

# Mass spectrometry approaches in proteomic and metabolomic studies

PAWEŁ RODZIEWICZ<sup>†\*</sup>, BARBARA SWARCEWICZ<sup>†</sup>, KLAUDIA CHMIELEWSKA<sup>†</sup>

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland

<sup>†</sup>These authors contributed equally to this work and should be considered co-first authors

\* Corresponding author: prod@ibch.poznan.pl

## Abstract

With the development of the systems biology concept proteomic and metabolomic studies have become even more attractive. The advancement in separation methods of proteins and metabolites, and particularly the progress that has been made in the field of mass spectrometry significantly facilitated high-throughput analyses and substantially increased both quality and quantity of the data. In this short review we discuss some aspects of the analytical strategies used in proteome and metabolome research in which mass spectrometry plays a crucial role.

## Introduction

“Omics” studies like genomics, transcriptomics, proteomics, metabolomics, lipidomics, etc. comprise the systems biology. Proteomics and metabolomics are analytical approaches focused on qualitative and/or quantitative overview of proteins and metabolites present in the organism, tissue, cell, organelle, sap in case of plants or body fluids in case of animals. The dynamics and variability of the proteome and metabolome are the results of changes in gene activities caused by various factors like growth stage/age, diseases and other environmental conditions. As a consequence researchers encounter a lot of hurdles in analyzing such complex levels of molecular organization. Growing interest in those fields of research during recent decades led to development of novel, sensitive and high-throughput analytical methods in proteome and metabolome science. Most of these strategies are based on mass spectrometry coupled to various separation techniques.

## Proteomics

The term “proteome” was coined by Marc Wilkins in 1995 to describe all the proteins encoded by the genome present in the cell, tissue or organism in specific external and internal conditions, at a specific time point and developmental stage. By analogy to the term “genomics” “proteomics” was introduced (Wilkins et al., 1996). Pro-

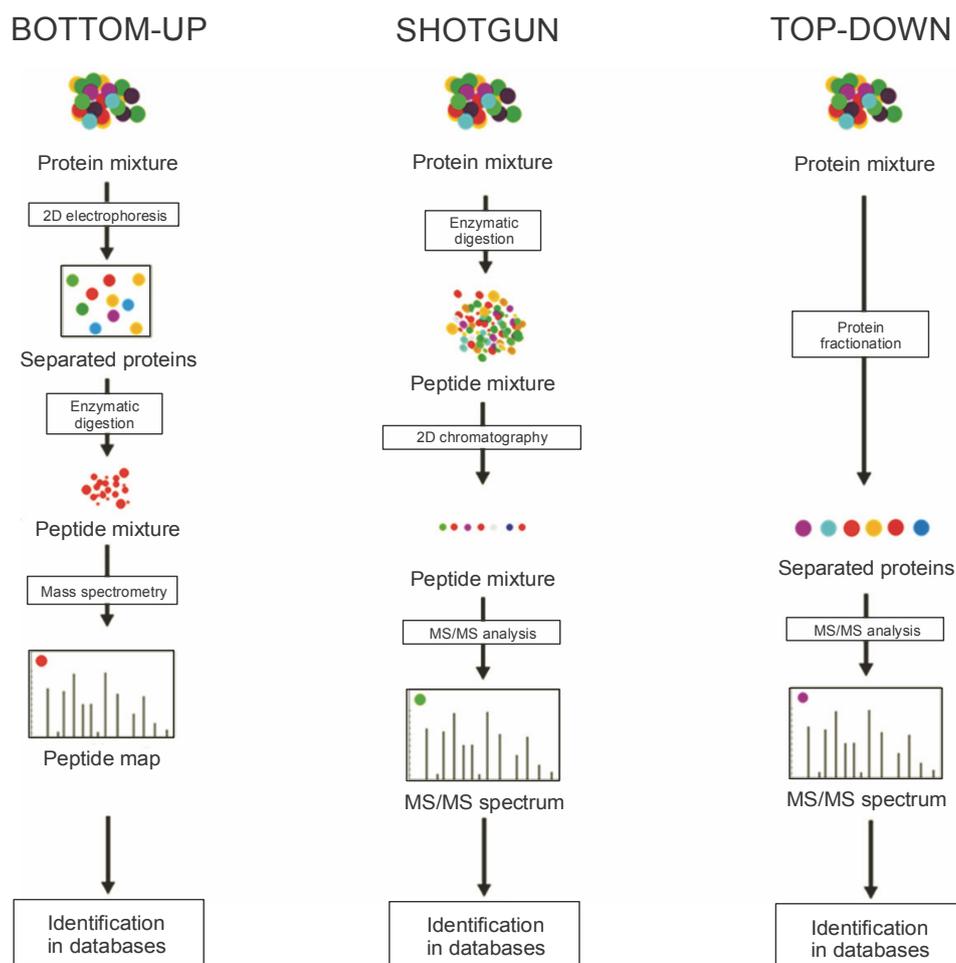
teomics primary goal is to rapidly identify all the proteins found in a cell, tissue or organism and determine possible interaction occurring between them and between proteins and other molecules. It also allows to follow the activity of proteins in various stages of growth and development and uncover role they play in specific environmental conditions. Proteins are assembled from 20 different amino acids linked by peptide bond. Those biopolymers are very heterogeneous in terms of weight, sequence, and physico-chemical properties. The number of genes encoding proteins range from a few hundred in bacteria to tens of thousands in mammals, whereas the number of proteins can be much higher (Thelen, 2007). It is estimated that the number of proteins encoded in the human genome may be even 100-fold greater than the number of genes due to alternative mRNA splicing and post-translational modifications (PTMs) (Neverova and Van Eyk, 2005). With the growing number of sequenced genomes and development of protein separation and identification methods proteomics has become an essential tool in molecular and systems biology (Abersold and Mann, 2003). The oldest approach in modern proteomics includes two-dimensional electrophoretic separations and identification by mass spectrometry techniques, usually MALDI-TOF (Matrix Assisted Laser Desorption/Ionization Time of Flight) mass spectrometer or tandem mass spectrometer ESI-MS/MS (Electrospray Tan-

dem Mass Spectrometry) are utilized (Görg et al., 2004). The introduction of ultra performance liquid chromatography (UPLC) and nano liquid chromatography (nLC) systems coupled to tandem mass spectrometer allowed to overcome the limitations of two-dimensional electrophoresis (2D-PAGE). The development of labeling techniques as well as label-free methods permitted accurate quantitative analysis (Patel et al., 2009). Nowadays proteomic studies are mainly realized by three strategies: the classical “bottom-up”, “shotgun proteomics” and “top-down” approach (Fig. 1).

