Exp Oncol 2008
30, 3, 171–180

REVIEWS

Experimental ONCOLOGY

PROTEOMIC APPROACHES IN BIOLOGICAL AND MEDICAL SCIENCES: PRINCIPLES AND APPLICATIONS

*P. Conrotto[1], *, S. Souchelnytskyi[2]*

[1]*Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zurich, Switzerland*

[2]*Karolinska Biomics Center, Dept. Oncology-Pathology, Karolinska Institutet, S17177 Stockholm, Sweden*

After the first introduction of the concept of "proteome" more than 10 years ago, large-scale studies of protein expression, localization, activities and interactions have gained an exponential increase of interest, leading to extensive research and technology development. Proteomics is expansively applied in many areas, ranging from basic research, various disease and malignant tumors diagnostic and biomarker discovery to therapeutic applications. Several proteomics approaches have been developed for protein separation and identification, and for the characterization of protein function and structure. Two-dimensional gel electrophoresis, chromatography, capillary electrophoresis and mass spectrometry have become the most used proteomics methods. These techniques are also under constant development. This review provides an overview of the main techniques and their combinations, used in proteomics. The emphasis is made on description of advantages and disadvantages of each technique, to navigate in selection of the best application for solving a specific problem.

Key Words: **proteomics, mass spectrometry, chromatography, two-dimensional gel electrophoresis, protein arrays, capillary electrophoresis.**

Advances in DNA analysis lead to annotations of genomes from human and several other species (http://www.gdb.org/), enabling to enter in the post-genomic era [1]. In this era, proteomics became a powerful tool for the understanding of cellular processes, as well as for clinical applications. The old correlation "1 gene — 1 protein" being overcome, it's now clear that the number of proteins produced by cells is subjected to many types of regulations. Splicing variants, fusions, post-translational modifications, all contribute to a highly dynamic and, at the same time, tightly regulated proteomics profiles [2]. The onset of pathological conditions directly correlates with an altered proteome; an understanding of these alterations and their causes will allow designing more efficient and specific therapies [3].

A rapidly emerging set of technologies is making possible to identify large numbers of proteins in complex mixtures, to map their cellular and subcellular interactions and to investigate their biological activities. In the last years proteomics techniques have been significantly improved and standardized, and new approaches have been introduced. Combinations of two-dimensional gel electorphoresis (2DE) with MALDI TOF mass spectrometry (MS), liquid chromatography (LC) with ESI MS, and capillary electrophoresis with ESI MS are of particular interest due to their frequent applications. The first combination, 2DE and MS, enables separation of proteins in a polyacrylamide gel followed by identification by MS. The coupling with MS is used for two main goals: the first is a large scale identification of proteins in complex mixtures, and the second is analysis of differential protein expression among two or more samples. The second combination is LC-ESI MS. LC resolves peptides after protease digestion of complex biological samples: each peptide retains a distinctive affinity toward a hydrophobic matrix, and it is eluted under specific hydrophilic conditions of a gradient buffer. The third combination is capillary electrophoresis and ESI MS: although less spread, it separates with high efficiency both proteins and peptides by electrophoresis in a liquid phase. Single peptides from LC or CE are then ionized and analyzed with MS. Protein microarray is the fourth often used proteomics technique, which gains high attention, especially for clinical applications. In protein microarrays the interaction of a protein to the cognate substrate(s) allows its isolation and/or characterization. These proteomics techniques have both strong sides and limitations; their applications can be discriminating in efficiently supplying information regarding the protein(s) of interest. Here these four main proteomics approaches are described, their advantages and disadvantages highlighted and examples of experimental applications are provided.

Received: May 14, 2008.

*Correspondence: Fax: 0041 446331358
E-mail: paolo.conrotto@pharma.ethz.ch

Abbreviations used: DIGE – difference in gel electrophoresis; ESI – electrospary ionization; FTICR – Fourier transform-ion cyclotron resonance; ICAT – isotope-coded affinity tag; iTRAQ – isobaric tag for relative and absolute quantitation; MALDI – matrix-assisted laser desorption ionization; MRM – multiple reaction monitoring; TOF – time of flight; SELDI – surface-enhanced laser desorption/ionization.

I. TWO DIMENSIONAL GEL ELECTROPHORESIS — MASS SPECTROMETRY

2DE allows separation of complex protein mixtures in two steps: in the first dimension proteins are resolved according to their isoelectric point (pI), and in the second dimension according to their molecular mass. Various gel image analysis softwares are used to detect differentially expressed proteins which are then identified by MS (Fig. 1) [4].

1. Technique overview. *A) Preparation of biological sample* - Biological samples undergo lysis, with inactivation of proteases, extraction of proteins, and eventually protein pre-fractionation. After this, proteins must be denatured and reduced in order to disrupt inter- and intramolecular interactions. Chaotropes, such

