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, compared 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)

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, proteinos, 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) 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 finalization of those objectives. Proteome analysis provides a view of biologic processes at their 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 Dreger7 and Huber et al8) 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...
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 coworkers9 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.10 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 35S or 32P, 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,11 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 α amines of peptides and the ε 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 DIGE12 (Figure 2B). We have used this approach with chronic myeloid leukemia CD34+ cells using 1 × 10⁶ cells/sample, visualizing approximately 1000 protein spots per 2-DE gel. This cell population (> 95% CD34+) 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 differential 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.13 A new cyanide dye derivative that labels cysteine residues has now been produced and offers enhanced sensitivity12 (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.9 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.14 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.15 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.16

Further application of gel-based technologies in hematology can be seen in a variety of different studies.17,18 Prelabeling cells with 32P has been used to recognize proteins phosphorylated as a consequence of cytokine action.19 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.20 Recently, a fluorescent dye that reversibly associates with phosphorylated proteins within gels has been characterized.21 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
epidermal growth factor (EGF) after culture in [13 C]-arginine

reticulum components sequentially, prior to engaging on the increase expression of metabolic enzymes and endoplasmic MALDI peptide proteins, tandem mass spectrometry (MS/MS) analysis is performed, generally using electrospray (ES) ionization. A prevalent method that separates proteins at the peptide hydrophobicity of the proteins). Recently, a new set of dyes intended to greater sensitivity. (C) Following separation, the protein spot is usually excised, subjected to in-gel enzymatic proteolysis, and analyzed by MS, usually MALDI-MS. The antibody production provides an elegant example of the applica-change in the proteome of lymphocytes as they prepare for...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. Untreated HeLa cells, grown in medium containing [13 C]-arginine, were mixed with cells treated with epidural growth factor (EGF) after culture in [13 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. 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. 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.

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. 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 organelar function. With respect to hematopoietic cells, phagosomes have been isolated from macrophages and over 140 proteins recognized in a 2-DE/MS study. 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. 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.
A nonquantitative approach to identification of such proteins within biologic material is exemplified by the work of Link et al\textsuperscript{32} who submitted a protein complex directly to enzymatic digestion within biologic material is exemplified by the work of Link et al\textsuperscript{32} 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\textsuperscript{33} to achieve full automation. The technique, termed multidimensional protein identification technology (MudPit), 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.\textsuperscript{34} In hematology, MS has been used for protein, peptide, and DNA sequencing,\textsuperscript{4} protein-folding studies, noncovalent interaction analysis, analysis of PTMs,\textsuperscript{35} identification of novel proteins,\textsuperscript{36} and in a variety of diagnostic and drug discovery projects.\textsuperscript{36,37,40}
The main components of a mass spectrometer are an ion source, one or several mass analyzers, and a detector. 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. 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. A similar study on prostate cancer and lung cancer reveals the general applicability of this technique. 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 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. Other configurations, however, have also been successfully applied for the analysis of peptides and proteins, such as MALDI-ion trap, MALDI-QqToF (where q represents a quadrupole ion decomposition region), and MALDI-ToF-ToF. 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 μL/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 reflection, 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.
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. More recent work, such as the study of non–small-cell lung tumors, 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 is a favorable alternative to the classical sequencing performed by the Edman degradation technique.

In contrast to Edman degradation, which remains valuable for peptide sequencing and has therefore evolved as an indispensable tool for protein recognition, structure elucidation, and characterization of PTMs.

Figure 5. Sequencing of peptides using MS/MS. Peptide fragmentation by collisionally activated dissociation (CAD) during MS/MS 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 recent advances in coupling a MALDI ion source to ion trap, ToF/ToF, and Q-ToF 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. Recently developed algorithms allow the search of uninterpreted or partially interpreted MS/MS spectra for example, SEQUEST, 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 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 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 ion 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 diagnostic purposes (“SELDI-ToF MS and the search for disease markers”).

Figure 6. Peptide fragmentation by collisional activation during MS/MS analysis. Panel A: Typical product ion spectrum of the doubly charged precursor ion selected for fragmentation in MS/MS analysis. Panel B: Product ion spectrum of the doubly charged precursor ion selected for fragmentation in MS/MS analysis. 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]+ denotes the doubly charged precursor ion selected for fragmentation, and the [M+2H]+·H3PO4 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.59 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.59,60 (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,61,62 or at the protein or peptide level.