### Bottom-up proteomics

The bottom-up proteomics is based on the analysis of peptides derived from enzymatic/non-enzymatic protein digestion in mass spectrometer. The obtained spectrum, called peptide mass fingerprint (PMF) is used for protein identification in databases. In this approach two dimensional polyacrylamide gel electrophoresis is used for protein separation (Han et al., 2008; Hirano et al., 2004; Kelleher et al., 1999). Resolved proteins are excised from gel, purified and subjected to disulfide bonds reduction and alkylation process in order to prevent their reformation. Prepared samples are submitted to enzymatic digestion and further analyzed (Fig. 1). 2D-PAGE consists of two independent separations. In the first dimension (IEF, isoelectrofocusing) proteins migrate in pH gradient formed by ampholytes, in the presence of urea and detergents, according to their isoelectric point (pI). In the second step proteins are separated by molecular weight in denaturing polyacrylamide gel containing SDS (Scheler et al., 1998; Smith, 2009; Thelen, 2007). Proteins can be labeled in various points of the analysis: before isoelectrofocusing by radioisotopes or cyanide dyes, which do not alter the protein pI; between first and second dimension; but usually proteins are visualized after both separations. One of the most commonly used and simplest method of protein detection is colloidal solution of Coomassie Brilliant Blue, which however has a moderate detection sensitivity (~100 ng) and requires long time procedure (Neuhoff et al., 1985). Silver staining has higher sensitivity (~1 ng), but narrow dynamic range and low repeatability (Chevallet et al., 2006). Alternatively, fluorescent dyes can be used, as they provide high sensitivity (~2-10 ng) and wide dynamic range. However, they are expensive and require additional hardware to visualize the sepa-

rated proteins (Steinberg et al., 1996). The resolution of 2D-PAGE largely depends on the pH range and size of the used strips in first dimension, and usually allows to resolve from 1000 to 5000 proteins. This technique also enables detection of protein species, which are the result of PTMs, alternative splicing or genetic variability (Görg et al., 2004; Newton et al., 2004; Sarma et al., 2008). The most commonly used protease is trypsin, which cleaves specifically peptide bonds at the C-terminal side of lysine or arginine, except when the next amino acid in the sequence is proline, then the hydrolysis does not occur. Chymotrypsin, which cleaves peptide bonds after hydrophobic amino acids and endoproteases like Lys-C or Arg-C are also utilized. It is also possible to use cyanogen bromide, which cleaves peptide bonds at the carboxyl side of methionine (Matthiesen and Mutenda, 2007). The resulting peptide mixture is analyzed by mass spectrometry. Most commonly used instrument in this approach is MALDI-TOF mass spectrometer (Henzel et al., 2003; Hirano et al., 2004; Yates et al., 2009). Peptide mass fingerprinting is based on the concept that each protein after proteolytic digestion will generate a unique set of peptides, which masses will provide a molecular signature enabling protein identification (Hjernø, 2007; Thiede et al., 2005). The list of peptide masses from generated spectrum is used for comparison with theoretical masses of peptides obtained *in silico* from proteins sequences present in databases (Lin et al., 2003; Sommerer et al., 2007). In the algorithms implemented in database search engines like MASCOT (<http://fields.scripps.edu/sequest>) or SEQUEST (<http://www.matrixscience.com>) various parameters are taken into account, including type and size of the database, taxonomy, used protease and number of miscleavages, modifications made in amino acid sequence, error tolerance of mass values, etc. (Sadygov et al., 2004). The advantages of PMF are fast data acquisition, ease of sample preparation, small amount of sample needed for analysis and ability to store sample after measurement (Cramer, 2009). This approach has also some limitations. Some of the peptides may be poorly ionized thus will be invisible in the spectrum. In case of additional signals on spectrum, effects of modified peptides, impurities or mixture of two and more proteins in one sample, accidental proteins may be identified or given data will result in no identification at all (Hamdan and Righetti, 2005; Resing and Ahn, 2005). This method is particularly successful



**Fig. 1.** Strategies for protein analysis (adapted from Sikorska and Rodziewicz, 2011)

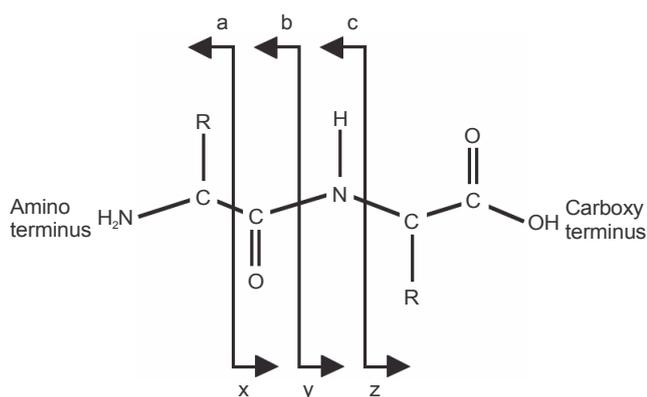
for identifying proteins from species with sequenced genomes. In other case identification is based on homology to related species (Hjerno, 2007; Knochenmuss, 2012; Sommerer et al., 2007). *De novo* protein sequencing may significantly improve the ratio of positive identifications. Such solution can be achieved by MALDI-TOF/TOF spectrometers equipped with collision cell. Peptide fragmentation by collision-induced dissociation (CID) provides capability to generate partial amino acid sequence. Combination of the PMF data with MS/MS spectra gives the best results and minimalizes the probability of random matches (Liu and Schey, 2005; Mathiesen and Mutenda, 2007). The limitations of the bottom-up approach are mainly related to 2D-PAGE and include difficulties in resolving small and large proteins with hydrophobic properties (e.g. membrane proteins). The most popular protein detection methods have low dynamic range, which hampers detection of low-abundant proteins and subtle differences in the protein ac-

cumulation between different cellular states. The weakness of the 2DE-PAGE is also related to the differences in the distribution of individual proteins (shifts in their position between the gels), which complicates subsequent statistical analysis. In the case of low-intensity proteins MS identification may be ineffective and ambiguity may occur due to presence of several proteins in the form of one spot. 2D-PAGE possess also many advantages, which include high resolution and information about protein mass and pI. Moreover, selective staining enables identification of PTMs like phosphorylation and glycosylation (Berth et al., 2007). Both steps of 2D-PAGE are separated in time and space, and the equipment is relatively inexpensive. This technique can also be combined with other used in molecular biology, e.g. electrotransfer and western-blotting. Despite its limitations the bottom-up approach is still widely used in proteomic research (Rabilloud et al., 2010).