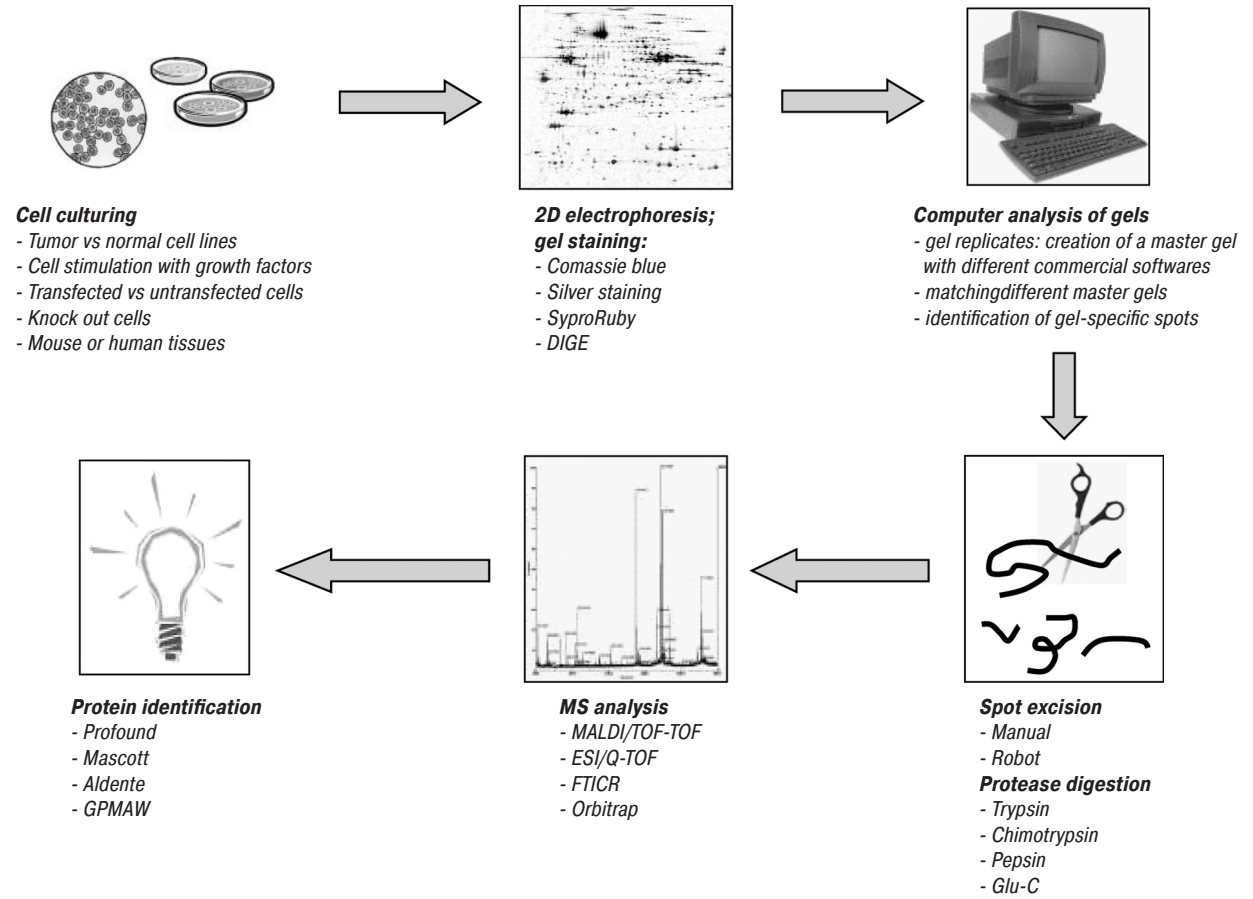

Fig. 1. Workflow of 2 dimensional gel electrophoresis coupled with mass spectrometry analysis

as urea or thiourea, break hydrogen bonds, leading to protein unfolding and denaturation. The addition of dithiothreitol (DTT) reduces sulfhydryl groups. Different detergents are added to keep proteins in solution: anionic surfactants such as sodium dodecyl sulphate (SDS) are efficient, but interfere with isoelectrofocusing, impairing spot resolution. Thus, non-anionic detergents (NP-40, Triton X-100) and particularly zwitterionic detergents, such as CHAPS, are used.

B) Protein separation — Protein mix is separated in a two-steps process. In the first dimension (isoelectric focusing, IEF) a pH gradient is generated. Two main alternatives are possible: 1) carrier ampholyte (CA), in which synthetic compounds covering a wide pH range are added to a polyacrylamide gel and the pH gradient is experimentally generated; 2) immobilized pH gradient (IPG), in which a series of bifunctional immobiline reagents with a broad pK spectra are chemically linked to an acrylamide matrix strip [4, 5]. Pre-casted IPG strips are predominantly used due to the formation of stable and reproducible pH gradient. Moreover, IPG strips can cover a large pH range (typically 3 to 10), or a narrow one (1 or 2 pH units; for ex 5.3–6.3; 6.1–7.1); in the last case a progressive series of strips is used, and corresponding gels are electronically merged (zoom-gel) [6], thus increasing the final number of detectable proteins. Upon application of an electric field, proteins migrate along the strip until they reach the point where the pH of the gel equals their isoelectric point: in that area proteins become neutral and stop to migrate. In the second dimension a polyacrylamide gel electrophoresis (PAGE) is carried out. Strips must be equilibrated with SDS to confer a negative charge to proteins, in combination with DTT and iodoacetamide (IAA), in order to ensure the complete reduction of disulphide bonds and the alkylation of cysteine residues, respectively.

C) Protein detection — Proteins, separated by 2DE, may then be stained in different ways. Comassie Blue is common, inexpensive, quick and easy to use, but not very sensitive (detection limit of 100–200 ng/spot). Metal cations (copper, zinc) give a negative staining by reacting with SDS all over the gel except dots with proteins; they are much less common, although they can detect proteins at 10 ng/spot. Silver staining is highly sensitive (1–2 ng/spot) and widely used in proteomics, but it is more time-consuming, and the staining is not quantitative over a broad range of protein amount. Fluorescent dyes like SyproRuby are as sensitive as silver staining, plus a high precision in protein quantification over a large range of concentrations; on the other hand they are very expensive, need additional instruments for detection, and make spot cutting more elaborated. In the "difference gel electrophoresis" (DIGE) [7] 2 diverse samples (plus an internal standard) are *in vitro* labeled with fluorescent dyes (CyDyes) differing in their excitation and emission wavelengths. The two samples and the internal standard are then mixed, loaded on the same gel and separated. Different fluorescence allows the identification and comparison of spots from two different samples in the

same gel. This technique has the considerable advantage to compare 2 samples visualized in the same gel with high sensitivity (1 ng/spot); however, it is relatively expensive, requires a specific instrumentation and proteins are chemically modified prior to separation and identification, with possible alteration of results. Pro-Q Diamond is a florescent dye which labels exclusively phosphorylated proteins; it is sensitive (1–16 ng/spot, according the phosphorylation level of the protein) and highly specific, making it suitable for signal transduction studies; commonly it must be used in combination with a more generic staining. Similarly, radioactive isotopes can be incorporated in sample before loading: after the separation, only proteins carrying the specific chemical group (*i. e.* phosphoric, sulphonic group) are detected, with high sensitivity (10–100 pg/spot).

D) Protein identification — According to the staining procedures used, 2D spot maps are generated and subjected to comparative image analysis with appropriate software [8]. Selected spots are then cut out from the gel and digested with protease(s). The pool of peptides produced from each protein is first desalted using specific tips partially filled with an hydrophobic resin (ZipTip) [9], then analyzed by MS. Mass spectrometers can differ mainly by the ionization of the analytes (MALDI and ESI are the 2 most used in proteomics), and by the separation of ions (TOF, quadrupole, ion trap; for MS theories and principles see [10, 11]). Mass spectrometer identifies the mass of each peptide, generating a list of measured masses. Matching the experimental list with the predicted mass lists from "in silico" digestion of all known proteins allows the identification of the protein(s) in the studied sample (peptide mass fingerprinting, PMF). An additional approach may take advantage of a combination of two (or more) MS in series (MS/MS, MS^n), in order to obtain the amino acid sequence of a selected peptide [11]. In this case, in the first MS peptides are separated according their mass-to-charge ratio; only a single selected peptide is allowed to enter the second MS, where it is fragmented and its amino acid sequence is determined. For detailed description of molecular mechanisms of fragmentation of peptides in mass spectrometry, see [12].

2. Applications. 2DE-MS combination is often used when analysis of full-length proteins is required. This combination is used in expression proteomics, analysis of post-translational modifications and protein complexes.

