnetic fields, and offers great opportunities for investigating protein interactions and PTMs, with high sensitivity, mass accuracy, and resolution.<sup>93</sup>

**Figure 4. Principle of surface-enhanced laser desorption/ionization (SELDI)-ToF MS.** This is a technique for the enrichment of proteins with specific chemical characteristics and combines chromatography with MS. ProteinChip arrays have been designed that contain chemically or biochemically treated surfaces. The sample, consisting of crude extracts or mixtures of whole proteins, is applied to the surface. After a series of washes, the targeted proteins are selectively retained. An energy-absorbing solution is added to the surface and the sample is subjected to laser desorption/ionization. The formed ions are measured using a ToF mass analyzer as previously described. Characteristic features of spectra can be used for prognostic purposes ("SELDI-ToF MS and the search for disease markers").



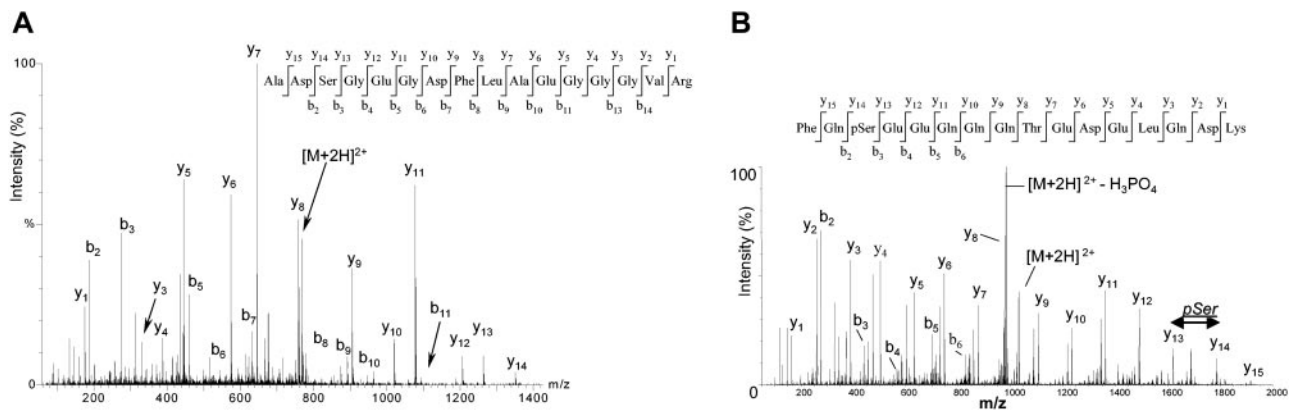
fashion. In an adaptation of the technique thin slices of solid tumors are also being used as targets for MALDI-ToF MS. The application of this mass spectrometric approach hopefully will deliver a raft of new clinical indicators and surrogate markers for response. A considerable effort, initiated by Caprioli and coworkers, has been applied for the direct analysis of tissue slices by MALDI MS.<sup>46</sup> More recent work, such as the study of non-small-cell lung tumors,<sup>47</sup> showed the potential of this technique for the detection of disease markers.

## Sequence analysis of peptides and proteins using MS

Tandem mass spectrometry (MS/MS) allows for protein and peptide sequencing and has therefore evolved as an indispensable tool for protein recognition, structure elucidation, and characterization of PTMs. Figure 5 shows examples of ES Q-ToF MS/MS analyses of peptides for amino acid sequence information (A) and detection of site of phosphorylation (B). Sequence analysis of peptides by collision-induced dissociation (CID) during MS/MS<sup>48</sup> is a favorable alternative to the classical sequencing performed by the Edman degradation technique.<sup>49</sup> In contrast to Edman degradation, which remains valuable for

N-terminal sequencing of large amounts of relatively pure peptides, MS/MS analysis can be used for the sequencing of peptides present in mixtures or blocked at the N-terminus of the sequence. The recent advances in coupling a MALDI ion source to ion trap,<sup>50</sup> ToF/ToF,<sup>51</sup> and Q-ToF<sup>52</sup> mass spectrometers provide enhanced opportunities for high-throughput analyses.

The development of computer algorithms that can correlate the data obtained from MS and MS/MS analyses with databases enables one to use either peptide mass fingerprint information or MS/MS CID data to recognize a protein.<sup>53,54</sup> Recently developed algorithms allow the search of uninterpreted or partially interpreted MS/MS spectra,<sup>55,56</sup> for example, SEQUEST,<sup>55</sup> Mascot, PepFrag, Sonar, XProteo, and MS-Tag. The programs search for all peptides present in databases that have the same mass as the precursor ion selected by the user, predict the principal fragment ions that would be expected for these database peptides, and then match these with the experimental data. In this way the original peptide can be identified in terms of amino acid sequence and the protein from which it originated deduced. The confidence in the result obtained can be increased using a variety of peptide derivatization methods<sup>57,58</sup> designed to obtain a specific set of parameters to be included in the algorithms for database search, which would reduce the ambiguity of the sequence.