## Shotgun proteomics

Shotgun proteomics is an emerging tool for analysing complex mixtures of proteins in samples prepared from tissues, cells, organelles or biofluids. The term “shotgun proteomics” was coined after shotgun sequencing of genomic DNA, where long sequences are computationally recovered from short fragment reads. In shotgun proteomics, proteins are identified from tandem mass spectra (MS/MS) of their proteolytic peptides (Yates, 1998). In basic, a typical design of shotgun proteomics experiment involves sample preparation, multi-dimensional chromatography, mass spectrometry analysis and protein identification in databases (Fig. 1). Purification and reduction of the sample complexity is advantageous and can be obtained by the application of chromatographic or electrophoretic fractionation, tandem affinity purification (Rigaut et al., 1999), and by chemical labelling like Isotope-Coded Affinity Tags (ICAT) (Gygi et al., 1999). The next step is the enzymatic or chemical digestion of the prepared protein mixture. The most utilized protease is trypsin because of its stability, specificity, low cost and it also produces short peptides appropriate for MS analysis. Other enzymes like chymotrypsin, Asn-N, Glu-C Lys-C or proteinase K and chemical agents like cyanogen bromide can also be used. The proteolysis of the sample increases its complexity and the resulted peptides require separation in more than one dimension prior MS/MS analysis to assure appropriate sensitivity and resolution of the peptides. Multi-dimensional Protein Identification Technology (MudPIT) is a chromatographic approach where two or more stationary phases are applied and usually different chemical or physical properties of the molecules are exploited to fractionate them (Giddings, 1987). A combination of two chromatographic techniques is commonly employed, including size exclusion (SE; separation of molecules is based of their size), anion exchange (AE; separation of the molecules is based on their charges using positively charged column), strong cation exchange (SCX; separation of the molecules is based on their charges using negatively charged column) and reversed phase (RP; separation of molecules is based on hydrophobic interactions with C4, C8, C18 alkyl chains of the stationary phase) (Swanson and Washburn, 2005). SCX as a first dimension and RP column as a second one is the most commonly implemented combination in shotgun ap-

proach. In general, three configuration of SCX/RP separation are possible. In off-line analysis fractions from SCX are first collected and then subjected to RP column. In on-line analysis SCX fractions are eluted to RP trap column to wash out the salts used as a mobile phase in first dimension and then to RP column. The third strategy takes advantage of biphasic columns packed with both chromatographic beds. This approach requires volatile salts like ammonium acetate or formate to be used as a mobile phase (Wolters et al., 2001). Separated peptides are then analysed in mass spectrometer. In shotgun proteomics soft ionization methods ESI and MALDI are applied. Shotgun analysis requires tandem mass spectrometer to be used in order to fragment analysed ions to enable peptide sequencing. Mostly utilized analysers are: quadrupole ion traps, linear ion traps, time-of-flight/time-of-flight and quadrupole-time-of-flight (Yates, 2004). Ionized peptide is isolated in mass spectrometer and excited by the CID with inert gas particles which results in fragmentation of the selected molecule. The development of electron capture dissociation (ECD) and electron transfer dissociation (ETD) enabled efficient fragmentation of larger peptides and facilitated the analysis of PTMs (Creese and Cooper, 2007; Syka et al., 2004). Fragmentation occurs by the breakage of the backbone of the two adjoining amino acids which generates fragments of different length. The sum of all obtained fragments will give MS/MS spectrum from which the sequence of the peptide can be read. In practice it is a complicated task because the breakage can occur between various amino acids and carboxyl sides of peptide leading to generation of *abc* ions if the charge is at amino terminus or *xyz* ions if the charge remains at the carboxyl terminus and with different efficiency – Figure 2 (Roepstorff and Fohlman, 1984). The obtained data is interpreted using various algorithms which involve comparison of experimental MS/MS data with a set of theoretically generated fragments for peptides present in the database. Most popular search engines are SEQUEST (Eng et al., 1994) and Mascot (Perkins et al., 1999). Typical shotgun analysis generates from 10 000 to 100 000 tandem spectra. Vast amount of generated MS/MS data forced development of statistical methods for validating protein identification (Nesvizhskii et al., 2003; Sadygov and Yates, 2003). To gain the full potential of shotgun proteomics further development of



**Fig. 2.** Schematic diagram of ion nomenclature (adapted from Roepstorff and Fohlman, 1984)

search algorithms and statistical tools is necessary. One of the most promising application of shotgun proteomics is large scale analysis of PTMs, which with some exceptions like glycosylation, are hard to be observed using bottom-up approach. Moreover, PTMs change masses of peptide fragments which additionally hamper the identification. For example, phosphorylation very often occur in hydrophilic regions of low-abundant proteins and exists at substoichiometric levels. The recent technological breakthroughs facilitated for example phosphorylation site mapping in various organisms (Nakagami et al., 2012). In large scale analysis of phosphoproteome from whole-cell lysate of *Saccharomyces cerevisiae*, analytical strategy, which was applied, enabled detection of more than 1000 phosphopeptides and determination of 383 sites of phosphorylation (Ficarro et al., 2002). In other study, in which for the first time ETD fragmentation was used to analyse plant material, more than 3000 phosphopeptides in *Medicago truncatula* roots was identified (Grimsrud et al., 2010). Another field of application of shotgun proteomics is quantitative analysis of the differences in protein expression pattern between samples after treatment or exposition to various stimuli (e.g. healthy vs diseased patients). To quantify such changes, proteins in one sample (representing one cellular state) have to be modified in order to be distinguished from the second sample (representing second cellular state). After labelling, samples are mixed together and analysed. Labelling can be achieved by various methods. For example, if the analysis concerns cells cultured *in vitro*, they can be grown on medium containing isotopic variants of light and heavy amino acids (SILAC, Stable Isotope Labelling by Amino Acids in Cell

Culture). Another approach involves chemical labelling after protein isolation. The most popular are ICAT (Gygi et al., 1999) and Isobaric Tag for Relative and Absolute Quantitation (iTRAQ) (Aggarwal et al., 2006; Ross et al., 2004). The ICAT is restricted only to proteins which contain cysteine residues in their sequence. Recently developed iTRAQ technique is free of that restriction and enables quantification even up to 8 different samples in single analysis. Quantitative shotgun proteomics can be also conducted by label-free methods. This strategy is realized by measuring changes in chromatographic peak areas or heights or by MS/MS spectral counting of the analysed peptides (Zhu et al., 2010).