Post-translational modifications (PTMs) shift the mass (and eventually the charge) of the protein; 2DE, by analyzing full proteins, can show different protein isoforms (originating from PTMs, alternative splicing or protein rearrangement), and it is therefore particularly suitable for these types of analysis. In a study of a liver cell line, Liu and coworkers combined the Pro-Q Diamond fluorescence dye for phosphoprotein detection with standard Comassie blue staining, followed by MS analysis [13]; 269 phosphorylated proteins expressed by the cell line were identified, together with the ratio phosphorylated/unphosphorylated for each protein.

2DE-MS applications to expression proteomics are numerous. As an example, by comparing 2DE from cells infected and non-infected by IBD virus, followed by MS identification, signal pathways specifically altered by the pathogen were identified [14]. The authors took advantage of proteins displayed on gel to point out differences in healthy versus diseased cells.

Protein complex formation is an important mechanism for regulation of protein activities. After pull-down of the exogenously-expressed constitutive active form of the Transforming Growth Factor-β type I receptor (TβRI), 2DE/MS allowed to identify new molecules which interact with the receptor, and therefore new possible functions controlled by TβRI [15]. Since a receptor interactome is restricted, but each component may play a pivotal function, the high identification rate achieved by 2DE (compared to other techniques) fits well with these types of experiments.

3. Advantages and Disadvantages. 2DE/MS is a well-established technique, with well developed scientific, reagents and instrumental basis. The trademark of 2DE is the possibility to separate proteins according to isoelectric point and molecular mass. This allows the researchers to work with highly complex protein mixtures, such as whole cell lysates, and to resolve routinely 2000–4000 proteins at once [4]. The possibility to employ IPG strips in a very acidic or basic range and to have PAGE specifically set up for very high or very low-Mr proteins, allow to broad the pool of detectable cellular proteins. Of major importance, 2DE gives a map of intact proteins, and the subsequent MS analysis is univocally related to each gel spot; moreover, size and pI of spots are an additional internal confirmation of MS data. This feature has 2 main applications: a comparison of samples from different metabolic or pathological conditions by semiquantitative analysis of protein expression levels; and the study of PTMs. The identification of a protein in more than 1 spot, or in a spot with an unexpected molecular mass or pI, can lead to the characterization of PTMs, and eventually correlate the protein modification with a specific cellular context (*i. e.* tumor environment). Moreover, for protein identification, it is possible to achieve both PMF study and MS/MS peptide sequences; in each case the choice among different MS variants and strategies [10] permits to draw the best experimental conditions. Since after 2DE each spot is treated separately, it is possible to apply chemical modifications which increase protein coverage and efficiency of peptide fragmentation [16]. Furthermore, a protein extracted from the gel can be analyzed by MS, eluted from the MS target, chemically modified and analyzed a second time by MS, in order to extract the maximum of informations [17]. The development of software for 2D gel image analysis and robots for the recover and treatment of selected spots enables to perform large scale experiments in a relatively short time and with good reproducibility. Overall, the combination of 2DE with MS is quite efficient, with a 50–70% of protein identification rate.

On the other side some drawbacks may impair final results. For high confidence protein identification, well-defined spots in polyacrylamide gels are required. Different factors, such as high salt concentration or the presence of polysaccharides or nucleic acids, affect protein separation. The presence of high abundant proteins in the complex mixture (*i. e.* albumin in serum) may interfere with detection of low abundant proteins, and a selective removal of the former ones must be performed. Membrane proteins, carrying highly hydrophobic regions, are difficult to solubilize in the gel loading buffer, and therefore they are often lost. Although the separation is carried out according two parameters, the presence of overlapping spots can not be excluded.

Thus, 2DE-MS combination offers robustness of 2DE and variability in application of MS techniques, including peptide chemistries. This combination is well suited to projects where analysis of full-length proteins is essential, whether it is expression, PTMs or complex formation (Table 1).

II. LIQUID CHROMATOGRAPHY — MASS SPECTROMETRY

LC-MS is often used to work with peptides originated from complex mixture of proteins. A typical example of LC-MS application is a shotgun approach where non-separated proteins are digested and the mixture of peptides from all proteins is subjected to LC and then MS.

LC is used for the separation of peptide mixtures prior to MS analysis. Even if different types of chromatography could be carried out (affinity, ion exchange, hydrophobic, gel filtration chromatography), for proteomics studies reverse phase-high pressure LC (RP-HPLC) is commonly used. Peptides eluted from the column are directly (on-line) conveyed to mass spectrometer, where they are identified by PMF and/or sequenced by fragmentation. This technique is particularly suitable for high-throughput peptide identification starting from complex protein mixtures, *i. e.* whole cell extracts (Fig. 2).

1. Technique overview. *A) Preparation of biological sample* — As standard procedure, proteins are extracted from biological samples. After purification and desalting, proteins are digested with protease. Typically, thousands of peptides, belonging to all proteins originally present in the sample, are mixed in a solution.

B) Peptide separation-chromatographic column — The peptide mixture is loaded onto the chromatographic column and subjected to HPLC separation. Many types of columns are available, varying both in length and diameter. Usually, RP columns are packed with silica particles, which can differ in diameter and pore size (http://www.ionsource.com/tutorial/chromatography/rphplc.htm). An optimal combination of all these parameters regarding column and silica particle dimensions is required for high efficiency separation of peptides (http://www.forumsci.co.il/HPLC/lcms_page.html). The stationary phase, which interacts with and captures the analyte, is made of hydrophobic alkyl chains; the length of these chains can vary from 4 carbon atoms (C4), used for proteins, to C8 and C18/C20, for peptides or small molecules. The direct use of underivatized silica particles as stationary phase has also been suggested [18]. The reverse phase solvents for HPLC are conventionally the aqueous solvent, and the organic solvent (acetonitrile, methanol, propanol). Both solvents contain an acid at low concentration, *e. g.* 0.1%: this improves the chromatographic peak shape and provides a source of protons for MS ionization. For an efficient separation, samples must be reconstituted in the aqueous buffer before loading. The elution is carried out with an

Table 1. Comparison of principle proteomic techniques. Advantages and disadvantages of the 4 main combinations of proteomics techniques are indicated

