Review

Proteomic approaches in the search for disease biomarkers

A. Vlahou*, M. Fountoulakis

Laboratotory of Biotechnology, Foundation for Biomedical Research of the Academy of Athens, 11527 Athens, Greece

Abstract

Significant technological advances in protein chemistry, physics and computer sciences in the last two decades have greatly improved protein separation methodologies, such as electrophoresis and chromatography, and have established mass spectrometry (MS) as an indispensable tool for protein study. The goal of this review is to provide a brief overview of the recent improvements in these methodologies and present examples from their application in proteome analysis and search for disease biomarkers.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Mass spectrometry; Proteomics; Electrophoresis; Liquid chromatography; Biomarkers

Contents

1. Introduction

Proteomics involves the study of the protein content of an organism including the identification of protein amino acid sequence, modifications, structure and ultimately assignment to functional pathways in the cell. Systematic proteomic studies are imperative since proteins perform the main cellular functions that lead to cell growth, differentiation, proliferation and death. Aberrations in protein structure and expression levels signal in most cases the presence of disease, rendering the proteins excellent targets in disease diagnostics, prognostics and therapeutics.

The complexity and variation in protein structure, the plethora of post-translational modifications as well as the presence of most regulatory proteins at very low levels in the cell, are the main challenges that proteomic research has to encounter. Undeniably, advances in the fields of chemistry, physics and computer sciences in the last decades have greatly

* Corresponding author. Tel.: +30 210 659 7506; fax: +30 210 659 7545.
  E-mail address: vlahoua@bioacademy.gr (A. Vlahou).

facilitated proteomic studies. Significant improvements in the protein separation methodologies such as two-dimensional electrophoresis (2DE) and chromatography, as well as in mass spectrometry systems, have increased the resolution, sensitivity and accuracy in protein detection. Nevertheless, some of the main challenges, such as the detection of the low abundance molecules, remain. In the same time, some new challenges have arisen, such as the development of the computing resources that could reliably analyze the vast amount of data that the application of the new technologies generate. The scope of this review is to outline some of the major technical developments in the fields of protein separation and identification, with an emphasis on mass spectrometry-based approaches, that occurred in the last decades and provide some examples of the applications of these new methodologies in the identification of disease biomarkers.

2. Protein separation by two-dimensional electrophoresis

One of the first steps in protein analysis is protein separation. Two-dimensional gel electrophoresis is the classical and principal tool for the protein profiling of biological samples [1]. This methodology provides the highest resolution in protein separation, capable of resolving thousands of proteins in one experiment. A brief overview of the principles of 2DE and advances of this methodology that have allowed its combination with mass spectrometry during the last few years, are provided below.

In 2DE, proteins are separated on the basis of differences in their net charge, through a technique known as isoelectric focusing (IEF) and differences in their molecular masses through polyacrylamide gel electrophoresis. IEF is an equilibrium process, during which, under the influence of a high electric field, proteins move along a stable pH gradient and focus at a position where they have no net charge. There are two ways to form a pH gradient: (i) with the use of carrier ampholytes; and (ii) with the use of an immobilized pH gradient (IPG) [2]. In the first case, upon application of the electric field the carrier ampholytes move and align themselves between the electrodes. In the second system, the IPG is formed prior to the IEF run by acrylamide derivatives, called immobilines. These are weak acids and bases with a buffering capacity, co-polymerized with acrylamide and fixed on plastic strips.

The IPG strips is the preferred method of IEF currently being utilized and it is safe to say that their application revolutionized 2DE. Among the benefits their use brought up, is a significant improvement in the reproducibility of the 2DE methodology. Furthermore, IPG strips are ideal for subsequent protein identification techniques as they allow application of relatively large amounts of protein [3,4].

In the second dimension of 2DE, the proteins are separated according to their size on porous polyacrylamide gels. Prior to this step, the anionic detergent sodium dodecyl sulfate (SDS) is added, so that the proteins are negatively charged. In comparison with IEF, the second dimension of separation on SDS-gels is relatively easy to control. Gels of a constant acrylamide concentration or gradient gels can be used. The latter such as the 9–16% linear gradient gels allow the visualization of proteins with a wide spectrum of molecular masses (5–200 kDa), whereas, gels of constant acrylamide concentration provide higher protein resolution at a specific mass [3].

Gels can be stained for protein detection by Coomassie blue, silver or fluorescent dyes. Images of the gels acquired by laser scanners or fluorescent imagers as applicable, are subsequently analyzed by specific software such as PDQuest, Melanie, etc. [5]. These software programs provide spot detection, and ratio analysis for the determination of quantitative protein changes between different samples. It should be noted however, that comparison of 2D gels for the identification of protein expression differences, as described above, involves the technical challenge of matching spots of different gels. This task is often considered extremely arduous due to the variability of the 2DE methodology as reflected by spot streaking, gels bending etc. To overcome this problem, the technique of differential in-gel electrophoresis (DIGE) may be applied which involves labeling of protein samples with fluorescent dyes, followed by their separation in the same gel. In this way, inter-gel variability is avoided. Further analysis includes visualization of the spots and signal quantification using fluorescent imagers [6–9].

The major advantage of 2DE is that it enables the simultaneous separation and visualization of thousands of unknown proteins. It should be noted, that on 2D gels, the proteins are often represented by more than one spot, so that the number of expressed products is much higher than the number of the corresponding encoding genes [10–12]. In most cases, we do not know either the origin or the biological message of the observed differences between the number of expressed products and the number of encoding genes.

Besides the inter-gel variability as mentioned above, there are some additional limitations of 2DE: (i) unless coupled to specific protein enrichment steps, 2DE analysis resolves only the major components of a protein mixture [12–18]; (ii) along the same lines, the detection of the low and high molecular mass, of basic and hydrophobic proteins is inefficient [18,19]; (iii) although automation of protein excision and development of image analysis software for the comparison of 2D gel protein maps have greatly facilitated the analysis of the results, 2D gel analysis remains labor intensive. Improvements of the technology are urgently required for an efficient proteome analysis [20].