**Figure 5. Sequencing of peptides using MS/MS.** Peptide fragmentation by collisionally activated dissociation (CAD) during MS/MS<sup>48</sup> allows the recognition of complete or partial peptide sequence. Following collisional activation, the peptides undergo fragmentation along the backbone or in their side chains giving rise to product ions that can be attributed to the amino acid sequence. The nomenclature for classifying these ions allows the labeling of all the various fragment ions. The 3 possible cleavage points of the peptide backbone are called  $a_n$ ,  $b_n$ , and  $c_n$  when the charge is retained at the N-terminal fragment of the peptide, and  $x_m$ ,  $y_m$ , and  $z_m$  when the charge is retained by the C-terminal fragment. The subscripts  $n$  and  $m$  indicate which amide bond is cleaved counting from the N- and C-terminus, respectively, and thus also the number of amino acid residues contained by the fragment ion. Immonium ions and internal fragments are usually labeled with the letter codes of the amino acids (not indicated in this figure). Product ions derived from fragmentation at the side chain of the peptide (labeled  $d$ ,  $v$ , and  $w$ ) are not discussed here. Panel A illustrates the ES Q-ToF product ion spectrum of  $[M+2H]^{2+}$  at 768  $m/z$ , corresponding to fibrinogen  $\alpha A$  from rat hepatocytes, following digestion with trypsin, where  $[M+2H]^{2+}$  denotes the doubly charged precursor ion selected for fragmentation. An almost complete series of  $b$  and  $y$  ions was obtained in this analysis resulting in a facile identification of the peptide sequence. Panel B illustrates the application of MS/MS analysis for the search for sites of phosphorylation. The product ion spectrum, recorded in positive mode, was obtained from the 1-dimensional liquid chromatography-tandem electrospray mass spectrometry analysis (LC-MS/MS) of a phosphorylated peptide. The signals between 1200 and 2000  $m/z$  are magnified by a factor of 2.5. The sites of phosphorylation are searched by scanning for loss of phosphate or for characteristic immonium ions (Figure 7).  $[M+2H]^{2+}$  denotes the doubly charged precursor ion selected for fragmentation, and the  $[M+2H]^{2+}-H_3PO_4$  denotes the characteristic loss of the phosphate moiety from the precursor ion.

## MS-based quantitative proteomics

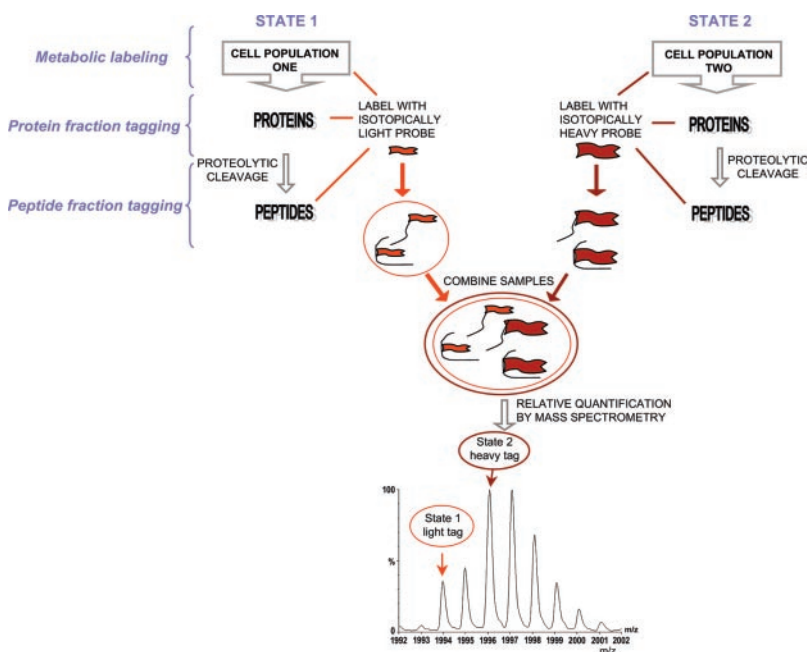
The methods most commonly used for quantitative proteome analysis are relative quantification methods.<sup>59</sup> Although densitometry or fluorescent analysis of spot intensity following 2-DE gel staining have an important niche in cell biology, a more versatile technique for precise relative quantification involves the differential labeling of each set of proteins or peptides derived from 2 different cell states with light and heavy isotopes of the same chemical reagent and subsequent MS analysis<sup>59,60</sup> (Figure 6). This approach allows for relative quantification of basic, hydrophobic, or large proteins excluded from analysis using 2-DE or DIGE. Two derivatized samples are combined and analyzed by MS, which allows the calculation of the relative abundance levels of the proteins. This method is based on the assumption that the structurally similar isotopically labeled peptides have identical ionization efficiencies during the mass spectrometric analysis. The light- and heavy-labeled peptides appear in the mass spectrum as doublets, and the peak heights illustrate the relative abundance of the protein in each cell state. In MS the intensity of a specific peptide signal is dependent on the chemical and physical properties of the analyte, the associated solvent, and other factors. The signals from 2 peptides, even from the same protein, will therefore differ in intensity. Only the use of a direct comparison of chemically similar entities can yield relative quantification. The isotopic labels can be incorporated either during cell culture,<sup>61,62</sup> or at the protein or peptide level.