### Top-down proteomics

Top-down strategy refers to the fragmentation of intact proteins in mass spectrometer without prior enzymatic/non-enzymatic digestion. Multiply charged proteins are used as precursors in tandem mass analysis (Fig. 1). Mass spectra of the proteins and their fragments allow to obtain the amino acid sequence and identify for example sites and type of PTMs. The very first structural analyses of intact proteins were conducted in 1990's on ribonuclease A (Loo et al., 1990) and serum albumin (Loo et al., 1991). Since then this approach started to gain more and more attention from scientific community. The breakthrough in top-down strategy was achieved by application of ECD, which outperformed the CID fragmentation in number of amino acid backbone cleavages, which resulted in better sequence coverage (Zubarev et al., 1998). In top-down proteomics ESI coupled to Fourier transformation ion cyclotron resonance mass spectrometer (FTICR MS) and hybrid ion-trap-orbitrap mass spectrometer are commonly used. Both of them are capable of performing measurements characterized by high accuracy, sensitivity and resolution, which enable determination of elemental composition of proteins up to 70 kDa. MALDI-TOF/TOF was also implemented in intact protein analysis and is mainly used in the field of tissue imaging. However its application in high-throughput proteomics is limited due to lower resolution and incomplete fragmentation of proteins (Seeley and Caprioli, 2011). One of the highest protein mass for which isotopic resolution was achieved was registered for 158 kDa tetrameric aldolase complex with resolving power of 520 000 at 6033 m/z for the 26-

times charged ion using ESI coupled to FTICR MS. This strategy also enabled 35% C-terminal amino acids to be sequenced (168 out of 463) (Li et al., 2014). High resolution mass spectra are crucial for characterization of various modifications; e.g. disulfide bonds ( $\Delta m = 2$  Da), deamination ( $\Delta m = 1$  Da), and distinction between different modification with similar molecular weights, e.g. trimethylation *vs* acetylation ( $\Delta m = 36$  mDa). Prior the top-down analysis complex protein mixtures have to be fractionated. Separation can be performed by liquid chromatography, hydrophobic interaction chromatography, ion exchange chromatography, electrophoresis, and isoelectrofocusing. However, to perform efficient analyses on proteome scale multidimensional separation methods are required (Tran et al., 2011). Fractions eluting from chromatograph can be either directly introduced to mass spectrometer (on-line) or collected and subjected for direct infusion analysis (off-line) (Catherman et al., 2014). In the past the top-down strategy was mainly implemented for targeted studies of single proteins. Nowadays, thanks to development of faster ion traps, automated fragmentation and efficient separation methods, this approach has been extended to whole proteome analysis (Macek et al., 2006; Tipton et al., 2012). For example, Smith and coworkers reported detection of 700 bacterial proteins (5-40 kDa) in single experiment. In other study 1034 gene products from human cells which accounted for more than 3000 of protein species up to 105 kDa were identified (Tran et al., 2011). The top-down strategy became extremely useful in identification of PTMs and proteins variants resulted from genetic variation and alternative splicing (Boyne et al., 2006; Medzihradzsky et al., 2004). In recent years substantial progress has been made in membrane proteins analysis using top-down approach, far exceeding the capability of bottom-up strategy (Armirotti and Damonte, 2010; Siuti and Kelleher, 2007). A big hurdle in this approach is difficult data interpretation due to complexity of generated spectra (Smith et al., 2013). However with the new software being developed the bioinformatic analysis of the top-down generated data will soon match the “shotgun proteomics” computations (Armirotti and Damonte, 2010).

## Metabolomics

Metabolomics studies due to the wide diversity of metabolites, both in terms of their number and chemical

nature, set analytical chemistry very ambitious task. It has been estimated that there are over 200 000 different metabolites within the plant kingdom, whilst the metabolome size of individual plant species range from 7000 to 15 000 compounds (Fernie et al., 2011; Weckwerth, 2003). Most bacteria produce between a few hundred to 2000 metabolites (Liebeke et al., 2012). Furthermore animals are believed to contain a few thousand small molecules they make themselves as well as other compounds coming from the environment (nutrients, pollutants, products of intestinal bacteria metabolism) (Baker, 2011).

The initial goal of metabolomic studies was to catalog all of the metabolites present in a given organism, but it quickly became apparent that this task would be extremely difficult. Due to the chemical complexity and heterogeneity of metabolites, there is no single analytical method that enables comprehensive metabolome analysis, therefore multiple technologies must be employed. However, the employment of even many analytical platforms still does not allow measuring the complete metabolite set of any organism. Additionally all of the techniques have advantages and limitations, therefore careful research approach must be chosen. Over the years several experimental strategies have been developed, which might be useful for answering the questions about the influence of environmental conditions on the composition and level of cellular metabolites, as well as impact of low-molecular-weight compounds on the cell.

## Approaches in metabolomics

Generally, the approaches applied in metabolome studies can be categorized into three classes: untargeted analysis, semi-targeted analysis and targeted analysis. These strategies differ in the number of detected metabolites, the level of metabolite identification and their quantitation (absolute or relative), as well as experimental precision.

In order to monitor low-molecular-weight compounds without the need for identification and quantification *metabolic fingerprinting* is applied. All steps, including: sample preparation, extraction procedures and all the other analytical methods used, should preferably be simple and high-throughput. To reduce time of the analysis chromatographic separation is usually neglected. Metabolic fingerprinting enables researchers to classify

samples according to their origin or biological relevance (e.g. control vs treatment) by searching chemical patterns, specific to an individual group. This approach is often used as the initial stage of metabolite profiling because it allows rapid sample comparison and discrimination analysis. Typically that kind of “snapshot” of the metabolic state of the cells is done by nuclear magnetic resonance (NMR), mass spectrometry (MS), Fourier transform infrared spectroscopy (FT-IR) or Raman spectroscopy (Goodacre et al., 2004). Since the metabolome refers to both intracellular and extracellular set of metabolites, to distinguish analysis of the endo- and exo-metabolome, screening approach of compounds present in the body fluids, or secreted to the external environment (e.g. culture medium), is called *metabolic footprinting* (Nielsen and Oliver, 2005).