3. Protein separation by liquid chromatography

Besides the 2DE, chromatographic strategies are also widely utilized to separate peptide and protein mixtures into their components. In this case, separation is based on differ-

ences in the peptide/protein physiochemical properties such as size (size exclusion chromatography), charge (ion exchange chromatography) and hydrophobicity (reverse phase chromatography) [7,21]. The experimental protocol involves dilution of the protein mixture with a buffer that is compatible with the type of column in use and then elution of the bound material based on some type of competitive attraction. For example, in the cation exchange chromatography elution may take place by the use of gradient of salt. Various types of detectors can be connected on line with the columns so as to provide qualitative and occasionally quantitative information on the separated molecules. Unless however, coupled to mass spectrometers, chromatography indicates the presence of a protein without providing its identity.

This method of separation is faster compared to 2DE and can be applied efficiently for the study of hydrophobic proteins and protein post-translational modifications [21–25]. In addition, multi-dimensional chromatographic methods (for example ion exchange and hydrophobic columns put in tandem) have been developed and successfully coupled to mass spectrometry systems. This methodology known as MUD-PIT (multi-dimensional protein identification technology) allows the rapid study and identification of proteins within complex protein mixtures [7,23–25]. When combined to metabolic, enzymatic or chemical protein-labeling methods, this methodology allows the detection of quantitative differences between proteins of different classes of samples [7,24,25] (more details on the isotope coded affinity tags (ICAT) system are provided in Section 7.1). Automated systems have been developed that perform all operations from sample application to fraction collection and on line connection to mass spectrometers for protein identification. However, the instrumentation that is needed for the automated MUD-PIT is expensive, and the computing resources for the identification of proteins from such complex mixtures are still under development. For that reasons, in many cases, 2DE remains the separation process of choice for the study of very complex protein mixtures such as biological fluids and tissue/cell extracts.

4. Mass spectrometry: advances in instrumentation

Mass spectrometry has been very effectively coupled to both 2DE and liquid chromatography (LC) methodologies so as to achieve protein identification. In brief, analysis of a protein/peptide molecule (analyte) by mass spectrometry occurs in three major steps: (i) ionization of the protein and generation of gas-phase ions; (ii) separation of ions according to their mass to charge ratio; and (iii) detection of the ions. These functions are performed by the three major components of a mass spectrometer: the ion source, the mass analyzer and the detector, respectively (reviewed in [26,27]).

Improvement of mass spectrometry in the last two decades, have resulted from changes in several areas of the mass spectrometer design. Development of the

"soft" desorption-ionization techniques namely, matrix-assisted laser desorption/ionization (MALDI) [28–30] and electrospray ionization (ESI) [31–34] in the late 1980s revolutionized proteomics research by rendering the study of macromolecules feasible.

A detailed description of the ionization processes followed in MALDI and ESI would be out of the realm of this brief review, but we would recommend the manuscripts by Knockenmuss and Zenobi [35], Karas et al. [36], Kebarle [37], Smith et al. [38] and Mann and Wilm [39], for the interested reader. To briefly mention, MALDI is based on the utilization of an energy absorbing molecule (matrix), usually a small organic molecule (like α-cyano-4-hydroxy-cinnamic acid). The matrix is highly absorbent at the laser wavelength as a means to facilitate the ionization of the analyte and also to protect it from laser-induced damage. The ionization and desorption processes that follow the laser irradiation of the analyte-matrix crystals result in the generation of their primarily singly charged ions [35,36].

In contrast to MALDI, during ESI, the analyte is ionized from a solution and transferred into the gas phase by spraying from a high voltage needle [37–39]. Usually multiple charging of the analyte occurs resulting in the generation of multiple consecutive ions differing by a single charge. NanoLC-ESI systems have recently been developed that enable sample introduction at very low flow rates (<25 nL/min). In this way, decreased sample quantity requirements and increased sensitivity are achieved [39–41].

Besides the development of the ESI and MALDI methodologies, improvement in mass spectrometry has also resulted from advancement in mass analyzers. The time-of-flight (TOF), ion trap, quadrupole and Fourier transform ion cyclotron (FTIC) are the four major types of mass analyzers currently in use. Each of them shows distinct advantages and weaknesses with regards to the sensitivity, mass accuracy, and mass resolving power they provide (reviewed in [27,42]). Notably, combination of different mass analyzers in tandem has recently become possible (for example, quadrupole–TOF, TOF–TOF, quadrupole–ion trap, etc.). In this way, combination of the strengths of each analyzer is achieved greatly improving the outcome of the proteomics research. Furthermore, although MALDI has been traditionally coupled to time-of-flight and ESI to quadrupole mass filter or ion trap analyzers, mass spectrometers have been recently developed that combine ESI with TOF as well as MALDI with quadrupole-ion traps. [27,43].

5. Peptide fingerprinting by mass spectrometry analysis

MALDI and ESI-based mass spectrometry technologies are being widely employed as protein identification tools. Most widely utilized is the "bottom-up" approach where proteins are digested with endoproteinases and their peptide maps (i.e. a list of peptide masses) subsequently obtained and

characterized by MS. These maps are then searched against theoretical maps of proteins included in databases (PROWL, EMBL, etc. reviewed in [27,44–46]).

In brief, in the case of 2D gels, for the generation of peptides, protein spots are excised from the gels and the proteins are in-gel digested with a protease, usually trypsin. The peptides generated during digestion can then be directly analyzed by mass spectrometry or first separated by LC. The experimental peptide masses are then correlated to the peptide fingerprints (i.e. list of peptides generated by the theoretical digest) of known proteins through search engines/algorithms such as Mascot, Sequest, MSFIT, etc. Matches are calculated on a probability basis and the degree of similarity of the experimental to the theoretical peptide mass is denoted by various mathematical parameters such as score and delta correlation, depending on the algorithm in use. Typically, a minimum of four peptide molecular weight matches is recommended to reduce the possibility of false-positive matches.