In our laboratory, we have used guanidination with [<sup>14</sup>N]- and [<sup>15</sup>N]-*O* methylisourea for a combined increased sensitivity of detection of lysine C-terminus peptides, relative quantification of 2 cell states, and de novo sequencing.<sup>63</sup> The 2-Da mass difference between the light and heavy isotope-labeled peptides allows for simultaneous selection and fragmentation of both isotopomers during MS/MS and assignment of C-terminus fragments is therefore readily achieved by virtue of their presence as doublets in the product ion spectra. Alternatively, the differential derivatization can be followed by an enrichment

step prior to the analysis by MS. Enzymatic digestion of proteins with either H<sub>2</sub><sup>16</sup>O or H<sub>2</sub><sup>18</sup>O, followed by immobilized metal-affinity chromatography, dephosphorylation, and mass spectrometric analysis has been successfully applied for the relative quantification of protein phosphorylation.<sup>64</sup> Proteins from 2 cell states can be labeled and separated using an isotope-coded affinity tag (ICAT) approach.<sup>13</sup> The chemical reagent initially used in this method consisted of a protein reactive group, a linker region, and a biotin tag and its use is demonstrated graphically (see the information available online at the *Blood* website; use the Supplemental Materials link at the top of the online article). The mixture of the 2 derivatized proteins is digested and separated by affinity extraction prior to tandem mass spectrometric analysis. The initial complexity of the mixture is reduced during the affinity extraction step.<sup>65</sup> ICAT technology has been applied for the separation and quantification of proteins with extreme pI values, low abundance,<sup>33</sup> and membrane-bound proteins.<sup>60</sup>

Although it is sometimes suggested that ICAT frees one from the constraints of the lesser technology of 2-DE, a recent study comparing ICAT and 2-DE demonstrated that neither offers comprehensive coverage of a proteome.<sup>66</sup> A full evaluation of the relative quantification offered by these 2 techniques is required and we are presently undertaking such a study. The use of an ICAT approach for the identification and characterization of 491 microsomal proteins and determination of changes in expression during development of the promyelocytic cell line HL-60 demonstrates the power of this technique, especially because many of the proteins identified were hydrophobic. Nonetheless, such studies are not a casual undertaking.<sup>60</sup> A paradigm study on the leukemogenic oncogene product Myc by Shii et al<sup>67</sup> shows the broad effects of a single protein on pathways as diverse as motile control and protein synthesis.

Recently, methods that use synthetic peptides, designed to mimic some of the naturally occurring peptides following the enzymatic digestion of the proteins of interest, have been developed for absolute quantification of proteins.<sup>68,69</sup> The synthetic peptide, labeled with a stable isotope, is mixed with the sample



**Figure 6. Isotopic labeling for relative quantification using MS.** Proteins or peptides derived from 2 different cell states are derivatized with light and heavy isotopes of the same chemical reagent. The samples are then combined and analyzed by MS. The relative abundance levels of the proteins are calculated by comparing the peak heights of the light- and heavy-labeled peptides. The isotopic label can be incorporated in vivo, in vitro, at the protein, or at the peptide level. In cell culture, the incorporation was achieved by using media that included the isotopic labels, such as <sup>15</sup>N or <sup>14</sup>N media, minimal medium, or minimal medium with isotopically labeled leucine (leuD10).<sup>61,62,94,95</sup> The stable isotope-labeling approach has also been used for detecting the numbers of the tagged amino acid residues and, therefore, for unambiguously recognition of the protein.<sup>62,95</sup> A method that gained increased popularity is the isotope-coded affinity tag (ICAT) approach.<sup>13</sup> The label contains a thiol group (which reacts with cysteine residues), 8 hydrogens (light) or deuteriums (heavy), used for relative quantification, and a biotin group that is selectively recognized during the affinity extraction step by an avidin moiety attached to the chromatographic column. Proteins from 2 cell states are derivatized with either the heavy or the light isotope and the mixture is digested and separated by affinity extraction prior to tandem mass spectrometric analysis. The ICAT analysis with the initial reagent used was complicated by the fact that heavy-labeled peptides elute earlier than the light-labeled peptides during LC separation, the so-called deuterium effect.<sup>96</sup> Newer cleavable <sup>13</sup>C-labeled ICAT reagents have significant advantages over the original reagents.<sup>97</sup>

prior to enzymatic digestion and analysis by LC-MS/MS. The peptide can be synthesized to contain a variety of modifications (such as phosphorylation) to allow quantification of PTMs. This will be a key aid in future pharmacokinetic studies and more generally in studies on signal transduction and other aspects of cell studies requiring absolute quantification.

---

## Protein arrays

Abundant proteins can obscure the quantification of lower abundance proteins, such as signaling molecules or kinases. This problem can be reduced using ICAT and MudPit or prefractionation prior to 2-DE.<sup>28-30,70</sup> Protein microarrays offer a different solution and have the potential for high-throughput applications to identify novel drug targets and disease markers and a sensitivity born of new technical approaches. The sequencing projects offer opportunities to manufacture protein chips that embrace the thousands of proteins encoded in the genome. Zhu et al in a landmark paper<sup>71</sup> used gene tagging for enrichment of thousands of proteins from yeast to create a protein array and used this to screen for calmodulin- and phospholipid-binding proteins. This study is but one in which key issues of the appropriate surface for protein attachment, the maintenance of native protein folding, and the methods of detection for interacting agents have been addressed. Exciting advances include application of planar waveguide technology with which the measurement of as few as 500 cytokine molecules can be achieved,<sup>72</sup> as reviewed by Mitchell.<sup>73</sup> Advances in surface chemistry, microfluidics, and detection techniques will make rapid analysis of proteins in samples a reality in the future.<sup>74,75</sup>

Protein array technologies using discrete sets of known proteins are plainly of potential value in diagnostic or prognostic testing. A microarray of 60 antibodies has already been constructed for cluster of differentiation (CD) antigens by binding antibodies to a nitrocellulose-coated glass slide. The results achieved with this antibody array compared well with flow cytometric assessment of CD expression on a number of leukemic cell types and normal leukocytes.<sup>76</sup> One strength of this approach is the relatively high number of markers that can be assayed per experiment compared to flow cytometry. However, discrete binding to subsets of cells is not easily assessed. The creation of high-density antibody microarrays using phage display technology and other methods, reviewed by Hebestreit,<sup>77</sup> remains a realistic proposition for the characterization of the proteome. Development of appropriate arrays for concerted proteomics research projects will continue and, in the future, simultaneous measurement of many proteins from limited volumes of samples of complex biologic mixtures (such as serum or cell lysates) will be possible. Similarly, proteins can be arrayed to search for interacting partners. For example, Procognia (Maidenhead, Berkshire, United Kingdom) has arrayed p53 proteins with germline single-nucleotide polymorphism (SNP), or functional potassium regulatory channels; these arrays identify drug binding and effects on proteins.