Technique	Advantages	Disadvantages
2D / MS	MS identifies peptides originating mainly from a single protein; identification of the protein more direct	Salt ions could interfere with protein separation
	Analysis of complex protein mixtures	Polysaccharides and nucleic acids impair the gel quality
	Study of intact proteins; match of identified proteins with experimentally measured Mr and pH	Suppression of signal by high abundant proteins
	Semi-quantitative measurement of protein expression	Problematic analysis of very low- or very high-Mr proteins
	Possibility to make chemical modification on a single protein (to improve identification or sequencing)	Manual work required
	Possibility to make more than 1 independent MS analysis on the same protein sample	
LC / MS / MS	High reproducibility	MS identify peptides; final software analysis groups peptides belonging to a protein; identification of proteins less direct
	High sensitivity	Low protein identification rate (<10%)
	Automation is developed, a little manual work is required	Samples can not be dissolved in organic solvents or detergent-containing solutions
	High throughput experiments	Many peptides eluted from LC in a narrow range
	Unique software might control both LC and MS	No match with experimentally measured Mr and pH
CE / MS	Good protein separation	Manual work required
	Very small volumes required	Low reproducibility,
	High sensitivity (samples concentrated during column loading)	Risk of protein adsorption to wall capillary (extreme pH needed in these cases)
	Broad range of applications (supramolecular structures, proteins, peptides, small artificial compounds)	
	Insensitive towards interfering substances (inorganic or organic)	
	Suitable for coupling with other methods	
Protein microarray	Specific for protein interactions with many class of molecules	Detects only known proteins
	Study of interactions along a time frame (dynamic interactions)	Antibody specificity
	Study of interactions under different physiological conditions	Uses biased capturing agents
	Large spectrum of applications	Difficult to have protein in native conformation
		Need of high-density microarray

increasing gradient of organic buffer, with run time and flow rate as critical parameters (http://www.ionsource.com/tutorial/chromatography/rphplc.htm). Several variants and implementations have been reported. A commonly used strategy is the multi-dimensional chromatography, where a first dimension, usually the strong cation-exchange (SCX) chromatography, is combined with the second dimension RP-HPLC [19]. Another approach is the nanoLC, where a microcapillary is used as column and an ultra-low flow is applied; this allows reaching the level of low femtomole/attomole sensitivity [20, 21]. Peptides eluted from LC must undergo ionization in order to be analyzed with MS.

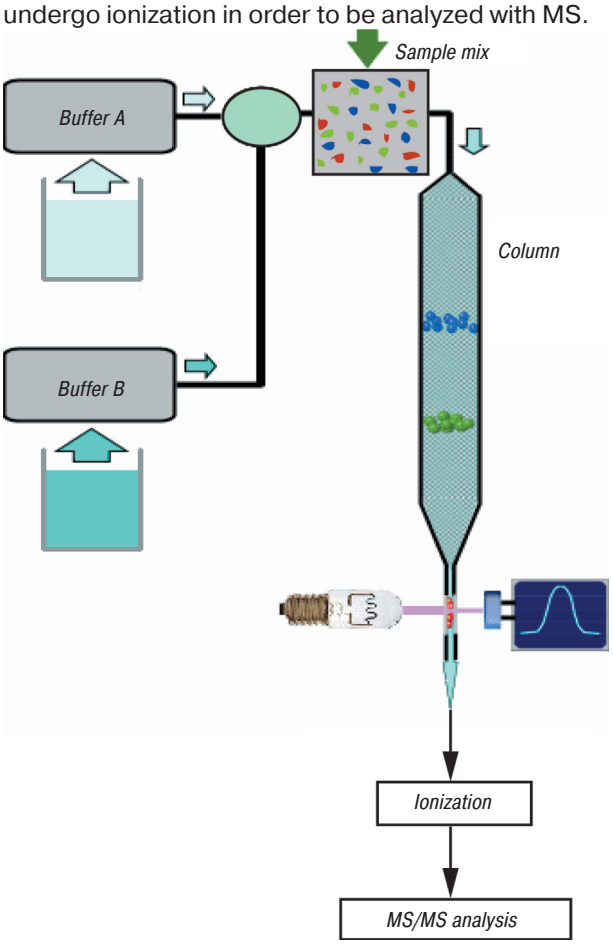

Fig. 2. Schematic representation of a liquid chromatography coupled to MS analysis. A gradient of buffer A / buffer B is generated. The sample mix is loaded, and separated in the column according the peptide affinity to hydrophobic phase. Peptides coming out from the column are registered, and ionized prior to mass analysis

C) LC-MS coupling — Chromatographic columns, producing a continuous flow of peptides, are usually coupled to ESI source, with separation and detection of peptides in various mass spectrometry instruments, such as Fourier transform-ion cyclotron resonance (FTICR) MS or quadrupole MS. Coupling nanoLC with MALDI have also been reported [22]. After ionization, charged peptides are separated in the first MS analyzer, and then fragmented and sequenced in the second MS [23].

2. Applications. LC-MS combination is particularly suitable for high throughput studies, as it allows high automation of sample separation and identification. Two-dimensional chromatography and MS/MS was applied to profiling of mice brain microsomes. More than 1900 different proteins were identified, and allowed to build networks of proteins involved in specific metabolic pathways [24].

LC-MS is also employed for identification of diagnostic and therapeutic biomarkers in high throughput screening. Soltermann and colleagues enriched, on a chromatographic column, N-linked glycoproteins, extracted from lung adenocarcinoma pleural effusions; the subsequent LC fractionation and MS analysis identified several disease-specific glycoproteins, some of them expressed at low concentration [25]. In this study, the authors took advantage of chromatography to selectively enrich glycosylated proteins on the column, prior to separation and identification.

3. Advantages and disadvantages. The LC fractionation step is often on-line coupled with a MS/MS instrument, and unique software controls both processes. The possibility of automation gives many advantages in term of reproducibility and standardization. LC-MS/MS is particularly suitable for high throughput studies of peptides, and many samples can be analyzed in series with limited manual work. Moreover, it is a sensitive technique, since studied molecules (*i.e.* peptides) are directly ionized and analyzed by MS.

Major limitations are related to the nature of chromatography: analytes have to be stable in solutions used in LC and should not interfere with the stationary phase and do not contain detergents. As the proteins eluted from the column are directly delivered into the ionization source, quantitative analysis is not trivial; radioactive labeling of sample with isotope coded affinity tag (ICAT) [26] can be the answer. Other solutions are iTRAQ and MRM, which allow quantification of peptides from different samples (iTRAQ) [27] and simplify MS spectra (MRM) for studied peptides [28]. In LC-MS/MS peptides, and not proteins, are separated by chromatography (in one or more dimensions) and ionized in the MS; only in the final step, with software analysis, single informations are merged in order to obtain the identification of proteins originally present in solution. Although the generation of the amino acid sequences provides tags for protein recognition, MS/MS analysis produces clear and complete spectra only for a limited number of peptides. In many situations software can only suggests possible sequences for each spectra, and a rate of false positives must always be considered. Proteins are often considered identified when only 2 peptides were detected by MS, strongly impairing the reliability of data. Furthermore, in LC-MS it is not possible to match the proteins identified by software with their Mr and pI, as well as it is questionable to discriminate between the same protein with various PTMs. Overall, the protein identification rate is relatively low, between 5–10% of the spectra produced, which is compensated but the large number of spectra generated.

Thus, LC-MS combination is suitable for large-scale studies requiring automation and high throughput. LC-MS can generate huge amount of MS spectra, it may be used to identify specific sites of PTMs, and

has developed tools for comparative quantification of peptides in different samples (see Table 1).