Notably, peptide mapping of spots detected in 2D gels has been adapted to high-throughput platforms. Coupling of MALDI-TOF-MS to 2D gels has been automated, enabling the analysis of thousands of samples per day with a minimal involvement of the operator [47–49]. With this approach, typically an identity is assigned to 60–80% of the spots analyzed.

In contrast to the "bottom-up" (i.e. from the peptides to identify the protein), an alternative "top-down" approach for protein identification has been recently developed which includes the direct MS analysis of intact proteins [50–52]. In this case, protein fragmentation takes place electronically during the MS analysis (for example using Fourier-transform mass spectrometers in conjunction with electron capture dissociation). This methodology, provides the advantage of studying the protein along with its post-translational modifications as a whole, however, its application is still limited, due to primarily the lack of optimized relevant experimental methodologies and, with the exception of the Prosight software developed by Taylor et al. [51], the lack of specific software for the analysis of MS data from intact proteins.

6. Peptide sequencing by tandem mass spectrometry

The drawback of peptide mapping is that protein identifications can be ambiguous. In this case, amino acid sequence analysis has to be performed. An alternative to the traditional Edman degradation procedure is the generation of peptide-sequence tags by tandem mass spectrometry ([53–56], reviewed in [27,34,44]). MS/MS or tandem mass spectrometry, is based on the detection and molecular mass determination of fragment ions that are formed either spontaneously (post-source decay or PSD) or by induced fragmentation (collision-induced dissociation or CID) from an intact ion. Simply stated, following their ionization and desorption, peptide ions may be fragmented to produce secondary ions which, can be separated in a second mass analyzer to give the mass per charge (m/z) of these fragment ions. Information obtained by

MS/MS can be utilized for the identification of proteins from protein databases, databases of expressed sequence tags or translated DNA databases. Besides allowing the more reliable identification of protein spots, the MS/MS approach enables the identification of post-translational modifications such as phosphorylation and glycosylation. With the current availability of identification algorithms that analyze the MS/MS results, there is a widespread application of this technology.

7. Application of mass spectrometry approaches in disease research

The application of mass spectrometry-based systems in disease research has been enormous. This methodology has been applied in the investigation of infectious diseases, diseases of the central nervous system, heart and cancer (reviewed in [49,56–65]). Many changes in protein levels resulting from specific disorders have been identified by proteomic approaches. This field is too broad to be covered in this review and our present work is confined to provide some highlights and specific examples drawn from the application of mass spectrometry-based systems in neuronal disease and cancer diagnostics.

7.1. Neurological diseases

Our group and others have applied proteomics in the investigation of protein changes in disorders of the central nervous system, like Alzheimer's disease (AD) and Down syndrome (DS; [64,66–73], reviewed in [49]). Samples from five regions of the brain of the control (i.e. with no history of neurological disease) and the disease groups were analyzed by 2D gels and the proteins with different levels in the disease groups were identified by MALDI-MS. In the disease groups, we found up to about 10-fold increased levels in the AD brain of the glial fibrillary acidic protein (GFAP), a known marker of neuronal decay and brain damage, which distinguishes astrocytes from other glial cells during development of the central nervous system [66]. We also observed increased levels for the 14-3-3γ and 14-3-3ε proteins in the Alzheimer's and Down syndrome groups by about 1.5-fold [67]. The 14-3-3 proteins exert complex functions in signal transduction pathways. Deranged levels may reflect impaired signaling and apoptosis in the brain. The levels of synaptosomal associated protein 25 kDa (snap-25) decreased in the two disease groups to approximately 40% [66]. Snap-25 is widely distributed in the brain, is an integral constituent of the synaptic core complex, participates in synaptic vesicles exocytosis and is involved in the formation of presynaptic sites. Decreased snap-25 levels may lead to deranged functions in exocytosis and neurotransmission. As an example, Fig. 1 shows the spots representing the 14-3-3 and snap-25 proteins with deranged levels in DS and AD. The reason for the change in the position of SNAP-25 in AD (Fig. 1C) is not currently known.
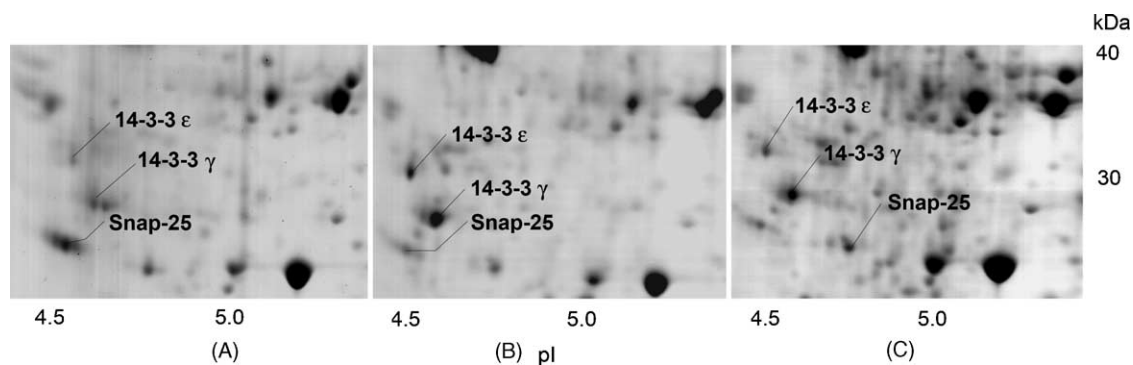
Fig. 1. Partial two-dimensional gel images of human brain proteins from a control (A), a patient with Down syndrome (B) and a patient with Alzheimer's disease (C). The proteins from the frontal lobe of cortex were separated on 3–10 non-linear IPG strips, followed by 9–16% SDS-polyacrylamide gels. The gels were stained with Coomassie blue. The proteins were identified by MALDI-MS. The spots representing proteins with different levels in the control and diseased groups, 14-3-3ε and γ and synaptosomal associated protein 25 (snap-25), are indicated.