Other approaches are being derived for global analysis of protein function and these include activity-based protein profiling (ABPP) for enzymes such as tyrosine phosphatases using selective modification of specific enzymatic active sites.<sup>78</sup> Aptamers, which are single-stranded oligonucleotides that bind to proteins of interest, also offer a route to array-based

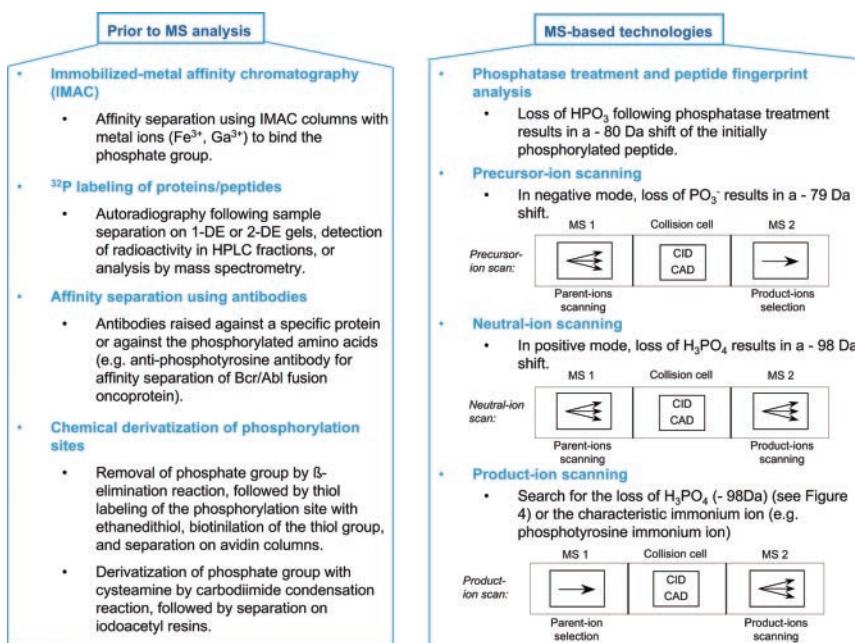
techniques, for example, in protein quantification. Directed evolutionary chemistry processes using systematic evolution of ligands by exponential enrichment (SELEX) identify aptamers by starting with a combinatorial library and the protein of interest (tethered to an appropriate substratum), a repetitive process,<sup>79</sup> which leads to the creation of a DNA or RNA molecule with a high affinity for a specific protein. There may be therapeutic potential; for example, a thrombin aptamer may have clinical value as an anticoagulant.<sup>80</sup> Somalogic (Boulder, CO) has developed aptamer based protein arrays for the measurement of thousands of proteins from a single sample.

---

## PTM of proteins

Following synthesis, the protein complement of a cell is subjected to a barrage of reversible or permanent modifications that control protein activity and interactions. To date, over 300 PTMs have been reported,<sup>81</sup> including phosphorylation, glycosylation, deamidation, ubiquitination, proteolytic processing, fatty acylation, and glycosylphosphatidylinositol lipid anchor attachment. These reactions are catalyzed by enzymes that can also be regulated by PTMs. The coining of the word "proteome" has spawned terms like "phosphoproteome" or "glycome." Inherent in these new terms is a miasma of PTMs that require systematic definition. The technologies for determination of these PTMs are in their infancy, which means throughput is slow, but significant advances are being made. The detection of PTMs using MS-based methods is challenging for several reasons. In contrast to the usual recognition of a protein, which requires the detection of only one or a few peptides, the identification of the PTM of a protein relies on the detection of the particular peptide(s) that bears the site of modification. A more complete coverage of the amino acid protein sequence can be usually achieved using individual or sequential proteolytic digestions with different enzymes. A variety of algorithms (eg, FindMod and FindPep tools, described in the Supplemental Document) have been developed that, using the peptide fingerprint of the analyzed protein, search databases for potential sites of PTMs. This technique is more reliable if it is applied following the initial identification of the protein and the search for potential PTMs is performed only using MS signals not otherwise assigned to peptide fragments. The covalent bonds between the PTMs and the peptides are of relatively low energy and therefore tend to dissociate during MS analysis, in some cases, giving rise to characteristic patterns in mass spectra that can be used for the identification of the modification.<sup>82</sup> Algorithms have been developed that search peptide fingerprints for pairs of peaks that are potential signatures of PTMs (eg, 79 Da for phosphopeptides; Figure 7). Unfortunately, most frequently, PTMs have a low stoichiometry. A variety of methods, such as affinity purification and chemical modifications, have been developed to be used prior to the MS analysis for the enrichment and separation of proteins/peptides having the modification of interest. Figure 7 illustrates some of these in the context of protein phosphorylation.<sup>83</sup> The variety of methodologies used for the analysis of protein phosphorylation has been thoroughly reviewed by McLachlin and Chait<sup>84</sup> (and references within). Affinity separation followed by MS/MS analysis for the detection of phosphotyrosine-specific immonium ions has been recently successfully used for the detection of 9 phosphorylation