III. CAPILLARY ELECTROPHORESIS — MASS SPECTROMETRY

In capillary electrophoresis, proteins or peptides are resolved by electrophoresis in a liquid phase. Single molecules exiting the electric field are detected by a sensor upon passing a detector, and are visualized as an electrophoregram. On-line coupling with the ionization source (*e. g.* ESI) conveys peptides to MS for identification. Although the term CE is often used to indicate capillary zone electrophoresis (CZE), it refers to a family of related techniques (Table 2) [29, 30].

Table 2. Capillary Electrophoresis variants. List of different types of capillary electrophoresis; information about proteins which can be obtained with these types of CE are indicated

Method	Information produced
CZE (Capillary Zone Electrophoresis) / MS	Charge-to-size ratio MW
CD (Constant Denaturant) / CE / MS	Isomer separation
CIEF (Capillary Isoelectric Focusing) / MS	Isoelectric point versus MW
CGE (Capillary Gel Electrophjoresis) / MS	Size versus MW
CITP (transient capillary isotachophoresis)	Concentration of sample
ACE (Affinity Capillary Electrophoresis) / MS	Affinity to a given ligand

1. Technique overview. *A) Sample preparation and loading of electrophoretic capillary* — Proteins, extracted from biological sample, are digested with protease. The loading of the column, called "sample stacking", is a critical point, since it is necessary to have the sample concentrated in a relative small volume, in order to obtain sharp peaks in the electrophoregram. The simplest strategy is to inject a sample in a buffer with a significantly lower conductivity than the running medium (called background electrolyte, BGE). Upon electrical potential, the electric field in the loading buffer is stronger than in the BGE, and the sample concentrates when entering the BGE (as it happens between stacking and running gel in gel electrophoresis). Capillary isotachophoresis (CITP [31]) or the preconcentration on solid phase [32] are alternative strategies.

B) CE notes — The electroosmosis is the theory behind CE [29]. The electroosmosis flow is a direct consequence of the surface charge on the wall of fused-silica capillary. The capillary surface carries silanol groups (pKa 3–5), which easily ionize at standard buffer conditions. The dissociation to silanate ions (SiO^-) creates a negatively charged surface, compensated by a layer of positively charged molecules from the BGE. Upon an electric field, the tangential flow of the medium over the capillary surface induces a distortion of the ion distribution, which ultimately results in the flow of the solution towards the detector. The interactions of peptides or proteins with capillary walls must be as little as possible. This could be achieved in two ways: 1) by tuning the BGE pH; 2) by coating capillary with very inert materials [33].

C) CE / MS coupling — After flowing and separating along the electric field, peptides exit the capillary column and they must be ionized for MS analysis. There are 2 main CE / MS interfaces [34]: with and without supplemental fluid. The formers improve system stability and are less sensitive to BGE composition, but with lower reproducibility; the latters offer higher sensitivity, but BGE must be compatible with MS. ESI is the most frequently used ionization method in combination with CE, allowing a direct passage of ions from liquid to gas phase, and making it suitable for analysis of high Mr molecules. Improvements led to the combination of CE with other separation techniques: LC and capillary isoelectric focusing (CIEF) are often coupled with CZE [35, 36].

2. Applications. CE revealed to be an excellent and sensitive approach for separation of proteins with similar size and charge. Moreover, proteins prone to form supramolecular complexes in buffers were successfully disaggregated and separated in CE, and the very high identification rate allowed identifying many -if not all- the proteins present in complexes. Among others, CE is used in clinical studies for drug characterization. In his review, Hartinger summarizes the use of CE / MS for studying the properties of metal-based drugs and drug candidates [37].

Zeins are a heterogeneous family of storage proteins in maize, assembled by many polypeptide combinations. Separation of Zein subunits with CE followed by MS allowed to identify known and unknown components, and to compare data from natural and transgenic maize [38].

The diffusion of pharmaceuticals or contaminants in food-producing animals is receiving enormous attention, and controls are routinely done on different types of artificial compounds. CE / MS, providing several benefits compared to other techniques (short times required, no need of expensive reagents, different classes and types of drugs separated and identified in a single experiment), is increasingly used for these analysis. In a study on fishes and bovines, Juan-Garcia and colleagues screened 12 antibiotics, some of them with similar structure, and they were able to achieve the detection of 20 μg of antibiotic per kilogram of animal sample [39].

3. Advantages and disadvantages. CE is a high sensitive technique, which permits fast and cheap separation of molecules with similar mass and charge. It is relatively insensitive towards interfering substances, like lipids, carbohydrates, and salts. CE is a non-perturbing technique, and it can separate big complexes such as polymers, nanoparticles or viruses as well as small artificial compounds; particularly, both full proteins and peptide mixtures can be efficiently resolved, making it suitable for a wide spectrum of analysis [40, 41]. In addition, small volumes are used (nanoliter range, compared to microliter range of LC), and larger volumes can be concentrated during the loading, thus strongly improving the overall sensitivity (see Table 1).

On the other side CE requires manual work, and little automation is possible. Technical characteristics, like electroosmosis flow variations, impair data reproducibility. Proteins could interact with capillary wall, thus perturbing the ion stream and the overall separation; to overcome this problem extreme pH is employed, with possible alteration and/or damage of proteins.

IV. PROTEIN MICROARRAYS

Protein microarrays are based on the interaction of known proteins with specific interacting, or capturing, molecule(s). This technique is gaining more and more interest for its high sensitivity and specificity, and it is now suitable for a broad range of applications.

1. Technique overview. *A) Microarray types* — Protein microarrays may be used in a wide spectrum of analysis, like antibody-antigen, protein-protein, protein-lipid, protein-nucleic acid, enzyme-substrate and protein-small molecules interactions (Fig. 3).