Additional changes in the levels of several other brain proteins, such as glycerol 3-phosphate dehydrogenase, dihydropyrimidinase-related protein (DRP) 2, synaptotagmin, voltage-dependent anion channel proteins VDAC1 and VDAC2 were also observed [68–74]. The technology is also being utilized to study changes in the brain proteins of animals serving as models for neurological diseases, such as ischemia, anxiety and pain.

An alternative approach including ICAT labeling of proteins in combination with micro-capillary LC-MS analysis was used by Johnson et al. [75], for the study of cell death processes that occur in neurons following DNA damage. Proteins from primary cultures of cortical neurons cultured under control or conditions that induce DNA damage, were labeled with light and heavy (with deuterium) isotopic versions of the ICAT reagents. These reagents contain a biotin affinity tag and a thiol reactive group so as to interact with the cysteine residues. Following their labeling, proteins were further digested with trypsin and the resulting peptides loaded onto an avidin column for enrichment for the ICAT-labeled forms. The peptides were then separated by the use of micro-capillary reverse phase LC coupled on line with either ESI-ion trap MS or ESI-FTIC mass spectrometer. Using this approach, identification and relative quantification of 150 proteins was achieved which indicated alterations in the levels of various proteins following DNA damage. These included enzymes involved in glycolysis, anti-oxidative and stress proteins, proteins involved in neurite degeneration, as well as proteins linked to neurodegenerative diseases such as Tau, 14-3-3ζ, cofilin, etc.

7.2. Cancer

The two-dimensional electrophoresis systems in combination with mass spectrometry have also been extensively utilized in the quest for cancer biomarkers. In the case of bladder cancer, extensive 2DE of bladder tumor cells and urine samples has led to a comprehensive 2D database for bladder cancer which includes profiles of both transitional cell carcinomas (TCC) and squamous cell carcinomas (SCC) (reviewed in [76–79]). The database contains profiles of 3159 cellular proteins with 578 proteins identified, and a urine database listing 339 proteins of which 124 have been identified. Comparing the SCC 2D gel protein expression profiles to the 2D gel proteome of normal biopsies, both quantitative and qualitative differences in several proteins were identified. Those not observed in the normal profiles were keratins 10, 14, 16, migration inhibitory factor-related proteins (MRPs) 8 and 14 and psoriasin [79–81].

Nuclear matrix proteins as bladder cancer specific markers have also been identified by 2DE and MS ([82,83], reviewed in [84,85]). Three of these proteins, BLCA-1, -4 and -6, were micro-sequenced and the peptide information utilized to develop specific antibodies. When comparing the BLCA-4 urine levels of bladder cancer patients with normal controls using an immunoassay, a sensitivity of 94.6% and a specificity of 100% in detecting bladder cancer were achieved [83].

Halmer et al. [86] have followed a different approach in the search for disease biomarkers: The have combined IEF, as the first dimension, with LC, as the second dimension in protein separation, with ESI-TOF and MALDI-TOF MS for protein quantification and identification. Applying this methodology to study protein lysates from benign and malignant breast epithelial cells, various potential markers for breast cancer were detected.

Besides the application of 2DE or liquid chromatography-based systems, recent studies have evaluated the surface-enhanced laser desorption/ionization time-of-flight (SELDI-TOF) mass spectrometry system in the search for disease biomarkers. This system which can be considered as a variation of the MALDI technology, has recently been developed and is based on utilization of chemically or biochemically predefined ProteinChip® Arrays as capturing surfaces for protein molecules (reviewed in [87]). The biological sample is directly applied on to the chips and proteins are captured according to their affinity for the chemical (hydrophobic, ion exchange, normal phase, immobilized metal)

or biological (carrying prebound antibodies, ligands, etc.) surface. Retained molecules are then washed with buffers of various stringencies and analyzed by MALDI-TOF. Although the resolving power of 2DE remains unchallenged, this approach is more rapid and has an increased throughput compared to 2D electrophoresis. In addition, due to the utilization of the protein chips it also overcomes several problems associated with the sample preparation for MALDI-MS, most notably enabling the direct analysis of "crude" samples such as body fluids and cell extracts.

Using this system, distinct protein patterns of normal, premalignant, and malignant cells for prostate, ovarian, esophageal, breast, and hepatic cancers have been uncovered ([87–89], reviewed in [87]). Fig. 2 shows SELDI mass spectra of sera from breast cancer patients. As shown, multiple protein peaks are differentially expressed in the cancer versus the control samples. The identity of these breast cancer-associated peaks is not yet known. Similar studies on bladder cancer have identified defensin-α as a potential marker for the disease in urine [90]. Recent reports also provide evidence that analysis of SELDI data by "learning" algorithms can lead to the identification of serum protein "fingerprints" of prostate, ovarian and breast

cancers as well as urine fingerprints for renal cell and bladder carcinomas [91–98]. In each of these studies a different computational approach was employed to analyze the SELDI data, including a genetic algorithm [93], a decision tree [91,92,95,96,98], a support vector machine algorithm [94] and neural networks [97]. Each method appeared to be effective in developing accurate classification systems, notably, without previous identification of the protein peaks that consist the cancer and non-cancer fingerprints.