**Figure 7. Overview of strategies used for the analysis of phosphorylated proteins.** This overview presents only a few of the strategies currently implemented for the analysis of phosphorylated proteins. Readers are referred to reviews that cover in greater detail the variety of technologies used for the phosphoproteome analysis.<sup>84,98</sup> The methods used prior to MS analysis are mainly concerned with the enrichment of phosphorylated proteins or peptides. (A) Immobilized metal affinity chromatography (IMAC) relies on the affinity for the phosphate group of metal ions ( $\text{Fe}^{3+}$  and  $\text{Ga}^{3+}$ ) that are bound to a chelating resin. Several limitations precluded the straightforward use of this method, such as binding of acidic unphosphorylated peptides that necessitate the esterification of the carboxylic acid groups and difficult elution of peptides with multiple phosphorylation sites. Another affinity purification method routinely used for the detection of sites of phosphorylation involves the use of antibodies against phosphorylated tyrosine residues (C). These have been found more efficient for the enrichment of phosphoproteins than phosphopeptides. Equivalent approaches for the enrichment of proteins or peptides containing phosphorylated serine or threonine are not yet in common use. MS-based techniques are mainly concerned with the selection, isolation, detection, and recognition of the site of modification. Precursor ion scanning has been typically used in a tandem quadrupole instrument. The first quadrupole is set to scan over the appropriate peptide  $m/z$  range, collision-induced dissociation (CID) is induced in the second quadrupole (or equivalent), and the third quadrupole is set to selectively transmit  $m/z$  79, corresponding to  $\text{PO}_3^-$ . Therefore, only the peptides that yield the characteristic  $\text{PO}_3^-$  product ion are detected. A conceptually similar approach may be taken to the detection of phosphotyrosine-containing peptides based on detection of the phosphotyrosine immonium ion.<sup>85</sup> Another MS/MS method useful in this context is neutral loss scanning, this time applied in positive ion mode. The first quadrupole is set to scan over the full  $m/z$  range, CID is achieved in the second quadrupole, whereas the third quadrupole is set to scan with an offset on the  $m/z$  scale of  $-98/z$  (where  $z$  represents the charge state) relative to the scan of the first quadrupole. Therefore, peptides that lose the neutral  $\text{H}_3\text{PO}_4$  (mass 98) are detected. These screening steps may be followed by MS/MS with scanning of all product ions derived from the putative intact phosphopeptide ion.

sites of Bcr/Abl fusion oncoprotein, of which 6 were novel.<sup>85</sup> Another elegant example is the sequential use of phosphotyrosine immunoprecipitation, methyl esterification, IMAC (Figure 7), and nanoflow reversed-phase high-performance liquid chromatography (HPLC) ES MS for the study of temporal changes in tyrosine phosphorylation following the inhibition of Bcr-Abl protein tyrosine kinase with imatinib mesylate in chronic myelogenous leukemia cells.<sup>86</sup> The key role of tyrosine kinases in leukemogenesis determines that such approaches will be invaluable in defining downstream targets of kinases in the future. The relative and absolute quantification of drug effects on specific phosphorylation events can then be pursued using the peptide standards approach described.<sup>68,69</sup>

## Proteomics experimental data archiving

The increased complexity of proteome technologies and the subsequent development of high-throughput proteome analysis resulted over the last years in an overwhelming amount of experimental data. The diversity of the data is also enhanced by the dynamics of the proteome context in which the experiments were carried out. There is a clear need to develop internationally agreed standards for archiving proteomics experimental data. Although comprehensive databases of 2-DE gels are present on the Web, such as the Expert Protein Analysis System (ExPASy),

the available information is not sufficient for confident comparison of data sets. On the contrary, very few MS data sets are publicly available. The recently proposed Proteomics Experiment Data Repository (PEDRo)<sup>87</sup> model resembles the minimum set of information about the microarray experiment (MIAME)<sup>88</sup> model that has already been implemented for the archiving of transcriptome data. The PEDRo model takes into consideration all the various steps involved in a proteomics experiment, such as sample generation and processing, and parameters for the MS analysis and data interpretation, to capture the relevant information required for the reproduction, comparison, and archiving of any proteomics experiment.

In many respects the repertoire of experiments that can be performed in hematology has been expanded greatly by advances in MS and proteomics. This review gives an overview of some of these approaches and where they can be applied to hematology. It is likely that within 5 years their application will have transformed our knowledge of key process in normal and malignant hematopoiesis and blood cell function.

## Acknowledgments

The authors are grateful to Elaine Spooncer and Wenzhu Zhang for critical reading of this manuscript.