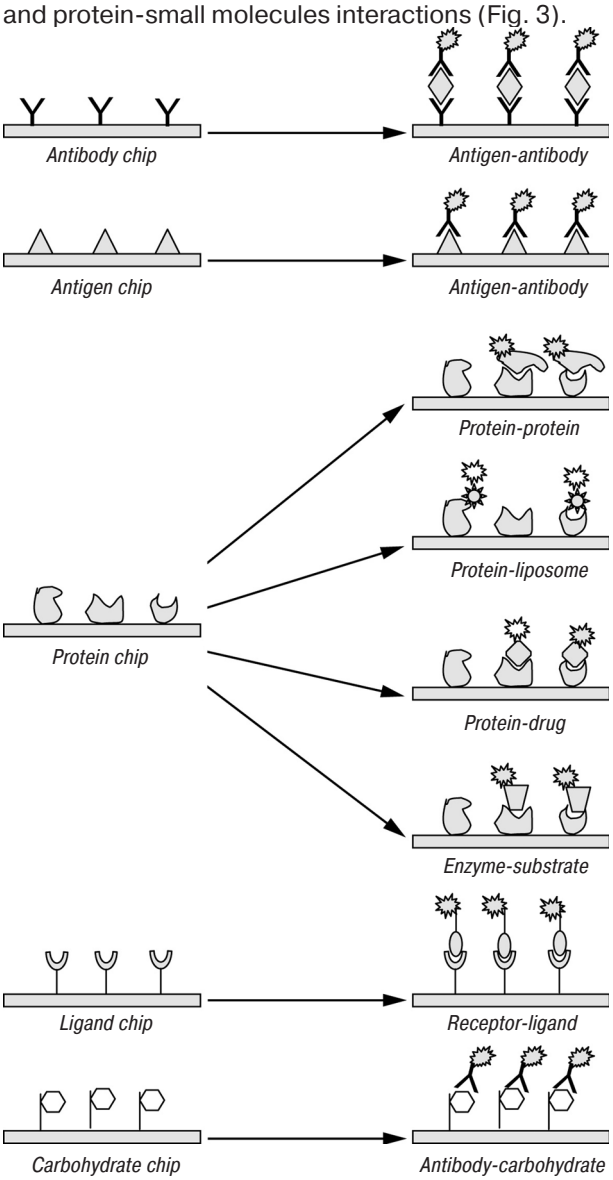

Fig. 3. Schematic representation of protein microarrays. Different molecules (antibodies, antigens, proteins, ligands, carbohydrates) are chemically linked to a chip by the use of various capturing agents, and then challenged with specific binders. Successful interactions are then detected by absorbance, luminescence or radioactive methods (star flashes in the figures)

Two main approaches are used: 1- quantitative-based and 2- functional-based microarrays. Currently there are two types of quantitative-based protein chips: 1a) capture microarrays, in which bait molecules, spotted on a support, are challenged by the sample [42]; 1b) reverse-phase protein blot (RPP), in which unknown experimental sample is loaded directly on the support, and then probed with analyte-specific reagents (*i. e.* antibodies [43]). Functional-based microarrays are focused on the study of biochemical activities and dynamic interactions of target proteins. Two different strategies are applied: 2a) protein spotting microarrays; a specific and unknown target protein is purified, spotted on the chip and challenged with different molecules/drugs, under various physiological conditions [44]; 2b) self-assembling protein microarrays; in order to avoid purification problems and to improve the purity of the target molecules, proteins are synthesized directly onto the chip surface using a cell-free transcription/translation system [45].

B) Microarray technologies — Two main issues must be considered when planning use of protein microarrays: the choice of the chip and the link of proteins on it. Ideally, microarrays supports must bind proteins with high density, and they must keep proteins in a physiological conformation, with active site(s) exposed and accessible to their cognate ligand(s). Two types of protein chip are available: two-dimensional support (polystyrene, PVDF or nitrocellulose), and three-dimensional (3D) matrixes, such as polyacrylamide gel packet or polyelectrolyte thin films [46]. 3D matrixes provide much higher protein binding capacity compared to 2D layer, and they create a hydrophilic environment which minimizes protein denaturation. Alternatively, direct attachment of proteins onto plain glass surface in the presence of specific glycerole-based buffer [47] has also been successfully tested. The second critical step in protein microarray is the anchor of the bait (in general proteins, see Fig. 3) to the support. The link through hydrophobic or ionic interactions (*i. e.* on a polylysine coat) is more direct but less stable, and it might be washed off under stringent washing conditions. An alternative is the use of cross-linkers. Bifunctional silane cross-linkers react with hydroxyl groups of glass surface and with primary amino groups of proteins; bifunctional thio-alkylene links bait molecules with SH-groups of gold-coated glass surface [48, 49]. However all these methods give a random orientation of the bait onto the support, thus generating an intra- and inter-matrix variability that impairs final result. Ideally, the molecule baits must be attached to the support at high density with the same orientation, exposing their active site(s). To achieve these goals, proteins might be fused at their amino- or carboxy-terminus with a high affinity tag, which guarantees at the same time a strong interaction with the support, a uniform orientation and a native conformation state [50].

C) Protein detection — After the preparation of the appropriate microarray, the sample, dissolved in a suitable buffer, is loaded onto the chip, and washing steps minimize unspecific interactions. Different options are possible for the detection and the identification of the bound proteins. Fluorescence is widely used for its simplicity, sensitivity and very high resolution; ELISA and radioisotope labeling are also employed. Since labeling molecules can affect three dimensional structure and/or functional activity, non-labeling methods such as surface-enhanced laser desorption/ionization (SELDI)/MS

have been used with low-density microarray [51]. For the kinetic studies of dynamic interactions the best choice is surface plasmon resonance (SPR [52, 53]).

2. Applications. Protein arrays are often used when a fast and reliable analysis of known molecules/proteins is required. This analysis includes measurements of protein expression, detection of PTMs and measurements of affinity of interactions.

As protein arrays are based on protein interactions, it is not surprising that most of the arrays employ antibodies. There are strong preclinical evidences that cancer undergoes immune surveillance. Although antibodies against many tumor antigens have been identified in the sera of cancer patients, very little is known about the specificity and clinical impact of anticancer auto-antibodies. Anderson and colleagues first build up a microarray system with over 1700 proteins, synthesized and assembled directly on the chip; these antigens carry a GST tag, for the expression control. Chips were then challenged with serum from patients, and bound antibodies detected [54]. The choice and set up of this system allows a very fast and sensitive screening of many hundreds of possible biomarkers, by the use of antibody-antigen interaction; the study was limited to pre-defined candidate targets, synthesized directly on a chip.

Protein phosphorylation is estimated to affect 30% of the proteome and is a major regulatory mechanism controlling many basic cellular processes. Although many studies focused on this subject, an overall picture is far to be achieved. In a large scale study on yeasts, protein microarray chips with 4400 different kinase substrates were prepared, and each chip challenged with 82 yeast kinases. The experiment identified 1325 kinase substrates, with over 4200 phosphorylation events [55]. Authors exploited protein microarray, as a fast and highly efficient system for identification of interactions among known, pre-defined pool of molecules.

3. Advantages and disadvantages. Protein microarrays allow investigating interaction of proteins with a wide spectrum of molecules, including lipids, carbohydrates, nucleic acids, synthetic drugs. Moreover, these interactions might be studied under different physiological conditions (pH, salts concentrations, etc.) and in a temporal range, following the dynamic evolution of the interactions (in the case of functional-based microarrays). The technique is fast, with a high identification rate (since it is based on pre-defined molecules), and it is particularly suitable for large screening of protein interactome, under different functional contexts. Protein microarrays are thus a very powerful tool, with a broad range of applications in biological, medical, pharmaceutical and chemical sciences.