As expected, besides its strengths, the SELDI approach has also several limitations. As summarized in the review by Diamandis [99] these include: (i) The identification of the protein peaks that are differentially expressed in the disease, requires their relative purification which is technically challenging; (ii) Similar to the 2D systems, the SELDI technology, unless coupled to fractionation methodologies, allows the detection of primarily the high abundance proteins; (iii) comparison of the expression levels of the protein peaks is difficult due to the inherent limitations of mass spectrometry in protein quantification [27,44]; (iv) improvements/adjustments in instrumentation as well as in the mathematical algorithms and computing resources associated with the SELDI analysis (i.e. algorithms for
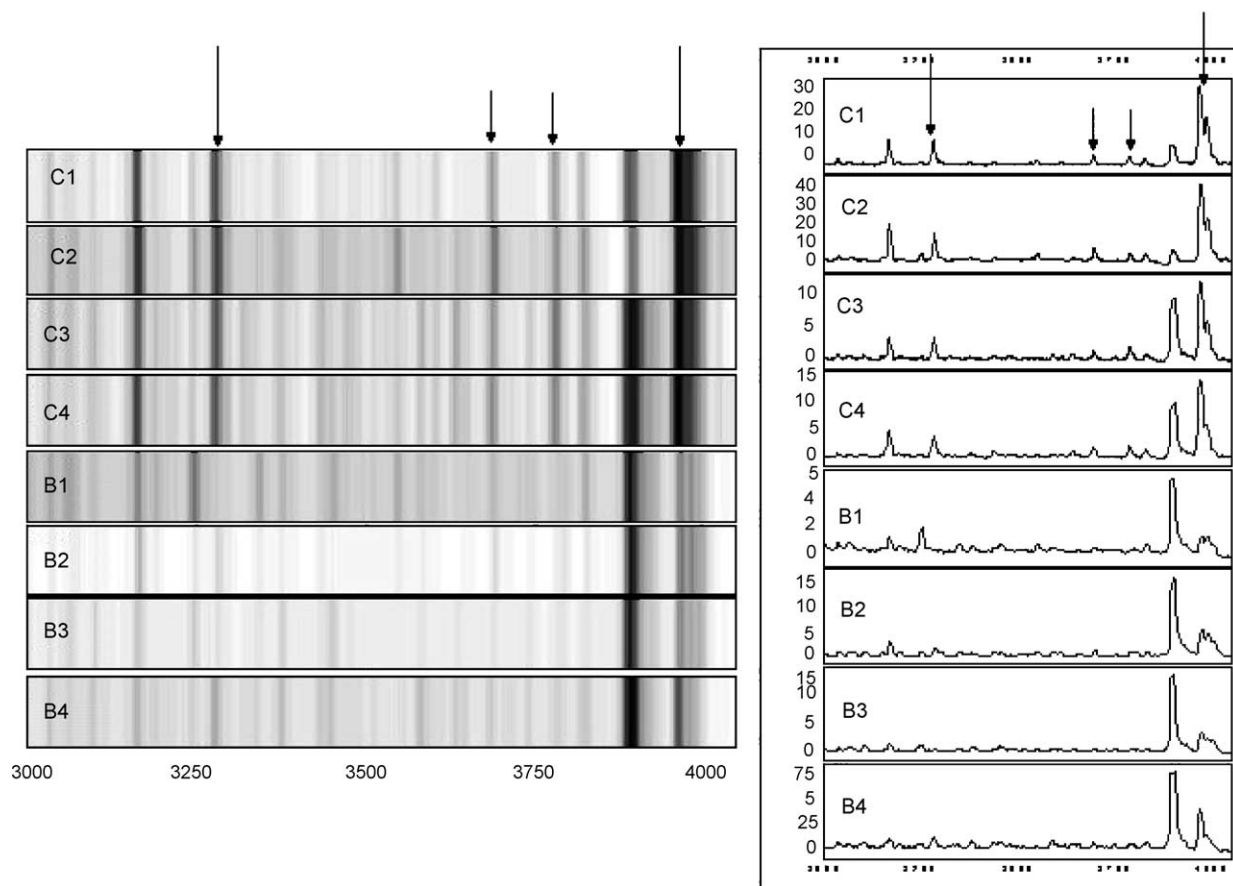

Fig. 2. Potential breast cancer biomarkers (arrows) detected by the SELDI technology. The mass spectra (right) can also be visualized in the grey scale/gel view (left) to facilitate the visualization of protein differences. Spectra from serum samples from four breast cancer patients (C1–C4) and four patients with benign breast disease (B1–B4) are shown. *m/z*: mass/charge.

peak identification, protein pattern recognition algorithms, etc.) will be required prior to the clinical application of this technology as a diagnostic/prognostic tool.

Besides the SELDI technology, imaging mass spectrometry is an exciting new technology for analysis of protein expression in mammalian tissues. Caprioli and colleagues recently demonstrated the use of this technology for directly mapping proteins present in tissue sections of various tumors [100–103]. Briefly, frozen tissue sections are applied on MALDI-plates, and following cell lysis, mass spectra are collected at different intervals. With this approach, the spatial distribution of protein peaks throughout the tissue section is provided and proteins with differential expression in the normal and diseased tissue are identified. Imaging MS shows potential for a number of applications including biomarker discovery and biomarker tissue localization as well as in providing a better correlation of the tumor histopathologic phenotype with its molecular profile.

8. Concluding remarks: future challenges

The complexity of protein molecules, the plethora of post-translational modifications and the frequent representation of proteins at very low amounts are some of the reasons that analyses at the proteomic level lagged behind genomic studies. Advances in 2D electrophoresis, chromatography and mass spectrometry technologies have provided tremendous opportunities in biological research. As shown, however, in the manuscript by Anderson et al. [104] where a comparison of four different methodologies for plasma proteome characterization is conducted, each separation methodology and MS technology has distinct strengths and limitations. A combinatorial approach that entails collaboration of different centers of expertise for each methodology, will be required for the successful characterization of the human proteome in the normal and disease state. At the same time, nano-technologies and microfluidics, currently in development, promise to facilitate the analysis of low abundant proteins.