## References

- Han KK, Martinage A. Post-translational chemical modification(s) of proteins. *Int J Biochem*. 1992; 24:19-28.
- Voehringer DW, Hirschberg DL, Xiao J, et al. Gene microarray identification of redox and mitochondrial elements that control resistance or sensitivity to apoptosis. *Proc Natl Acad Sci U S A*. 2000;97:2680-2685.
- Blohm DH, Guiseppi-Elie A. New developments in microarray technology. *Curr Opin Biotechnol*. 2001;12:41-47.
- Futcher B, Latter GI, Monardo P, McLaughlin CS, Garrels JI. A sampling of the yeast proteome. *Mol Cell Biol*. 1999;19:7357-7368.
- Gygi SP, Rochon Y, Franz BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol*. 1999;19:1720-1730.
- Lian Z, Kluger Y, Greenbaum DS, et al. Genomic and proteomic analysis of the myeloid differentiation program: global analysis of gene expression during induced differentiation in the MPRO cell line. *Blood*. 2002;100:3209-3220.
- Dreger M. Proteome analysis at the level of sub-cellular structures. *Eur J Biochem*. 2003;270:589-599.
- Huber LA, Pfaller K, Vietor I. Organelle proteomics: implications for subcellular fractionation in proteomics. *Circ Res*. 2003;92:962-968.
- Gorg A, Obermaier C, Boguth G, et al. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis*. 2000; 21:1037-1053.
- Patton WF. Detection technologies in proteome analysis. *J Chromatogr B Anal Technol Biomed Life Sci*. 2002;771:3-31.
- Berggren K, Steinberg TH, Lauber WM, et al. A luminescent ruthenium complex for ultrasensitive detection of proteins immobilized on membrane supports. *Anal Biochem*. 1999;276:129-143.
- Tonge R, Shaw J, Middleton B, et al. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics*. 2001;1:377-396.
- Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol*. 1999;17:994-999.
- Cochran DAE, Evans CA, Blinco D, et al. Proteomic analysis of chronic lymphocytic leukemia subtypes with mutated or unmutated Ig VH genes. *Mol Cell Proteomics*. 2004;2:1331-1341.
- Hoving S, Voshol H, van Oostrum J. Towards high performance two-dimensional gel electrophoresis using ultrazoom gels. *Electrophoresis*. 2000;21:2617-2621.
- Adkins JN, Varnum SM, Auberry KJ, et al. Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Mol Cell Proteomics*. 2002;1:947-955.
- Stulik J, Tichy M, Kovarova H. Two-dimensional gel electrophoresis of four serum samples from patients with IgD myeloma. *Clin Chim Acta*. 1993; 218:149-158.
- Skibeli V, Nissen-Lie G, Torjesen P. Sugar profiling proves that human serum erythropoietin differs from recombinant human erythropoietin. *Blood*. 2001;98:3626-3634.
- Csar XF, Wilson NJ, Strike P, et al. Copper/zinc superoxide dismutase is phosphorylated and modulated specifically by granulocyte-colony stimulating factor in myeloid cells. *Proteomics*. 2001;1:435-443.
- Lewis TS, Hunt JB, Aveline LD, et al. Identification of novel MAP kinase pathway signaling targets by functional proteomics and mass spectrometry. *Mol Cell*. 2000;6:1343-1354.
- Martin K, Steinberg TH, Cooley LA, Gee KR, Beecher JM, Patton WF. Quantitative analysis of protein phosphorylation status and protein kinase activity on microarrays using a novel fluorescent phosphorylation sensor dye. *Proteomics*. 2003;3: 1244-1255.
- Blagoev B, Kratchmarova I, Ong SE, Nielsen M, Foster LJ, Mann M. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat Biotechnol*. 2003;21: 315-318.
- van Anken E, Romijn EP, Maggioni C, et al. Sequential waves of functionally related proteins are expressed when B cells prepare for antibody secretion. *Immunity*. 2003;18:243-253.
- Hanash SM, Madoz-Gurpide J, Misek DE. Identification of novel targets for cancer therapy using expression proteomics. *Leukemia*. 2002;16:478-485.
- Antonucci F, Chilosi M, Santacatterina M, Herbert B, Righetti PG. Proteomics and immunomapping of reactive lymph-node and lymphoma. *Electrophoresis*. 2002;23:356-362.
- Wang HX, Ma XC, Deng QL, Li D. Cytotoxicity of flutamide and 2-hydroxyflutamide and their effects on CYP1A2 mRNA in primary rat hepatocytes. *Acta Pharmacol Sin*. 2002;23:562-566.
- Levi-Schaffer F, Temkin V, Simon HU, Kettman JR, Frey JR, Lefkowitz I. Proteomic analysis of human eosinophil activation mediated by mast cells, granulocyte macrophage colony stimulating factor and tumor necrosis factor alpha. *Proteomics*. 2002;2:1616-1626.
- Bell AW, Ward MA, Blackstock WP, et al. Proteomics characterization of abundant Golgi membrane proteins. *J Biol Chem*. 2001;276:5152-5165.