An important limitation is that only a pool of known proteins can be analyzed. Although this pool can be extended up to few thousands different proteins, it will be always limited compared to the potential of other techniques to identify whatever protein involved, regardless *a priori* constrictions. In order to achieve a larger diffusion of the procedure, technical advances are needed. As antibodies are commonly used, directly or indirectly, in protein microarrays, the antibody specificity is a pivotal issue. Often available antibodies do not show a good specificity, thus lowering the number of proteins that could be analyzed. On microarray chip, proteins must be kept in the native conformation, they must have functional sites and/or docking site accessible to substrate(s) and they must be uniformly oriented and anchored to the support; all these requirements do not easily match with the high structural and functional plasticity of proteins (see Table 1).

In Table 3 internet links to web sites are provided with detailed information on described proteomics techniques.

Table 3. Links to web sites with detailed information on described proteomic techniques. The discussion forums allow all participants exchange questions and suggestions

Web site	Reference
abrf.org/	Proteomics
expasy.org/	Proteomics
weihenstephan.de/blm/deg/manual/manfrm.htm	2D gel Electrophoresis
expasy.ch/ch2d/protocols/protocols.fm.html	2D gel Electrophoresis
chm.bris.ac.uk/ms/theory/theory.html	MS
astbury.leeds.ac.uk/facil/MStut/mstutorial.htm	MS
forumsci.co.il/HPLC/lcms_page.html	LC / MS
microsolvtech.com/ce.asp	CE
robinsonlab.stanford.edu/microarrays/index.htm	Protein microarray
abrf.org/index.cfm/list.index	Proteomics – Discussion forum
scientistsolutions.com/	Proteomics – Discussion forum
amersham.zeroforum.com/zeroforum?id=1	2D gel Electrophoresis – Discussion forum

CONCLUSIONS

Proteome plays the central role in regulation of cell functions, and this boosts studies of protein expression, structure, organization and interactions. Proteome studies put also pressure on development of technologies suitable to solve specific problems. Although each proteomics approach has intrinsic limitations and advantages, the available techniques show a great potential in terms of sensitivity, accuracy and reproducibility. As proteome study consists of preparation of a sample, separation and identification of analytes, various technical solutions have been introduced. 2DE-MS, LC-MS, CE-MS and protein microarrays are the 4 most often used applications. Selection of the most suitable of these approaches may define successfulness of a study.

REFERENCES

1. **Kahn P.** From genome to proteome: looking at a cell's proteins. Science 1995; **270**: 369–70.

2. **Seo J, Lee KJ.** Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. J Biochem Mol Biol 2004; **37**: 35–44.

3. **Espina V, Geho D, Mehta AI, *et al.*** Pathology of the future: molecular profiling for targeted therapy. Cancer Invest 2005; **23**: 36–46.

4. **Gorg A, Weiss W, Dunn MJ.** Current two-dimensional electrophoresis technology for proteomics. Proteomics 2004; **4**: 3665–85.

5. **Pastwa E, Somiari SB, Czyz M, *et al.*** Proteomics in human cancer research. Proteomics Clin App 2007; **1**: 4–17.

6. **Hoving S, Gerrits B, Voshol H, *et al.*** Preparative two-dimensional gel electrophoresis at alkaline pH using narrow range immobilized pH gradients. Proteomics 2002; **2**: 127–34.

7. **Unlu M, Morgan ME, Minden JS.** Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis 1997; **18**: 2071–77.

8. **Palagi PM, Hernandez P, Walther D, *et al.*** Proteome informatics I: bioinformatics tools for processing experimental data. Proteomics 2006; **6**: 5435–44.

9. **Tiss A, Smith C, Camuzeaux S, *et al.*** Serum peptide profiling using MALDI mass spectrometry: avoiding the pitfalls of coated magnetic beads using well-established ZipTip technology. Proteomics 2007; **7:** 77–89.

10. **Aebersold R, Mann M.** Mass spectrometry-based proteomics. Nature 2003; **422**: 198–207.

11. **Steen H, Mann M.** The ABC's (and XYZ's) of peptide sequencing. Nat Rev Mol Cell Biol 2004; **5**: 699–711.

12. **Paizs B, Suhai S.** Fragmentation pathways of protonated peptides. Mass Spectrom Rev 2005; **24**: 508–48.

13. **Liu J, Cai Y, Wang J, *et al.*** Phosphoproteome profile of human liver Chang's cell based on 2-DE with fluorescence staining and MALDI-TOF/TOF-MS. Electrophoresis 2007; **28**: 4348–58.

14. **Zheng X, Hong L, Shi L, *et al.*** Proteomics analysis of host cells infected with infectious bursal disease virus. Mol Cell Proteomics 2008; **7**: 612–25.

15. **Conrotto P, Yakymovych I, Yakymovych M, *et al.*** Interactome of transforming growth factor-beta type I receptor (TbetaRI): inhibition of TGFbeta signaling by Epac1. **J** Proteome Res 2007; **6**: 287–97.

16. **Leitner A, Lindner W.** Chemistry meets proteomics: the use of chemical tagging reactions for MS-based proteomics. Proteomics 2006; **6**: 5418–34.

17. **Conrotto P, Hellman U.** Lys Tag: an easy and robust chemical modification for improved de novo sequencing with a matrix-assisted laser desorption ionization tandem time-of-flight mass spectrometer. Rapid Comm. Mass Spec 2008; **22**: 1823–33.

18. **Naidong W.** Bioanalytical liquid chromatography tandem mass spectrometry methods on underivatized silica columns with aqueous/organic mobile phases. J Chromatogr B Analyt Technol Biomed Life Sci 2003; **796**: 209–24.

19. **Gu X, Deng C, Yan G, *et al.*** Capillary array reversed-phase liquid chromatography-based multidimensional separation system coupled with MALDI-TOF-TOF-MS detection for high-throughput proteome analysis. J Proteome Res 2006; **5**: 3186–96.

20. **Ishihama Y.** Proteomic LC-MS systems using nanoscale liquid chromatography with tandem mass spectrometry. J Chromatogr A 2005; **1067**: 73–83.

21. **Zimmer JS, Monroe ME, Qian WJ, *et al.*** Advances in proteomics data analysis and display using an accurate mass and time tag approach. Mass Spectrom Rev 2006; **25**: 450–82.

22. **Mueller DR, Voshol H, Waldt A, *et al.*** LC-MALDI MS and MS/MS-an efficient tool in proteome analysis. Subcell Biochem 2007; **43**: 355–80.

23. **Fröhlich T, Arnold GJ.** Proteome research based on modern liquid chromatography-tandem mass spectrometry: separation, identification and quantification. J Neural Transm 2006; **113**: 973–94.

24. **Stevens SM Jr, Duncan RS, Koulen P, *et al.*** Proteomic analysis of mouse brain microsomes: identification and bioinformatic characterization of endoplasmic reticulum proteins in the Mammalian central nervous system. J Proteome Res 2008; **7**: 1046–54.

25. **Soltermann A, Ossola R, Kilgus-Hawelski S, *et al.*** N-glycoprotein profiling of lung adenocarcinoma pleural effusions by shotgun proteomics. Cancer 2008; **114**: 124–33.

