

## METHODS

## A Path or a New Road in Laboratory Diagnostics? Biological Mass Spectrometry: Facts and Perspectives

Gábor ELEK<sup>1</sup> and Károly LAPIS

<sup>1</sup>st Institute of Pathology and Experimental Cancer Research and <sup>1</sup>Department of Pathology  
at Kútvölgyi Clinical Center, Semmelweis University, Budapest, Hungary

Proteins in tissues and biofluids and their many attributes define the proteome. Proteome can be directly correlated to known diseases and histological regions allowing the diagnosis and monitoring of disease progression as well predicting the patient's response to specific treatments. Proteomics performs large-scale, high-throughput characterization of the human proteome, among others by biological mass spectrometry. Protein-

chip technology coupled with bioinformatics is able to screen any protein source for putative disease biomarkers from a small sample volume (microliter range) by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS). This article discusses on a basic level both the technology and reliability of these methods. (Pathology Oncology Research Vol 12, No 3, 179–183)

*Key words:* serum proteome, tissue proteome, mass spectrometry, matrix, protein biomarker, SELDI, ProteinChip® system, proteomics, genomics

### Introduction and aims

Since the completion of human genome mapping,<sup>17,32</sup> continuous efforts have been made to associate these data with DNA functions (genomics); thus defining more closely the role of genetic functions in the universal system of biology. The new names reflect these efforts: transcriptomics describing messenger RNA functions, metabolomics for the regulation of the metabolism, omics technologies, which is a collective name of all these methods, etc. From time to time, a new method emerges in the course of molecular biological studies which may become an everyday tool in applied sciences. An important aim in medical practices and research is to obtain data from human samples (body fluids or tissues). In recent years, improvements in mass spectrometric methods have made it possible to quickly obtain a large amount of data from biological samples. Below we outline the practical consequences of these methodological developments.

*Received:* May 4, 2006; *accepted:* August 15, 2006

*Correspondence:* Prof emeritus Dr Károly LAPIS, <sup>1</sup>st Institute of Pathology and Experimental Cancer Research, Semmelweis University, Üllői út 26., Budapest, H-1085, Hungary. Tel: 36-1-266-1912, Fax: 36-1-266-1641, E-mail: [klapis@korbl.sote.hu](mailto:klapis@korbl.sote.hu)

### Proteome and proteomics

Genome is the total chromosomal (hereditary) content of a biological system. Consequently, proteome is the total of proteins the genome produces through biological transcription and translation (PROTEin complement to the genOME). Today we already know that the human genome consists of 30–35 thousand chromosomal genes.<sup>17,32</sup> The number of proteins constituting the proteome, however, is almost in the million range, as precursor proteins, mature proteins, degraded forms, immunoglobulins, functional and post-translational modifications (acetylation, methylation, deamination, glycosylation, oxidation, etc.) as well as secondary, tertiary and quaternary structural alterations caused by gene polymorphism all count as distinct molecules.<sup>27</sup> The genome, compared to the proteome, is stable. Proteomes are diverse depending on the type of the cell (there are approximately 250 types of cells in the human body) and its functional state; moreover, stimuli from the environment (diseases, medication, life style, or daily fluctuation) also have characteristic impacts. The amounts of distinct protein molecules compared to each other may show a ten-fold difference (e.g. serum albumin and a hormone). Human proteomes may also vary by the individuals. Hence, proteome contains a large amount of

information that may provide a tool to understanding gene functions more clearly, and could also prove useful in practice. This was already a goal even before the completion of the human genome mapping (Human Protein Index), but no methodology was available to carry it out.<sup>1</sup> Biological mass spectrometric methods represent a breakthrough exactly in this field. Proteomics research obtains information with modern analytical methods and aims to connect these data to gene functions.<sup>20</sup> As this research compares, both quantitatively and qualitatively, proteomes characteristic of different conditions to reveal biological processes, it has an important role in medical practice as well. Thus, proteomics can be applied in clinical chemistry, pharmacology, toxicology, hematology, immunology, microbiology, oncology, and even pathology.

### *Mass spectrometry*

Mass spectrometry provides measuring methods to determine the ratio of particles of different masses in a sample. The mass spectrometer (spectroscope) was first applied in the 1920s to study the isotope ratio of chemical elements. A characteristic of this method was,<sup>12</sup> as it is now, that it studied charged, i.e. ionized particles, in gas phase, with the air excluded, i.e. in vacuum. The ions are accelerated by an electrostatic field (approximately 25 kV potential difference) and they gain a velocity that is inversely proportional to their mass. In other words, smaller ions run at a higher speed, whereas bigger ones at a lower speed in the vacuum, where air molecules do not interfere with their courses. The speed of the ions in the spectrometer is a few multiples of  $10^4$  m/s. The accelerated ions pass through the accelerating grid and then those with the same speed fly together through an approximately 1-m-long field-free tube to the detector. Their impact in the detector induces a multiplying cascade, which amplifies the signal. It is an important requirement that this amplification retain the ratio and the characteristics of the signals. The software built in the spectrometer adds up the subsequent identical signals and decreases interfering noise. Smaller ions hit the detector sooner, bigger ones later; therefore, the spectrum reflects their masses (and flight times).