- Andersen JS, Lyon CE, Fox AH, et al. Directed proteomic analysis of the human nucleolus. *Curr Biol*. 2002;12:1-11.
- Garin J, Diez R, Kieffer S, et al. The phagosome proteome: insight into phagosome functions. *J Cell Biol*. 2001;152:165-180.
- Desjardins M. ER-mediated phagocytosis: a new membrane for new functions. *Nat Rev Immunol*. 2003;3:280-291.
- Link AJ, Eng J, Schieltz DM, et al. Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol*. 1999;17:676-682.
- Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol*. 2001;19:242-247.
- Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003;422:198-207.
- Rees DC, Rochette J, Schofield C, et al. A novel silent posttranslational mechanism converts methionine to aspartate in hemoglobin Bristol (beta 67[E11] Val-Met → Asp). *Blood*. 1996;88:341-348.
- Troxler H, Neuheiser F, Kleinert P, et al. Detection of a novel variant human hemoglobin by electrospray ionization mass spectrometry. *Biochem Biophys Res Commun*. 2002;292:1044-1047.
- Clark RE, Dodi IA, Hill SC, et al. Direct evidence that leukemic cells present HLA-associated immunogenic peptides derived from the BCR-ABL b3a2 fusion protein. *Blood*. 2001;98:2887-2893.
- Lasonder E, Ishihama Y, Andersen JS, et al. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature*. 2002; 419:537-542.
- Demirev PA, Feldman AB, Kongkasuriyachai D, Scholl P, Sullivan D Jr, Kumar N. Detection of malaria parasites in blood by laser desorption mass spectrometry. *Anal Chem*. 2002;74:3262-3266.
- Smith DL, Evans CA, Pierce A, Gaskell SJ, Whetton AD. Changes in the proteome associated with the action of bcr-abl tyrosine kinase are not related to transcriptional regulation. *Mol Cell Proteomics*. 2002;1:876-884.
- Issaq HJ, Veenstra TD, Conrads TP, Felschow D. The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. *Biochem Biophys Res Commun*. 2002;292:587-592.
- Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*. 2002;359:572-577.
- Petricoin EF 3rd, Ornstein DK, Pawletz CP, et al. Serum proteomic patterns for detection of prostate cancer. *J Natl Cancer Inst*. 2002;94:1576-1578.
- Zhukov TA, Johanson RA, Cantor AB, Clark RA, Tockman MS. Discovery of distinct protein profiles specific for lung tumors and premalignant lung lesions by SELDI mass spectrometry. *Lung Cancer*. 2003;40:267-279.
- Dayal B, Ertel NH. ProteinChip technology: a new and facile method for the identification and measurement of high-density lipoproteins apoA-I and apoA-II and their glycosylated products in patients with diabetes and cardiovascular disease. *J Proteome Res*. 2002;1:375-380.
- Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem*. 1997;69:4751-4760.
- Yanagisawa K, Shyr Y, Xu BJ, et al. Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet*. 2003;362:433-439.
- Hunt DF, Yates JR 3rd, Shabanowitz J, Winston S, Hauer CR. Protein sequencing by tandem mass spectrometry. *Proc Natl Acad Sci U S A*. 1986;83:6233-6237.
- Fields CG, VanDrise VL, Fields GB. Edman degradation sequence analysis of resin-bound peptides synthesized by 9-fluorenylmethoxycarbonyl chemistry. *Pept Res*. 1993;6:39-47.
- Qin J, Chait BT. Matrix-assisted laser desorption ion trap mass spectrometry: efficient isolation and effective fragmentation of peptide ions. *Anal Chem*. 1996;68:2108-2112.
- Medzihradsky KF, Campbell JM, Baldwin MA, et al. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal Chem*. 2000;72:552-558.
- Loboda AV, Krutichinsky AN, Bromirski M, Ens W, Standing KG. A tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: design and performance. *Rapid Commun Mass Spectrom*. 2000; 14:1047-1057.
- Mann M, Hojrup P, Roepstorff P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol Mass Spectrom*. 1993;22:338-345.
- Yates JR 3rd, Speicher S, Griffin PR, Hunkapiller T. Peptide mass maps: a highly informative approach to protein identification. *Anal Biochem*. 1993;214:397-408.
- Eng JK, McCormack AL, Yates JR 3rd. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom*. 1994;5:976-989.
- Fenyo D, Qin J, Chait BT. Protein identification using mass spectrometric information. *Electrophoresis*. 1998;19:998-1005.
- van der Rest G, He F, Emmett MR, Marshall AG, Gaskell SJ. Gas-phase cleavage of PTC-derivatized electrosprayed tryptic peptides in an FT-ICR trapped-ion cell: mass-based protein identification without liquid chromatographic separation. *J Am Soc Mass Spectrom*. 2001;12:288-295.