26. **Cho H, Smalley DM, Theodorescu D, *et al.*** Statistical identification of differentially labeled peptides from liquid chromatography tandem mass spectrometry. Proteomics 2007; **7**: 3681–92.

27. **Wiese S, Reidegeld KA, Meyer HE, *et al.*** Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. Proteomics 2007; **7**: 340–50.

28. **Ahmed N, Thornalley PJ.** Quantitative screening of protein biomarkers of early glycation, advanced glycation, oxidation and nitrosation in cellular and extracellular proteins by tandem mass spectrometry multiple reaction monitoring. Biochem Soc Trans 2003; **31**: 1417–22.

29. **Schmitt-Kopplin P, Frommberger M.** Capillary electrophoresis-mass spectrometry: 15 years of developments and applications. Electrophoresis 2003; **24**: 3837–67.

30. **Schmitt-Kopplin P, Englmann M.** Capillary electrophoresis — mass spectrometry: survey on developments and applications 2003–2004. Electrophoresis 2005; **26**: 1209–20.

31. **Stutz H.** Advances in the analysis of proteins and peptides by capillary electrophoresis with matrix-assisted laser desorption/ionization and electrospray-mass spectrometry detection. Electrophoresis 2005; **26**: 1254–90.

32. **Priego Capote F, Luque de Castro MD.** On-line preparation of microsamples prior to CE. Electrophoresis 2007; **28**: 214–20.

33. **Simpson DC, Smith RD.** Combining capillary electrophoresis with mass spectrometry for applications in proteomics. Electrophoresis 2005; **26**: 1291–305.

34. **Hernández-Borges J, Neusüss C, Cifuentes A, *et al.*** On-line capillary electrophoresis-mass spectrometry for the analysis of biomolecules. Electrophoresis 2004; **25**: 2257–81.

35. **Bushey MM, Jorgenson JW.** Automated instrumentation for comprehensive two-dimensional high-performance liquid chromatography of proteins. Anal Chem 1990; **62**: 161–7.

36. **Janini GM, Chan KC, Conrads TP, *et al.*** Two-dimensional liquid chromatography-capillary zone electrophoresis-sheathless electrospray ionization-mass spectrometry: evaluation for peptide analysis and protein identification. Electrophoresis 2004; **25**: 1973–80.

37. **Hartinger CG, Keppler BK.** CE in anticancer metallodrug research-an update. Electrophoresis 2007; **28**: 3436–46.

38. **Erny GL, Marina ML, Cifuentes A.** CE-MS of zein proteins from conventional and transgenic maize. **Electro**phoresis 2007; **28**: 4192–201.

39. **Juan-García A, Font G, Picó Y.** Simultaneous determination of different classes of antibiotics in fish and livestock by CE-MS. Electrophoresis 2007; **28**: 4180–91.

40. **Desai MJ, Armstrong DW.** Separation, identification, and characterization of microorganisms by capillary electrophoresis. Microbiol Mol Biol Rev 2003; **67**: 38–51.

41. **Rodriguez MA, Armstrong DW.** Separation and analysis of colloidal/nano-particles including microorganisms by capillary electrophoresis: a fundamental review. J Chromatogr B Analyt Technol Biomed Life Sci 2004; **800**: 7–25.

42. **Bradner JE, McPherson OM, Koehler AN.** A method for the covalent capture and screening of diverse small molecules in a microarray format. Nat Protoc 2006; **1**: 2344–52.

43. **Speer R, Wulfkuhle J, Espina V, *et al.*** Development of reverse phase protein microarrays for clinical applications and patient-tailored therapy. Cancer Genomics Proteomics 2007; **4**: 157–64.

44. **LaBaer J, Ramachandran N.** Protein microarrays as tools for functional proteomics. Curr Opin Chem Biol 2005; **9**: 14–9.

45. **Ramachandran N, Hainsworth E, Bhullar B, *et al.*** Self-assembling protein microarrays. Science 2004; **305**: 86–90.

46. **Zhou X, Zhou J.** Antibody-microarrays on hybrid polymeric thin film-coated slides for multiple-protein immunoassays. Methods Mol Biol 2007; **382**: 259–71.

47. **Liu Y, Li CM, Yu L, *et al.*** Optimization of printing buffer for protein microarrays based on aldehyde-modified glass slides. Front Biosci 2007; **12**: 3768–73.

48. **Seurynck-Servoss SL, White AM, Baird CL, *et al.*** Evaluation of surface chemistries for antibody microarrays. Anal Biochem 2007; **371**: 105–15.

49. **Kim DH, Shin DS, Lee YS.** Spot arrays on modified glass surfaces for efficient SPOT synthesis and on-chip bioassay of peptides. J Pept Sci 2007; **13**: 625–33.

50. **Tomizaki KY, Usui K, Mihara H.** Protein-detecting microarrays: current accomplishments and requirements. Chembiochem 2005; **6**: 782–99.

51. **Espina V, Dettloff KA, Cowherd S, *et al.*** Use of proteomic analysis to monitor responses to biological therapies. Expert Opin Biol Ther 2004; **4**: 83–93.

52. **Yu X, Xu D, Cheng Q.** Label-free detection methods for protein microarrays. Proteomics 2006; **6**: 5493–503.

53. **Nedelkov D.** Mass spectrometry-based immunoassays for the next phase of clinical applications. Expert Rev Proteomics 2006; **3**: 631–40.

54. **Anderson KS, Ramachandran N, Wong J, *et al.*** Application of Protein Microarrays for Multiplexed Detection of Antibodies to Tumor Antigens in Breast Cancer. J Proteome Res 2008; **7**: 1490–9.

55. **Ptacek J, Devgan G, Michaud G, *et al.*** Global analysis of protein phosphorylation in yeast. Nature 2005; **438**: 679–84.

МЕТОДЫ ПРОТЕОМИКИ В БИОЛОГИИ И МЕДИЦИНЕ: ПРИНЦИПЫ И ПРИМЕНЕНИЕ

Развитие протеомики за последние 10 лет стимулировало возрастающий интерес к ее применению для решения важных проблем современной биологии и медицины. Сегодня протеомика активно используется в фундаментальных исследованиях, поиске биомаркеров, диагностике опухолей. Широкомасштабное изучение экспрессии, локализации, активности и взаимодействий белков привело к специализации протеомных технологий и методов. Поэтому вопрос подбора наиболее подходящих методов для решения специфических проблем является особо актуальным. Двухмерный электрофорез, жидкостная хроматография, капиллярный электрофорез и масс-спектрометрия являются наиболее развитыми технологиями протеомики. В этом обзоре проанализированы сильные и слабые стороны этих технологий. Авторы надеются, что этот анализ поможет читателям выбрать наиболее эффективную методику для решения соответствующих научных задач.

Ключевые слова: **протеомика, масс-спектрометрия, хроматография, двухмерный электрофорез, белковые чипы, капиллярный электрофорез.**