*Metabolite profiling* restricts metabolome studies to a selected number of pre-defined chemical compounds, for example metabolites related to specific pathways or classes (e.g. carbohydrates, amino acids, alkaloids). Although this analysis does not need to be quantitative, typically it is at least semi-quantitative. Usually metabolite profiling employs hyphenated techniques which combine chromatographic methods for separation with spectral methods for detection. Following platforms are commonly used: one-dimensional or two-dimensional gas chromatography coupled to mass spectrometry (GC-MS, GC×GC-MS) (Mohler et al., 2006; Welthagen et al., 2005), liquid chromatography coupled to mass spectrometry (LC-MS, LC-MS<sup>n</sup>) (Lewis et al., 2010; Staszko et al., 2011), liquid chromatography with nuclear magnetic resonance spectroscopy (LC-NMR) (Agnolet et al., 2010) and capillary electrophoresis-mass spectrometry (CE-MS) (Ramautar et al., 2009). New developments in MS-based techniques enable qualitative and quantitative analysis of a relatively large number of metabolites.

Another strategy in metabolome studies – *metabolite targeted analysis* – focuses on the analysis of single or small group of compounds that participate in a specific part of the metabolism (e.g. substrates and products involved in specific enzymatic reaction). Such an approach provides detailed information, both qualitative and quantitative (absolute values), about selected compounds. Typically, targeted analysis refers to more traditional methods of chemical analysis and involves optimized procedures for the extraction, separation and detection of chosen compounds (Roberts et al., 2012).

*Metabolic flux analysis* focuses on elucidating the structure of metabolic pathways, cross-talk between metabolic networks and studying of mass flow in metabolic systems (Crown and Antoniewicz, 2013). An essential step of this approach is the introduction into biological systems (*in vitro* or *in vivo*) compounds labeled with stable-isotopes (<sup>13</sup>C, <sup>2</sup>H, <sup>15</sup>N, <sup>18</sup>O) or radioisotopes (<sup>14</sup>C, <sup>3</sup>H). After incubation with labeled compounds, metabolites are isolated from the biological sample, and then compounds with incorporated tracers are analyzed. This type of analysis is mostly based on NMR (Szyperki, 1995) and MS instrumentation (Hofmann et al., 2008; Maier et al., 2008).

To get some information about the spatial distribution of metabolites within tissue, which is not provided by coupled techniques like LC-MS or GC-MS, *MS imaging* (MSI) technologies are used. Imaging of small molecules within tissue sections is an important tool in various areas of research and includes studies on toxicology and drug metabolism (Parson et al., 2012), lipidomics (Eberlin et al., 2010) and biomarker discovery (Ruh et al., 2013). The most widely used approach for imaging is MALDI. In this procedure chemical compound called matrix is applied by spotting or spraying to the thin tissue sections. Then laser is fired at the sample, which leads to ionization of the molecules and mass spectra generation. Scanning laser over the sample results in many spectra, used to create image which visualizes distribution and abundance of ions (Cornett et al., 2007). Other MS-based methods used for imaging include: desorption electrospray ionization (DESI) (Takats, 2004), liquid extraction surface analysis (LESA) (Eikel et al., 2011), nanostructure initiator mass spectrometry (NIMS) (Yanes et al., 2009) and secondary ion mass spectrometry (SIMS) (Winograd, 2005).

## Conclusions

Proteomics and metabolomics together with other “omics” approaches such as genomics and transcriptomics are a part of systems biology concept. The aim of systems biology is to integrate existing knowledge about biological components of the organism, and thus generate a holistic view of the organism, with all the interactions occurring between individual elements. The ultimate goal is to formulate universal principles explaining the structure, organization and functioning of a living organism (Coruzzi et al., 2009). Substantial pro-

gress towards mass spectrometry instruments and analytical software which has been made in recent years, improved the quality of the acquired data in proteomic and metabolomic studies. The development of ionization techniques and analyzers in MS enabled faster ion scanning, improved sensitivity, selectivity and more efficient fragmentation methods. Together with advancement in separation strategies (e.g. multidimensional chromatography) more information about biological samples is now accessible. New approaches in bioinformatics and increase in number and quality of databases facilitate identification of proteins and metabolites as well as their correlation. Further integration of proteomic and metabolomic data with results generated from other molecular levels and phenotype observations will contribute to explanation of the complex regulatory mechanisms of cellular networks. The gained knowledge will help to solve problems concerning for example impact of environmental factors on plants, influence of drug treatment or molecular bases of various diseases in humans.

This publication was also supported by the Polish Ministry of Science and Higher Education, under the KNOW program.