58. Brancia FL, Oliver SG, Gaskell SJ. Improved matrix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-containing peptides. *Rapid Commun Mass Spectrom*. 2000;14:2070-2073.
59. Steen H, Pandey A. Proteomics goes quantitative: measuring protein abundance. *Trends Biotechnol*. 2002;20:361-364.
60. Han DK, Eng J, Zhou H, Aebersold R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat Biotechnol*. 2001;19:946-951.
61. Marshall AG, Senko MW, Li W, et al. Protein molecular mass to 1 Da by  $^{13}\text{C}$   $^{15}\text{N}$  double depletion and FT-ICR mass spectrometry. *J Am Chem Soc*. 1997;119:433-434.
62. Pratt JM, Robertson DH, Gaskell SJ, et al. Stable isotope labelling in vivo as an aid to protein identification in peptide mass fingerprinting. *Proteomics*. 2002;2:157-163.
63. Cristea IM. Proteome changes during early stages of cell injury [doctoral dissertation]. Manchester, United Kingdom: University of Manchester Institute of Science and Technology. 2002.
64. Bonenfant D, Schmelzle T, Jacinto E, et al. Quantitation of changes in protein phosphorylation: a simple method based on stable isotope labeling and mass spectrometry. *Proc Natl Acad Sci U S A*. 2003;100:880-885.
65. Griffin TJ, Gygi SP, Rist B, et al. Quantitative proteomic analysis using a MALDI quadrupole time-of-flight mass spectrometer. *Anal Chem*. 2001;73:978-986.
66. Patton WF, Schulenberg B, Steinberg TH. Two-dimensional gel electrophoresis: better than a poke in the ICAT? *Curr Opin Biotechnol*. 2002;13:321-328.
67. Shiio Y, Donohoe S, Yi EC, Goodlett DR, Aebersold R, Eisenman RN. Quantitative proteomic analysis of Myc oncoprotein function. *EMBO J*. 2002;21:5088-5096.
68. Barnidge DR, Dratz EA, Martin T, Bonilla LE, Moran LB, Lindall A. Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. *Anal Chem*. 2003;75:445-451.
69. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A*. 2003;100:6940-6945.
70. Lamond AI, Mann M. Cell biology and the genome projects—a concerted strategy for characterizing multiprotein complexes by using mass spectrometry. *Trends Cell Biol*. 1998;7:135-174.
71. Zhu H, Bilgin M, Bangham R, et al. Global analysis of protein activities using proteome chips. *Science*. 2001;293:2101-2105.
72. Pawlak M, Schick E, Bopp MA, Schneider MJ, Oroszlan P, Ehrt M. Zeptosens' protein microarrays: a novel high performance microarray platform for low abundance protein analysis. *Proteomics*. 2002;2:383-393.
73. Mitchell P. A perspective on protein microarrays. *Nat Biotechnol*. 2002;20:225-229.
74. Taussig MJ. Protein arrays: issues to be addressed: The chairperson's introduction to the session peptide and protein chips from the ESF workshop Proteomics: focus on protein interactions. *Comp Functional Genomics*. 2001;2:298-300.
75. MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. *Science*. 2000;289:1760-1763.
76. Belov L, de la Vega O, dos Remedios CG, Mulligan SP, Christopherson RI. Immunophenotyping of leukemias using a cluster of differentiation antibody microarray. *Cancer Res*. 2001;61:4483-4489.
77. Hebestreit HF. Proteomics: an holistic analysis of nature's proteins. *Curr Opin Pharmacol*. 2001;1:513-520.
78. Adam GC, Sorensen EJ, Cravatt BF. Chemical strategies for functional proteomics. *Mol Cell Proteomics*. 2002;1:781-790.
79. Brody EN, Willis MC, Smith JD, Jayasena S, Zichi D, Gold L. The use of aptamers in large arrays for molecular diagnostics. *Mol Diagn*. 1999;4:381-388.
80. Tasset DM, Kubik MF, Steiner W. Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J Mol Biol*. 1997;272:688-698.
81. Krishna RG, Wold F. Post-translational modification of proteins. *Adv Enzymol Relat Areas Mol Biol*. 1993;67:265-298.
82. Qin J, Chait BT. Identification and characterization of posttranslational modifications of proteins by MALDI ion trap mass spectrometry. *Anal Chem*. 1997;69:4002-4009.
83. Zhou H, Watts JD, Aebersold R. A systematic approach to the analysis of protein phosphorylation. *Nat Biotechnol*. 2001;19:375-378.
84. McLachlin DT, Chait BT. Analysis of phosphorylated proteins and peptides by mass spectrometry. *Curr Opin Chem Biol*. 2001;5:591-602.
85. Steen H, Fernandez M, Ghaffari S, Pandey A, Mann M. Phosphotyrosine mapping in Bcr/Abl oncoprotein using phosphotyrosine-specific ammonium ion scanning. *Mol Cell Proteomics*. 2003;2:138-145.
86. Salomon AR, Ficarro SB, Brill LM, et al. Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. *Proc Natl Acad Sci U S A*. 2003;100:443-448.
87. Taylor CF, Paton NW, Garwood KL, et al. A systematic approach to modeling, capturing, and disseminating proteomics experimental data. *Nat Biotechnol*. 2003;21:247-254.
88. Brazma A, Hingamp P, Quackenbush J, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet*. 2001;29:365-371.
89. Walker J, Flower D, Rigley K. Microarrays in hematology. *Curr Opin Hematol*. 2002;9:23-29.
90. Steiner S, Witzmann FA. Proteomics: applications and opportunities in preclinical drug development. *Electrophoresis*. 2000;21:2099-2104.
91. Shaw J, Rowlinson R, Nickson J, et al. Evaluation of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes. *Proteomics*. 2003;3:1181-1195.
92. Hillenkamp F, Karas M. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. *Methods Enzymol*. 1990;193:280-295.
93. Stensballe A, Jensen ON, Olsen JV, Haselmann KF, Zubarev RA. Electron capture dissociation of singly and multiply phosphorylated peptides. *Rapid Commun Mass Spectrom*. 2000;14:1793-1800.
94. Oda Y, Nagasu T, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat Biotechnol*. 2001;19:379-382.
95. Martinovic S, Veenstra TD, Anderson GA, Pasatolic L, Smith RD. Selective incorporation of isotopically labeled amino acids for identification of intact proteins on a proteome-wide level. *J Mass Spectrom*. 2002;37:99-107.
96. Zhang R, Sioma CS, Wang S, Regnier FE. Fractionation of isotopically labeled peptides in quantitative proteomics. *Anal Chem*. 2001;73:5142-5149.
97. Hansen KC, Schmitt-Ulms G, Chalkley RJ, Hirschi J, Baldwin MA, Burlingame AL. Mass spectrometric analysis of protein mixtures at low levels using cleavable  $^{13}\text{C}$ -isotope-coded affinity tag and multidimensional chromatography. *Mol Cell Proteomics*. 2003;2:299-314.
98. Mann M, Ong SE, Gronborg M, Steen H, Jensen ON, Pandey A. Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends Biotechnol*. 2002;20:261-268.
99. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem*. 1996;68:850-858.