Improvements/development of bioinformatics packages are urgently needed for the conduction of all steps of proteomic studies. In the case of 2DE, software packages have been and are being developed for spot detection, comparison, quantification, and statistical analysis of 2D electrophoresis data [105]. In the case of data acquired by MS, automatic mass assignment, denoising (i.e. removal of background), mass correction based on standard peptides and mass clustering are very crucial factors for the successful application of mass spectrometry-based approaches for protein profiling. Furthermore, the vast amount of data acquired by the contemporary mass spectrometry methods require databases and mining tools that will allow comparisons of different samples, generation of reference profiles as well as both statistical and graphical information about the sample under examination.

Data presentation, sharing, querying and retrieval present another challenge to be encountered by researchers. Various types of 2D protein maps have been generated and become available through the world-wide web [49,76]. These tools greatly facilitate research by providing protein map resources and efforts are underway to further expand these resources [106,107]. With the advent of the available high throughput approaches there is a definite need for database integration from multiple sources and development of user interfaces that allow data entry, visualization and retrieval. Simply stated, an open-minded database has to be established to allow continual additions of data and integration of information acquired by different centers of expertise.

It should not also be overlooked the fact, that the ability of new technologies to identify large number of proteins, brings up the new challenge of marker validation. In the short term, this process entails multi-institutional collaborations to address the portability of the methodology, as well as the presence of the marker in a geographically diverse cohort of patients, whereas in the long run, well-designed clinical trials will have to be established.

References

[1] P.H. O'Farrell, J. Biol. Chem. 250 (1975) 4007.
[2] B. Bjellqvist, K. Ek, P.G. Righetti, E. Gianazza, A. Görg, R. Westermeier, W. Postel, J. Biochem. Biophys. Methods 6 (1982) 317.
[3] H. Langen, D. Röder, J.-F. Juranville, M. Fountoulakis, Electrophoresis 18 (1997) 2085.
[4] A. Görg, C. Obermaier, G. Boguth, A. Harder, B. Scheibe, R. Wildgruber, W. Weiss, Electrophoresis 21 (2000) 1037.
[5] A.C. Bergman, T. Benjamin, A. Alaiya, M. Walthman, K. Sakaguchi, B. Franzen, et al., Electrophoresis 21 (2000) 679.
[6] M. Unlu, M.E. Morgan, J.S. Minden, Electrophoresis 18 (1997) 2071.
[7] T.C. Hunter, N.L. Andon, A. Koller, J.R. Yates, P.A. Haynes, J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 782 (1/2) (2002) 165.
[8] D.B. Friedman, S. Hill, J.W. Keller, N.B. Merchant, S.E. Levy, R.J. Coffey, R.M. Caprioli, Proteomics 4 (3) (2004) 793.
[9] D.P. Kreil, N.A. Karp, K.S. Lilley, Bioinformatics (2004) [Epub ahead of print].
[10] M. Fountoulakis, B. Takács, Electrophoresis 22 (2001) 1593.
[11] M. Fountoulakis, H. Langen, C. Gray, B. Takács, J. Chromatogr. A 806 (1998) 279.
[12] S.P. Gygi, G.L. Corthals, Y. Zhang, Y. Rochon, R. Aebersold, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 9390.
[13] P.G. Righetti, A. Castagna, B. Herbert, Anal. Chem. 73 (2001) 321.
[14] P.G. Righetti, A. Castagna, B. Herbert, F. Reymond, J.S. Rossier, Proteomics 3 (2003) 1397.
[15] M. Fountoulakis, M.-F. Takács, B. Takács, J. Chromatogr. A 833 (1999) 157.
[16] M. Fountoulakis, J.-F. Juranville, Anal. Biochem. 313 (2003) 267.
[17] M. Fountoulakis, J.-F. Juranville, G. Tsangaris, L. Suter, Amino Acids 26 (2004) 27.
[18] M. Fountoulakis, R. Gasser, Amino Acids 24 (2003) 19.
[19] M. Fountoulakis, B. Takács, Methods Enzymol. 358 (2002) 288.
[20] G.L. Corthals, V.C. Wasinger, D.F. Hochstrasser, J.-C. Sanchez, Electrophoresis 21 (2000) 1104.
[21] C.C. Wu, J.R. Yates III, Nat. Biotechnol. 21 (3) (2003) 262.
[22] D. Lin, D.L. Tabb, J.R. Yates III, Biochim. Biophys. Acta 1646 (1/2) (2003) 1.

[23] J. Kim, D.G. Camp, R.D. Smith, J. Mass Spectrom. 39 (2) (2004) 208.

[24] X. Yao, A. Freas, J. Ramirez, P.A. Demirev, C. Fenselau, Anal. Chem. 76 (9) (2004) 2675.

[25] D.K. Han, J. Eng, H. Zhou, R. Aebersold, Nat. Biotechnol. 19 (2001) 946.

[26] S.A. Carr, R.S. Annan, in: F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (Eds.), Current Protocols in Molecular Biology, Wiley, New York, 1998, unit 10.21, p. 10.21.1.

[27] R. Aebersold, M. Mann, Nature 422 (2003) 198.

[28] F. Hillenkamp, M. Karas, R.C. Beavis, B.T. Chait, Anal. Chem. 63 (1991) 1193.

[29] M. Mann, G. Talbo, Curr. Opin. Biotechnol. 7 (1996) 11.

[30] S. Cristoni, L.R. Bernard, Mass Spectrom. Rev. 22 (2003) 369.

[31] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science 246 (1989) 64.

[32] M. Mann, S. Shen, J.B. Fenn, in: M.L. Gross (Ed.), Mass Spectrometry in the Biological Sciences, Kluwer Academic Publishers, New York, 1992, p. 145.

[33] H.W. Lahm, H. Langen, Electrophoresis 21 (2000) 2105.

[34] D. Lin, D.L. Tabb, J.R. Yates III, Biochim. Biophys. Acta 1646 (2003) 1.