In our laboratory, we have used guanidination with [15 N]-and [14 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.53 The 2 Da mass difference between the light and heavy isotope-labeled peptides allows for simultaneous selection and fragmentation of both isotopeomers 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[D]O or H[D]O, followed by immobilized metal-affinity chromatography, dephosphorylation, and mass spectrometric analysis has been successfully applied for the relative quantification of protein phosphorylation.64 Proteins from 2 cell states can be labeled and separated using an isotope-coded affinity tag (ICAT) approach.13 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.65 ICAT technology has been applied for the separation and quantification of proteins with extreme pI values, low abundance,33 and membrane-bound proteins.60

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.66 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.60 A paradigm study on the leukemogenic oncogene product Myc by Shiio et al67 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.68,69 The synthetic peptide, labeled with a stable isotope, is mixed with the sample

![Figure 6. Isotopic labeling for relative quantification using MS](image-url)蛋白质或肽段来自2个不同细胞状态被标识化并带有轻重同位素的相同化学试剂。实验样品被结合并用MS分析，使相对量的计算成为可能。这种方法是基于的假设是，结构上相似的同位素标记肽有相同的电离效果。同位素标记可以被整合在细胞培养的任何时期，或者在蛋白质或肽段水平。

在我们的实验室，我们使用了15N-和14N-O甲基异硫脲的组合增加了检测 lysine C-terminus peptides的敏感性，相对量化2个细胞状态，和自发 sequencing。53 2 Da质荷比差适用于轻重同位素标记的肽，允许同时选择和碎片化这两种同位素，因此可以轻易识别同一蛋白的轻重双峰在产物离子谱中。另一种选择是通过MS/MS的直接对比Chemically similar entities可以获取相对量的。等同位素标记物可以在细胞培养阶段，61,62或在蛋白质或肽段水平整合。

在我们的实验室，我们使用了15N-和14N-O甲基异硫脲的组合增加了检测 lysine C-terminus peptides的敏感性，相对量化2个细胞状态，和自发 sequencing。53 2 Da质荷比差适用于轻重同位素标记的肽，允许同时选择和碎片化这两种同位素，因此可以轻易识别同一蛋白的轻重双峰在产物离子谱中。另一种选择是通过MS/MS的直接对比Chemically similar entities可以获取相对量的。等同位素标记物可以在细胞培养阶段，61,62或在蛋白质或肽段水平整合。

虽然ICAT释放了对2-DE技术的约束，最近的研究比较了ICAT和2-DE，显示两者均不提供全面的蛋白质组覆盖。66 对这些2种技术的全面评估是需要的，我们目前正在进行这样的研究。ICAT技术在491个微囊状蛋白上的应用和表达变化的检测在HL-60的白血病样细胞系中展示了这种技术的力量，特别是因为被识别的蛋白质是疏水性的。然而，这样的研究并不简单。60 一项关于白血病 oncogene Myc的标志物Shiio et al67展示了单一蛋白质对途径的广泛影响，如细胞的迁移控制和蛋白质合成。

最近，一些使用合成肽的方法发展出来了，这些合成肽设计模仿自然肽，尤其是遵循蛋白质的酶切和消化。68,69 合成肽，用稳定同位素标记，被与样本混合。
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.28,30,70 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 paper71 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,72 as reviewed by Mitchell.73 Advances in surface chemistry, microfluidics, and detection techniques will make rapid analysis of proteins in samples a reality in the future.74,75

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.76 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,77 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, Pro cognia ( 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.78 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,79 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.80 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,81 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.82 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.83 The variety of methodologies used for the analysis of protein phosphorylation has been thoroughly reviewed by McLachlin and Chait84 (and references within). Affinity separation followed by MS/MS analysis for the detection of phosphotyrosine-specific immunonium ions has been recently successfully used for the detection of 9 phosphorylation...
sites of Bcr/Abl fusion oncoprotein, of which 6 were novel. Another elegant example is the sequential use of phosphotyrosine immunoprecipitation, methyl esteriﬁcation, IMAC (Fig-ure 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. 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. Another elegant example is the sequential use of phosphotyrosine immunoprecipitation, methyl esteriﬁcation, IMAC (Fig-ure 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. 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. Another elegant example is the sequential use of phosphotyrosine immunoprecipitation, methyl esteriﬁcation, IMAC (Fig-ure 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. 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.

**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) model resembles the minimum set of information about the microarray experiment (MIAME) 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


drial proteins that control resistance or sen- 
sitivity to apoptosis. Proc Natl Acad Sci U S A.  
2000;97:2680-2685.