## References

- Aebersold R., Mann M. (2003) *Mass spectrometry-based proteomics*. Nature 422: 198-207.
- Aggarwal K., Choe L.H., Lee K.H. (2006) *Shotgun proteomics using the iTRAQ isobaric tags*. Brief. Funct. Genomic. Proteomic. 5: 112-120.
- Agnolet S., Jaroszewski J.W., Verpoorte R., Staerk D. (2010) *<sup>1</sup>H NMR-based metabolomics combined with HPLC-PDA-MS-SPE-NMR for investigation of standardized Ginkgo biloba preparations*. Metabolomics 6: 292-302.
- Amirotti A., Damonte G. (2010) *Achievements and perspectives of top-down proteomics*. Proteomics 10: 3566-3576.
- Baker M. (2011) *Metabolomics: from small molecules to big ideas*. Nat. Meth. 8: 117-121.
- Berth M., Moser F., Kolbe M., Bernhardt J. (2007) *The state of the art in the analysis of two-dimensional gel electrophoresis images*. Appl. Microbiol. Biotechnol. 76: 1223-1243.
- Boyne M.T., Pesavento J.J., Mizzen C.A., Kelleher N.L. (2006) *Precise characterization of human histones in the H2A gene family by top down mass spectrometry*. J. Proteome Res. 5: 248-253.
- Catherman A.D., Skinner O.S., Kelleher N.L. (2014) *Top Down proteomics: Facts and perspectives*. Adv. OMICS-Based Discip. 445: 683-693.
- Chevallet M., Luche S., Rabilloud T. (2006) *Silver staining of proteins in polyacrylamide gels*. Nat. Protoc. 1: 1852-1858.
- Cornett D.S., Reyzer M.L., Chaurand P., Caprioli R.M. (2007) *MALDI imaging mass spectrometry: molecular snapshots of biochemical systems*. Nat. Meth. 4: 828-833.
- Coruzzi G.M., Burga A.R., Katari M.S., Gutiérrez R.A. (2009) *Systems biology: principles and applications in plant research*. In: *Plant Systems Biology*. Ed. G.M. Coruzzi and R.A. Gutiérrez, Oxford: Wiley-Blackwell: 3-40.
- Cramer R. (2009) *MALDI MS*. In: *Proteomics*. Ed. J. Reinders, A. Sickmann, Humana Press: 85-103.
- Creese A.J., Cooper H.J. (2007) *Liquid chromatography electron capture dissociation tandem mass spectrometry (LC-ECD-MS/MS) versus liquid chromatography collision-induced dissociation tandem mass spectrometry (LC-CID-MS/MS) for the identification of proteins*. J. Am. Soc. Mass Spectrom. 18: 891-897.
- Crown S.B., Antoniewicz M.R. (2013) *Parallel labeling experiments and metabolic flux analysis: Past, present and future methodologies*. Metab. Eng. 16: 21-32.
- Eberlin L.S., Dill A.L., Costa A.B., Ifa D.R., Cheng L., Masterston T., Koch M., Ratliff T.L., Cooks R.G. (2010) *Cholesterol Sulfate Imaging in Human Prostate Cancer Tissue by Desorption Electrospray Ionization Mass Spectrometry*. Anal. Chem. 82: 3430-3434.
- Eikel D., Vavrek M., Smith S., Bason C., Yeh S., Korfmacher W.A., Henion J.D. (2011) *Liquid extraction surface analysis mass spectrometry (LESA-MS) as a novel profiling tool for drug distribution and metabolism analysis: the terfenadine example: LESA-MS distribution analysis of terfenadine*. Rapid Commun. Mass Spectrom. 25: 3587-3596.
- Eng J.K., McCormack A.L., Yates III J.R. (1994) *An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database*. J. Am. Soc. Mass Spectrom. 5: 976-989.
- Fernie A.R., Aharoni A., Willmitzer L., Stitt M., Tohge T., Kopka J., Carroll A.J., Saito K., Fraser P.D., DeLuca V. (2011) *Recommendations for Reporting Metabolite Data*. Plant Cell 23: 2477-2482.
- Ficarro S.B., McClelland M.L., Stukenberg P.T., Burke D.J., Ross M.M., Shabanowitz J., Hunt D.F., White F.M. (2002) *Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae*. Nat. Biotech. 20: 301-305.
- Giddings J.C. (1987) *Concepts and comparisons in multidimensional separation*. J. High Resolut. Chromatogr. 10: 319-323.
- Goodacre R., Vaidyanathan S., Dunn W.B., Harrigan G.G., Kell D.B. (2004) *Metabolomics by numbers: acquiring and understanding global metabolite data*. Trends Biotechnol. 22: 245-252.
- Görg A., Weiss W., Dunn M.J. (2004) *Current two-dimensional electrophoresis technology for proteomics*. Proteomics 4: 3665-3685.
- Grimsrud P.A., Os D. den, Wenger C.D., Swaney D.L., Schwartz D., Sussman M.R., Ane J.-M., Coon J.J. (2010) *Large-Scale Phosphoprotein Analysis in Medicago truncatula*