[35] R. Knochenmuss, R. Zenobi, Chem. Rev. 103 (2) (2003) 441.

[36] M. Karas, M. Gluckmann, J.J. Schafer, Mass Spectrom. 35 (1) (2000) 1.

[37] P.J. Kebarle, Mass Spectrom. 35 (7) (2000) 804.

[38] R.D. Smith, J.A. Loo, C.G. Edmonds, C.J. Barinaga, H.R. Udseth, Anal. Chem. 62 (9) (1990) 882.

[39] M. Mann, M. Wilm, Trends Biochem. Sci. 20 (6) (1995) 219.

[40] G.A. Valaskovic, N.L. Kelleher, F.W. McLafferty, Science 273 (5279) (1996) 1199.

[41] T.A. Fligge, K. Bruns, M.J. Przybylski, J. Chromatogr. B: Biomed. Sci. Appl. 706 (1) (1998) 91.

[42] S.D. Patterson, R.H. Aebersold, Nat. Genet. 33 (Suppl) (2003) 311.

[43] S.A. McLuckey, J.M. Wells, Chem. Rev. 101 (2) (2001) 571.

[44] S.D. Patterson, in: F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (Eds.), Current Protocols in Molecular Biology, Wiley, New York, 1998, unit 10.22, p. 10.22.1.

[45] M.S. Boguski, M.W. Mcintosh, Nature 422 (2003) 233.

[46] P.R. Shrinivas, S. Srivastava, S. Hanash, G.L. Wright Jr., Clin. Chem. 47 (2001) 1901.

[47] M. Fountoulakis, H. Langen, Anal. Biochem. 250 (1997) 153.

[48] M. Fountoulakis, H.-W. Lahm, J. Chromatogr. A 826 (1998) 109.

[49] M. Fountoulakis, Mass Spectrom. Rev. 23 (2004) 231.

[50] A.J. Forbes, S.M. Patrie, G.K. Taylor, Y.B. Kim, L. Jiang, N.L. Kelleher, Proc. Natl. Acad. Sci. U.S.A. 101 (9) (2004) 2678.

[51] G.K. Taylor, Y.B. Kim, A.J. Forbes, F. Meng, R. McCarthy, N.L. Kelleher, Anal. Chem. 75 (16) (2003) 4081.

[52] F. Meng, B.J. Cargile, S.M. Patrie, J.R. Johnson, S.M. McLoughlin, N.L. Kelleher, Anal. Chem. 74 (13) (2002) 2923.

[53] D.F. Hunt, J.R. Yates, J. Shabanowitzz, S. Winston, C.R. Hauer, Proc. Natl. Acad. Sci. U.S.A. 83 (1986) 6233.

[54] A. Shevchenko, I. Chernushevic, A. Shevchenko, M. Wilm, M. Mann, Mol. Biotechnol. 20 (1) (2002) 107.

[55] A. Shevchenko, I. Chernushevich, M. Wilm, M. Mann, Methods Mol. Biol. 146 (2000) 1.

[56] A.A. Alaiya, B. Franzén, G. Auer, S. Linder, Electrophoresis 21 (2000) 1210.

[57] V.E. Bichsel, L.A. Liotta, E.F. Petricoin III, Cancer J. 7 (2001) 69.

[58] S. Hanash, Nature 422 (2003) 226.

[59] E.F. Petricoin, K.C. Zoon, E.C. Kohn, J.C. Barrett, L.A. Liotta, Nat. Rev. Drug Disc. 1 (2002) 683.

[60] M. Tyers, M. Mann, Nature 422 (2003) 193.

[61] R.E. Banks, M.J. Dunn, D.F. Hochstrasser, J.-C. Sanchez, W. Blackstock, D.J. Pappin, P.J. Selby, Lancet 356 (2000) 1749.

[62] C. Rohlff, C. Southan, Curr. Opin. Mol. Ther. 4 (2002) 251.

[63] M. Fountoulakis, Amino Acids 21 (2001) 363.

[64] E. Engidawork, G. Lubec, Amino Acids 21 (2001) 331.

[65] J.W. Zolg, L.H. Langen, Mol. Cell. Proteomics 3 (4) (2004) 345.

[66] S. Greber, G. Lubec, N. Cairns, M. Fountoulakis, Electrophoresis 20 (1999) 928.

[67] M. Fountoulakis, N. Cairns, G. Lubec, J. Neural Transm. Suppl. 57 (1999) 323.

[68] G. Lubec, O. Labudova, N. Cairns, M. Fountoulakis, Neurosci. Lett. 260 (1999) 141.

[69] G. Lubec, O. Labudova, N. Cairns, P. Berndt, H. Langen, M. Fountoulakis, J. Neural Transm. Suppl. 57 (1999) 21.

[70] G. Lubec, M. Nonaka, K. Krapfenbauer, M. Gratzer, N. Cairns, M. Fountoulakis, J. Neural Transm. Suppl. 57 (1999) 161.

[71] B.C. Yoo, M. Fountoulakis, N. Cairns, G. Lubec, Electrophoresis 22 (2001) 172.

[72] B.C. Yoo, S.H. Kim, N. Cairns, M. Fountoulakis, G. Lubec, Biochem. Biophys. Res. Commun. 280 (2001) 249.

[73] E. Engidawork, T. Gulesserain, M. Fountoulakis, G. Lubec, Neuroscience 122 (2003) 145.

[74] R. Weitzdoerfer, D. Stolzlechner, M. Dierssen, J. Ferreres, M. Fountoulakis, G. Lubec, J. Neural Transm. Suppl. 61 (2001) 347.

[75] M.D. Johnson, L.R. Yu, T.P. Conrads, Y. Kinoshita, T. Uo, J.D. Matthews, S.W. Lee, R.D. Smith, T.D. Veenstra, R.S. Morrison, J. Biol. Chem. (2004) [Epub ahead of print].