5. Lian Z, Kluger Y, Greenbaum DS, et al. Genomic and proteomic analysis of the myelodysplasia differen-
tiation program: global analysis of gene expression during induced differentiation in the MPRO cell 

6. Drescher M. Proteome analysis at the level of subcellular structures. Eur J Biochem. 2003;270:589- 
599.

7. Huber LA, Pfaller K, Vietor I. Organelle proteo-
tics: implications for subcellular fractionation in 

rent state of two-dimensional electrophoresis with 
immobilized pH gradients. Electrophoresis. 2000; 
21:1037-1053.

9. Patton WF. Detection technologies in proteome 

Luminescent rheumatoid complex for ultrasensitive 
detection of proteins immobilized on membrane 

and development of fluorescence two-dimen-
sional differential gel electrophoresis proteomics 

12. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, 

13. Cochran DAE, Evans CA, Blinco D, et al. Pro-
teomic analysis of chronic lymphocytic leukemia 
subtypes with mutated or unmutated Ig VH genes. Mol Cell Proteomics. 2002;4:1313-1341.

high performance two-dimensional gel electro-
phoresis using ultrazoom gels. Electrophoresis. 

15. Wang HK, Ma XG, Deng GL, D. Cytotoxicity of 
flutamide and 2-hydroxyflutamide and their ef-
cets on FVP1A2 mRNA in primary rat hepatoc-

16. Levi-Schaffer F, Temvik V, Simon HU, Kettman 
JR, Frey JR, Leffkows I. Proteomic analysis of 
human eosinophil activation mediated by mast 
cells, granulocyte macrophage colony stimulating 
factor and tumor necrosis factor alpha. Proteome-

17. Bell AW, Ward MA, Blackstock WP, et al. Pro-
teomics characterization of abundant Golgi mem-
brane proteins. J Biol Chem. 2001;276:5152- 
5155.

proteomic analysis of the human nucleolus. Curr 

proteome: insight into phagosome functions. 

20. Desjardins M. ER-mediated phagocytosis: a 

of protein complexes using mass spectrometry. 

22. Washburn MP, Wolters D, Yates JR 3rd. Large-
scale proteomic analysis of the myeloid differentia-
tion in yeast. Mol Cell Biol. 1999;19:1720-
1730.

23. Antonucci F, Chilosi M, Santacaterina M, Herbert 
C, Righetti PG. Proteomics and immunomapping 
of reactive lymph-node and lymphoma. Electro-

2003;3:149-158.

that leukemic cells present HLA-associated im-
modification(s) of proteins. Int J Biochem. 1992; 
24:149-158.

26. Skibeli V, Nissen-Lie G, Torjesen P. Sugar pro-
tection against red cell agglutination mediated by 
the action of bcr-abl tyrosine kinase are not re-
lated to transcriptional regulation. Mol Cell Pro-

27. Issaq HJ, Veenstra TD, Conrads TP, Felschow D. 
The SELDI-TOF MS approach to proteomics: 
protein profiling and biomarker identification. Bio-

28. Petricoin EF, Ardekani AM, Hilt BA, et al. Use of 
proteomic patterns in serum to identify ovarian 

 serum proteomic patterns for detection of pro-
tate cancer. J Natl Cancer Inst. 2002;94:1576- 
1578.

30. Dayal B, Ertel NH. ProteinChip technology: a new 
and facile method for the identification and mea-
surement of high-density lipoproteins apoA-I and 
apoA-II and their glycosylated products in pa-
tients with diabetes and cardiovascular disease. J 

31. Caprioli RM, Farmer TB, Gile J. Molecular imag-
ing of biological samples: detection of peptides 
and proteins using MALDI-TOF MS. Anal Chem. 
1997;69:4751-4760.

patterns of tumour subsets in non-small-cell lung 

33. Flohr DF, Yates JR 3rd, Shabanowitz J, Winston 
S, Hauer CR. Protein sequencing by tandem mass 
 spectrometry. Proc Natl Acad Sci U S A. 

34. Fields CG, VanDrisse V, Fields GB. Edman degr-
adation sequence analysis of resin-bound pep-
tides synthesized by 9-fluorenlymethoxy carbonyl 

35. Qin J, Chai BT. Matrix-assisted laser desorption ion 
trap mass spectrometry: efficient isolation and ef-
fective fragmentation of peptide ions. Anal Chem. 

36. Medzhibuzldsky KF, Campbell JM, Baldwin MA, et al. The characteristics of malignant differen-
ced dissociation using a high-performance 
MALDI-TOF/TQ tandem mass spectrometer. Anal 

37. Loboda AV, Krutchinsky AN, Bromirski M, Ens W, 
Standing KG. A tandem quadrupole mass detector 
with a matrix-assisted laser desorption/ionization 
source: design and perfor-
14:1047-1057.