- tula Roots Provides Insight into in Vivo Kinase Activity in Legumes*. *Plant Physiol.* 152: 19-28.
- Gygi S.P., Rist B., Gerber S.A., Turecek F., Gelb M.H., Aebersold R. (1999) *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags*. *Nat. Biotech.* 17: 994-999.
- Hamdan M., Righetti P.G. (2005) In: *Proteomics Today: Protein Assessment and Biomarkers Using Mass Spectrometry, 2D Electrophoresis, and Microarray Technology*, Hoboken: John Wiley & Sons: 209-224.
- Han X., Aslanian A., Yates J.R. (2008) *Mass spectrometry for proteomics*. *Curr. Opin. Chem. Biol.* 12: 483-490.
- Henzel W.J., Watanabe C., Stults J.T. (2003) *Protein identification: the origins of peptide mass fingerprinting*. *J. Am. Soc. Mass Spectrom.* 14: 931-942.
- Hirano H., Islam N., Kawasaki H. (2004) *Technical aspects of functional proteomics in plants*. *Phytochemistry* 65: 1487-1498.
- Hjerno K. (2007) *Protein Identification by Peptide Mass Fingerprinting*. In: *Mass Spectrometry Data Analysis in Proteomics*. Ed. R. Matthiesen, Humana Press: 61-75.
- Hofmann U., Maier K., Niebel A., Vacun G., Reuss M., Mauch K. (2008) *Identification of metabolic fluxes in hepatic cells from transient <sup>13</sup>C-labeling experiments: Part I. Experimental observations*. *Biotechnol. Bioeng.* 100: 344-354.
- Kelleher N.L., Lin H.Y., Valaskovic G.A., Aaserud D.J., Fridriksson E.K., McLafferty F.W. (1999) *Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry*. *J. Am. Chem. Soc.* 121: 806-812.
- Knochenmuss R. (2012) *MALDI Ionization Mechanisms: An Overview*. In: *Electrospray and MALDI Mass Spectrometry*. Ed. R.B. Cole, Hoboken: John Wiley & Sons, Inc.: 147-183.
- Lewis G.D., Farrell L., Wood M.J., Martinovic M., Arany Z., Rowe G.C., Souza A., Cheng S., McCabe E.L., Yang E., et al. (2010) *Metabolic Signatures of Exercise in Human Plasma*. *Sci. Transl. Med.* 2: 33ra37.
- Li H., Wolff J.J., Van Orden S.L., Loo J.A. (2014) *Native Top-Down Electrospray Ionization-Mass Spectrometry of 158 kDa Protein Complex by High-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. *Anal. Chem.* 86: 317-320.
- Liebeke M., Dörries K., Meyer H., Lalk M. (2012) *Metabolome Analysis of Gram-Positive Bacteria such as Staphylococcus aureus by GC-MS and LC-MS*. In: *Functional Genomics*. Ed. M. Kaufmann, C. Klinger, New York: Springer: 377-398.
- Lin D., Tabb D.L., Yates J.R. (2003) *Large-scale protein identification using mass spectrometry*. *Biochim. Biophys. Acta* 1646: 1-10.
- Liu Z., Schey K.L. (2005) *Optimization of a MALDI TOF-TOF mass spectrometer for intact protein analysis*. *J. Am. Soc. Mass Spectrom.* 16: 482-490.
- Loo J.A., Edmonds C.G., Smith R.D. (1990) *Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry*. *Science* 248: 201-204.
- Loo J.A., Edmonds C.G., Smith R.D. (1991) *Tandem mass spectrometry of very large molecules: serum albumin sequence information from multiply charged ions formed by electrospray ionization*. *Anal. Chem.* 63: 2488-2499.
- Macek B., Waanders L.F., Olsen J.V., Mann M. (2006) *Top-down protein sequencing and MS3 on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer*. *Mol. Cell. Proteomics MCP* 5: 949-958.
- Maier K., Hofmann U., Reuss M., Mauch K. (2008) *Identification of metabolic fluxes in hepatic cells from transient <sup>13</sup>C-labeling experiments: Part II. Flux estimation*. *Biotechnol. Bioeng.* 100: 355-370.
- Matthiesen R., Mutenda K. (2007) *Introduction to proteomics*. In: *Mass Spectrometry Data Analysis in Proteomics*. Ed. R. Matthiesen, Humana Press: 1-4.
- Medzihradsky K.F., Zhang X., Chalkley R.J., Guan S., McFarland M.A., Chalmers M.J., Marshall A.G., Diaz R.L., Allis C.D., Burlingame A.L. (2004) *Characterization of Tetrahymena histone H2B variants and posttranslational populations by electron capture dissociation (ECD) Fourier transform ion cyclotron mass spectrometry (FT-ICR MS)*. *Mol. Cell. Proteomics MCP* 3: 872-886.
- Mohler R.E., Dombek K.M., Hoggard J.C., Young E.T., Synovec R.E. (2006) *Comprehensive Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry Analysis of Metabolites in Fermenting and Respiring Yeast Cells*. *Anal. Chem.* 78: 2700-2709.
- Nakagami H., Sugiyama N., Ishihama Y., Shirasu K. (2012) *Shotguns in the Front Line: Phosphoproteomics in Plants*. *Plant Cell Physiol.* 53: 118-124.
- Nesvizhskii A.I., Keller A., Kolker E., Aebersold R. (2003) *A Statistical Model for Identifying Proteins by Tandem Mass Spectrometry*. *Anal. Chem.* 75: 4646-4658.
- Neuhoff V., Stamm R., Eibl H. (1985) *Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: A systematic analysis*. *Electrophoresis* 6: 427-448.
- Neverova I., Van Eyk J.E. (2005) *Role of chromatographic techniques in proteomic analysis*. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* 815: 51-63.
- Newton R.P., Brenton A.G., Smith C.J., Dudley E. (2004) *Plant proteome analysis by mass spectrometry: principles, problems, pitfalls and recent developments*. *Phytochemistry* 65: 1449-1485.
- Nielsen J., Oliver S. (2005) *The next wave in metabolome analysis*. *Trends Biotechnol.* 23: 544-546.
- Parson W.B., Koeniger S.L., Johnson R.W., Erickson J., Tian Y., Stedman C., Schwartz A., Tarcsa E., Cole R., Van Berkel G.J. (2012) *Analysis of chloroquine and metabolites directly from whole-body animal tissue sections by liquid extraction surface analysis (LESA) and tandem mass spectrometry: Chloroquine in tissue sections measured by LESA-MS/MS*. *J. Mass Spectrom.* 47: 1420-1428.
- Patel V.J., Thalassinos K., Slade S.E., Connolly J.B., Crombie A., Murrell J.C., Scrivens J.H. (2009) *A Comparison of La-*