[76] J.E. Celis, M. Ostergaard, H.H. Rasmussen, P. Gromov, I. Gromova, H. Varmark, H. Palsdottir, N. Magnusson, I. Andersen, B. Basse, J.B. Lauridsen, G. Ratz, H. Wolf, T.F. Orntoft, P. Celis, A. Celis, Electrophoresis 20 (1999) 300.

[77] J.E. Celis, P. Gromov, Cancer Cell 3 (2003) 9.

[78] P.S. Gromov, M. Ostergaard, I. Gromova, J.E. Celis, Prog. Biophys. Mol. Biol. 80 (2002) 3.

[79] J.E. Celis, H. Wolf, M. Ostergaard, Electrophoresis 21 (2000) 2115.

[80] M. Ostergaard, H.H. Rasmussen, H.V. Nielsen, H. Vorum, T.F. Orntoft, H. Wolf, J.E. Celis, Cancer Res. 57 (1997) 4111.

[81] M. Ostergaard, H. Wolf, T.F. Orntoft, J.E. Celis, Electrophoresis 20 (1999) 349.

[82] R.H. Getzenberg, B.R. Konety, T. Oeler, M.M. Quigley, A. Hakam, M.J. Becich, R.R. Bahson, Cancer Res. 56 (1996) 1690.

[83] B.R. Konety, T.-S.T. Nguyen, G. Brenes, A. Sholder, N. Lewis, S. Bastacky, D.M. Potter, R.H. Getzenberg, J. Urol. 164 (2000) 634.

[84] P. Oehr, Anticancer Res. 23 (2003) 805.

[85] T. Davido, R.H. Getzenberg, J. Cell Biochem. Suppl. 35 (2000) 136.

[86] R.L. Hamler, K. Zhu, N.S. Buchanan, P. Kreunin, M.T. Kachman, F.R. Miller, D.M. Lubman, Proteomics 4 (3) (2004) 562.

[87] G.R. Wright Jr., Expert Rev. Mol. Diagn. 2 (2002) 549.

[88] L.H. Cazares, B.-L. Adam, M.D. Ward, S. Nasim, P.F. Schellhammer, O.J. Semmes, G.L. Wright Jr., Clin. Cancer Res. 8 (2002) 2541.

[89] C.P. Paweletz, B. Trock, M. Pennanen, T. Tsangaris, C. Magnant, L.A. Liotta, E.F. Petricoin III, Dis. Markers 17 (2001) 201.

[90] A. Vlahou, P.F. Schellhammer, S. Mendrinos, K. Patel, F.I. Kondylis, L. Gong, G.L. Wright Jr., Am. J. Pathol. 158 (2001) 1491.

[91] B.-L. Adam, Y. Qu, J.W. Davis, M.D. Ward, M.A. Clements, L.H. Cazares, O.J. Semmes, P.F. Schellhammer, Y. Yasui, Z. Feng, G.L. Wright Jr., Cancer Res. 62 (2002) 3609.

[92] Y. Qu, B.L. Adam, Y. Yasui, M.D. Ward, L.H. Cazares, P.F. Schellhammer, Z. Feng, O.J. Semmes, G.L. Wright Jr., Clin. Chem. 48 (2002) 1835.

[93] E.F. Petricoin, A. Ardekani, B. Hitt, P.J. Levine, V.A. Fusaro, S.M. Steinberg, et al., Lancet 359 (2002) 572.

[94] J. Li, Z. Zhang, J. Rosenzweig, Y.Y. Wang, D.W. Chan, Clin. Chem. 48 (2002) 1296.

[95] A. Vlahou, J. Schorge, B.W. Gregory, R. Coleman, J. Biomed. Biotechnol. 5 (2003) 308.

[96] A. Vlahou, C. Laronga, L.L. Wilson, B. Gregory, D. McGaugher, R.R. Perry, G.L. Wright Jr., O.J. Semmes, Clin. Breast Cancer 4 (2003) 203.

[97] M.A. Rogers, P. Clarke, J. Noble, N.P. Munro, A. Paul, P.J. Selby, R.E. Banks, Cancer Res. 63 (2003) 6971.

[98] A. Vlahou, A. Giannopoulos, B.W. Gregory, T. Manousakas, F.I. Kondylis, L.L. Wilson, P.F. Schellhammer, O.J. Semmes, G.L. Wright Jr., Clin. Chem. 50 (8) (2004) 1438.

[99] E.P. Diamandis, Mol. Cell. Proteomics 3 (2004) 367.

[100] M. Stoeckli, P. Chaurand, D.E. Hallahan, R.M. Caprioli, Nat. Med. 7 (2001) 493.

[101] P. Chaurand, S.A. Schwartz, R.M. Caprioli, J. Proteome Res. 3 (2) (2004) 245.

[102] P. Chaurand, S.A. Schwartz, D. Billheimer, B.J. Xu, A. Crecelius, R.M. Caprioli, Anal. Chem. 76 (4) (2004) 1145.

[103] P. Chaurand, S.A. Schwartz, R.M. Caprioli, Curr. Opin. Chem. Biol. 6 (5) (2002) 676.

[104] N.L. Anderson, M. Polanski, R. Pieper, T. Gatlin, R.S. Tirumalai, T.P. Conrads, T.D. Veenstra, J.N. Adkins, J.G. Pounds, R. Fagan, A. Lobley, Mol. Cell. Proteomics 3 (4) (2004) 311.

[105] P. Berndt, U. Hobohm, H. Langen, Electrophoresis 20 (1999) 3521.

[106] M. Fountoulakis, B. Takács, Protein Expr. Purif. 14 (1998) 113.

[107] H. Langen, B. Takács, S. Evers, P. Berndt, H.-W. Lahm, B. Wipf, C. Gray, M. Fountoulakis, Electrophoresis 21 (2000) 411.