38. Mann M, Hnjup P, Roepstorff P. Use of mass 
spectrometric molecular weight information to iden-

39. Yates JR 3rd, Speicher S, Griffin PR, Hunkapiller 
M. Tandem masses: a new approach to protein 

40. Van den Rest G, He F, Emmett MR, Marshall AG, 
Gaskell SJ. Gas-phase cleavage of PTC-derivatized 
electrospray-ionized tryptic peptides in an FT-ICR 
trapped-ion cell: mass-balance identification 
without liquid chromatographic separation. J Am 
70. Lamond AI, Oliver SG, Gaskell SJ. Improved ma-
trix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-contain-

sis of protein activities using proteome chips. Sci-
ence. 2001;293:2101-2105.

72. Pawlik M, Schick E, Bopp MA, Schneider MJ, Oroszlan P, Ehrat M, Zepolons’ protein microar-
ays: a novel high performance microarray plat-
form for low abundance protein analysis. Pro-


74. Tausig MJ. Protein arrays: issues to be ad-
dressed. The chairperson’s introduction to the session peptide and protein chips from the ESF
workshop Proteomics: focus on protein interac-
tions. Comp Functional Genomics. 2002;1:298-
304.

75. MacBeath G, Schreib L. Printing proteins as microarrays for high-throughput function determi-

76. Betov L, de la Vega O, dos Remedios CG, Mulli-
gan SP, Christopherson RI. Immunophenotyping of leukemias using a cluster of differentiation anti-
body microarray. Cancer Res. 2001;61:4483-
4489.

513-520.

78. Adam GC, Sorensen EJ, Cravatt BF. Chemical
strategies for functional proteomics. Mol Cell Pro-
teomics. 2002;1:761-790.

388.

80. Tasset DM, Kubik MF, Steiner W. Oligonucleotide
microarrays: a novel high performance microarray plat-
form for low abundance protein analysis. Pro-

81. Krishna RG, Wold F. Post-translational modifica-
tion of proteins. Adv Enzymol Relat Areas Mol

82. Qin J, Chait BT. Identification and characteriza-
tion of posttranslational modifications of proteins by MALDI ion trap mass spectrometry. Anal

83. Zhou H, Watts JD, Aebbersold R. A systematic ap-
proach to the analysis of protein phosphorylation.

84. McLachlin DT, Chait BT. Analysis of phosphory-
lated proteins and peptides by mass spectrom-

85. Steen H, Fernandez M, Ghaffari S, Pandey A, Mann M. Phosphoryrosine mapping in Bcr/Abl
oncoprotein using phosphotyrosine-specific im-
munion scanning. Mol Cell Proteomics. 2003;
2:138-145.

of tyrosine phosphorylation pathways in human
cells using mass spectrometry. Proc Natl Acad

87. Taylor CF, Paton NW, Garwood KL, et al. A sys-
tematic approach to modeling, capturing, and dis-
seminating proteomics experimental data. Nat

88. Brazina A, Hingamp P, Quackenbush J, et al. Min-
imum information about a microarray experi-
ment (MIAME)-toward standards for microarray

89. Walker J, Flower D, Rigley K. Microarrays in he-

90. Steiner S, Wittmann FA. Proteomics: applica-
tions and opportunities in preclinical drug devel-

of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes. Proteomics.

92. Hillenkamp F, Karas M. Mass spectrometry of peptides and proteins by matrix-assisted ultra-

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, Pasa-
Tovic L, Smith RD. Selective incorporation of iso-
topically labeled amino acids for identification of intact proteins on a proteome-wide level. J Mass

tionation of isotopically labeled peptides in quanti-
tative proteomics. Anal Chem. 2001;73:5124-
5149.

J, Baldwin MA, Burlingame AL. Mass spectromet-
ric analysis of protein mixtures at low levels using cleavable 13C-isotope-coded affinity tag and mul-
tidimensional chromatography. Mol Cell Proteom-

98. Mann M, Ong SE, Gronborg M, Steen H, Jensen
ON, Pandey A. Analysis of protein phosphoryla-
tion using mass spectrometry: deciphering the phospophoproteome. Trends Biotechnol. 2002;20:
261-268.

yrosine mass spectrometry: quantification of phos-
phorylated proteins in a proteome-wide study. Mol
Proteomics techniques and their application to hematology

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