- beling and Label-Free Mass Spectrometry-Based Proteomics Approaches*. J. Proteome Res. 8: 3752-3759.
- Perkins D.N., Pappin D.J.C., Creasy D.M., Cottrell J.S. (1999) *Probability-based protein identification by searching sequence databases using mass spectrometry data*. Electrophoresis 20: 3551-3567.
- Rabilloud T., Chevallet M., Luche S., Lelong C. (2010) *Two-dimensional gel electrophoresis in proteomics: Past, present and future*. J. Proteomics 73: 2064-2077.
- Ramautar R., Somsen G.W., de Jong G.J. (2009) *CE-MS in metabolomics*. Electrophoresis 30: 276-291.
- Resing K.A., Ahn N.G. (2005) *Proteomics strategies for protein identification*. FEBS Lett. 579: 885-889.
- Rigaut G., Shevchenko A., Rutz B., Wilm M., Mann M., Serafini B. (1999) *A generic protein purification method for protein complex characterization and proteome exploration*. Nat Biotech 17: 1030-1032.
- Roberts L.D., Souza A.L., Gerszten R.E., Clish C.B. (2012) *Targeted Metabolomics*. In: *Current Protocols in Molecular Biology*. Ed. F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl, Hoboken: John Wiley & Sons, Inc.
- Roepstorff P., Fohlman J. (1984) *Proposal for a common nomenclature for sequence ions in mass spectra of peptides*. Biomed. Mass Spectrom. 11: 601.
- Ross P.L., Huang Y.N., Marchese J.N., Williamson B., Parker K., Hattan S., Khainovski N., Pillai S., Dey S., Daniels S., et al. (2004) *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents*. Mol. Cell. Proteomics MCP 3: 1154-1169.
- Ruh H., Salonikios T., Fuchser J., Schwartz M., Sticht C., Hochheim C., Wirtzner B., Gretz N., Hopf C. (2013) *MALDI imaging MS reveals candidate lipid markers of polycystic kidney disease*. J. Lipid Res. 54: 2785-2794.
- Sadygov R.G., Yates J.R. (2003) *A Hypergeometric Probability Model for Protein Identification and Validation Using Tandem Mass Spectral Data and Protein Sequence Databases*. Anal. Chem. 75: 3792-3798.
- Sadygov R.G., Cociorva D., Yates J.R. (2004) *Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book*. Nat. Meth. 1: 195-202.
- Sarma A.D., Oehrle N.W., Emerich D.W. (2008) *Plant protein isolation and stabilization for enhanced resolution of two-dimensional polyacrylamide gel electrophoresis*. Anal. Biochem. 379: 192-195.
- Scheler C., Lamer S., Pan Z., Li X.P., Salnikow J., Jungblut P. (1998) *Peptide mass fingerprint sequence coverage from differently stained proteins on two-dimensional electrophoresis patterns by matrix assisted laser desorption/ionization-mass spectrometry (MALDI-MS)*. Electrophoresis 19: 918-927.
- Seeley E.H., Caprioli R.M. (2011) *MALDI imaging mass spectrometry of human tissue: method challenges and clinical perspectives*. Trends Biotechnol. 29: 136-143.
- Sikorska K., Rodziewicz P. (2011) *Apekty metodyczne badań proteomów roślinnych*. In: *Na pograniczu chemii i biologii, tom XXVI*. Ed. Barciszewski J., Koroniak H., Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza: 163-192.
- Siuti N., Kelleher N.L. (2007) *Decoding protein modifications using top-down mass spectrometry*. Nat. Meth. 4: 817-821.
- Smith R. (2009) *Two-Dimensional Electrophoresis: An Overview*. In: *Two-Dimensional Electrophoresis Protocols*. Ed. D. Sheehan, R. Tyther, New York: Humana: 2-17.
- Smith L.M., Kelleher N.L., Linial M., Goodlett D., Langridge-Smith P., Ah Goo Y., Safford G., Bonilla L., Kruppa G., Zubarev R., et al. (2013) *Proteiform: a single term describing protein complexity*. Nat. Meth. 10: 186-187.
- Sommerer N., Centeno D., Rossignol M. (2007) *Peptide Mass Fingerprinting*. In: *Plant Proteomics: Methods and Protocols*. Ed. Thiellement H., Zivy M., Damerval C., Mechlin V., Totowa: Humana Press: 219-234.
- Staszków A., Swarczewicz B., Banasiak J., Muth D., Jasiński M., Stobiecki M. (2011) *LC/MS profiling of flavonoid glycoconjugates isolated from hairy roots, suspension root cell cultures and seedling roots of Medicago truncatula*. Metabolomics 7: 604-613.
- Steinberg T.H., Jones L.J., Haugland R.P., Singer V.L. (1996) *SYPRO Orange and SYPRO Red Protein Gel Stains: One-Step Fluorescent Staining of Denaturing Gels for Detection of Nanogram Levels of Protein*. Anal. Biochem. 239: 223-237.
- Swanson S.K., Washburn M.P. (2005) *The continuing evolution of shotgun proteomics*. Drug Discov. Today 10: 719-725.
- Syka J.E.P., Coon J.J., Schroeder M.J., Shabanowitz J., Hunt D.F. (2004) *Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry*. Proc. Natl. Acad. Sci. U. S. A. 101: 9528-9533.
- Szyperski T. (1995) *Biosynthetically Directed Fractional <sup>13</sup>C-labeling of Proteinogenic Amino Acids. An Efficient Analytical Tool to Investigate Intermediary Metabolism*. Eur. J. Biochem. 232: 433-448.
- Takats Z. (2004) *Mass Spectrometry Sampling Under Ambient Conditions with Desorption Electrospray Ionization*. Science 306: 471-473.
- Thelen J.J. (2007) *Introduction to proteomics: A brief historical perspective on contemporary approaches*. In: *Plant Proteomics*. Springer: 1-13.
- Thiede B., Höhenwarter W., Krah A., Mattow J., Schmid M., Schmidt F., Jungblut P.R. (2005) *Peptide mass fingerprinting*. Meth. San Diego Calif. 35: 237-247.
- Tipton J.D., Tran J.C., Catherman A.D., Ahlf D.R., Durbin K.R., Lee J.E., Kellie J.F., Kelleher N.L., Hendrickson C.L., Marshall A.G. (2012) *Nano-LC/FTICR Tandem Mass Spectrometry for Top-Down Proteomics: Routine Baseline Unit Mass Resolution of Whole Cell Lysate Proteins up to 72 kDa*. Anal. Chem. 84: 2111-2117.
- Tran J.C., Zamdborg L., Ahlf D.R., Lee J.E., Catherman A.D., Durbin K.R., Tipton J.D., Vellaichamy A., Kellie J.F., Li M., et al. (2011) *Mapping intact protein isoforms in discovery mode using top-down proteomics*. Nature 480: 254-258.

- Weckwerth W. (2003) *Metabolomics in systems biology*. Annu. Rev. Plant Biol. 54: 669-689.
- Welthagen W., Shellie R.A., Spranger J., Ristow M., Zimmermann R., Fiehn O. (2005) *Comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry (GC × GC-TOF) for high resolution metabolomics: biomarker discovery on spleen tissue extracts of obese NZO compared to lean C57BL/6 mice*. Metabolomics 1: 65-73.
- Wilkins M.R., Pasquali C., Appel R.D., Ou K., Golaz O., Sanchez J.C., Yan J.X., Gooley A.A., Hughes G., Humphery-Smith I., et al. (1996) *From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis*. Biotechnol. Nat. Publ. Co. 14: 61-65.
- Winograd N. (2005) *The Magic of Cluster SIMS*. Anal. Chem. 77: 142 A – 149 A.
- Wolters D.A., Washburn M.P., Yates J.R. (2001) *An Automated Multidimensional Protein Identification Technology for Shotgun Proteomics*. Anal. Chem. 73: 5683-5690.
- Yanes O., Woo H.-K., Northen T.R., Oppenheimer S.R., Shriver L., Apon J., Estrada M.N., Potchoiba M.J., Steenwyk R., Manchester M., et al. (2009) *Nanostructure Initiator Mass Spectrometry: Tissue Imaging and Direct Biofluid Analysis*. Anal. Chem. 81: 2969-2975.
- Yates J.R. (1998) *Mass spectrometry and the age of the proteome*. J. Mass Spectrom. 33: 1-19.
- Yates J.R. (2004) *Mass spectral analysis in proteomics*. Annu. Rev. Biophys. Biomol. Struct. 33: 297-316.
- Yates J.R., Ruse C.I., Nakorchevsky A. (2009) *Proteomics by mass spectrometry: approaches, advances, and applications*. Annu. Rev. Biomed. Eng. 11: 49-79.
- Zhu W., Smith J.W., Huang C.-M. (2010) *Mass Spectrometry-Based Label-Free Quantitative Proteomics*. J. Biomed. Biotechnol. ID 840518.
- Zubarev R.A., Kelleher N.L., McLafferty F.W. (1998) *Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process*. J. Am. Chem. Soc. 120: 3265-3266.