An important part of the spectrometer is the ionization energy source that transforms the molecules of the sample material into detectable ions. The first spectrometers used high-energy ionization methods (cathode-rays or plasma) but the methods have been refined and today the ions of bigger molecules are produced by laser beams.

A further improvement in ionization was the application of a matrix, when the material was not directly ionized but it was diluted with a substance or was enveloped in it. In the mass spectrometry of protein molecules, small-molecule organic acids were first used as matrix.<sup>15</sup> The significance of using a matrix is that it absorbs the ionizing energy

and transmits it to the macromolecules in a delicate, resonant way; therefore, it does not damage, polymerize, or denature the molecules, it only places a proton on (or takes an electron from) it.

### *The proteome of the human serum and its examination*

In the past few years, the number of known serum proteins has considerably changed, depending on the examination methods. Immunohistochemistry can distinguish about 300, gel electrophoresis almost 500, and mass spectrometry 800-1000 serum proteins.<sup>30</sup> Since the occurrence of the different proteins is not the same, if we first remove the ones that are present in highest quantity and then condense the rest, the number of distinct proteins (and peptides) reaches about 3,700. However, the examination of the amino acid structure has revealed that these are not all distinct molecules: more than one third of them are the results of post-translational (epigenetic) modifications.<sup>24</sup> The maximum number of distinct (non-redundant) serum molecules is now estimated to be 1,175.<sup>27</sup> One of the most interesting results of these examinations is that the proteins in highest number are not really serum proteins, rather cell-originated, membrane, cytoplasmic, cytoskeletal, endoplasmic reticular, or lysosomal proteins, which are present in minute quantities but may be important from the view of diagnosis.

With mass spectrometry only a part of the serum proteome can be analyzed directly. This large amount of molecules is examined with the so-called retentate chromatography mass spectrometry (RC-MS), when we analyze only one chromatographic fraction. The commercially available retentate chromatography mass spectrometry is SELDI-TOF MS (surface-enhanced laser desorption/ionization time-of-flight mass spectrometry).<sup>14</sup> The preparative and analytic processes can be sped up considerably with automation.<sup>29,33</sup>

What can we exactly obtain with this method? A spectrum profile that shows hundreds of bars. First we do not even know what bar represents what protein. However, a spectrum is as characteristic to a sample as a barcode is to a commercial item, or as a fingerprint to a person, with the difference that a spectrum can also show changes in the condition or pathological alterations. Since we see hundreds of bars in the diagram, it contains a large amount of information (part of the serum proteome), and today the main problem is that we do not know yet the normal state of the proteome.

### *Serum proteome as reflection of pathological conditions*

Most of the clinical examinations have aimed to identify spectrum bars that can be used as tumor markers. The first such project has been carried out to identify ovarian carcinoma.<sup>21</sup> In the case of this kind of cancer, the positive prediction rate of the traditionally used tumor marker, CA

125, does not even reach 30%. First, the selection of the diagnostic spectrum bars was carried out in a preliminary study, where half of the group of 100 individuals had the tumor, and the other half was the control. Among the tumorous individuals, all four clinical stages were represented, whereas in the control group cases with ovarian cysts below and above 2.5 cm in size were also included. An important criterion in this study was that all these individuals must have gone through a thorough examination (therefore, the control group also included high-risk cases). The spectra of the groups were compared using a computer software adapted from genetic studies. This software was not only able to select bars characteristic to the tumorous group, it was also capable to group the bars belonging together, thus creating a quasi family tree, and when with the expansion of the group new bars appeared in the spectrum, it evaluated that bar too; therefore, it was more and more useful in distinguishing the two groups. In the end, out of approximately 15,000 bars 8 were selected as "markers". In the second phase of the study, a previously not examined group was evaluated with mass spectrometry. It has been established that the sensitivity of this method was 95%, and the specificity and positive prediction 95% and 94%, respectively.

The number of similar studies has been growing. In 2004, a comprehensive publication<sup>27</sup> listed more than 20 articles (carcinomas of the prostate, breast, lung, kidney, urinary bladder, head and neck), out of which 14 had studied the spectrum of the serum. The sensitivity and specificity were 80-90%. The article also mentions the kinds of chips that have been used. In other studies, not only serum but urine<sup>25</sup> and other body fluids (pancreatic fluid)<sup>26</sup> have also been examined. Beside oncological studies, gerontological and toxicological<sup>34</sup> findings have also been published. Computer programs have been standardized and can now be purchased together with the chips. So far none of the authors have stated that their method can be used for screening.

#### *The use of tissue proteome in diagnosis*

Tissue proteome is more complex and harder to obtain than serum proteome. In the case of carcinoma, we need to produce pure epithelium. Tumor cell lines are available in tissue culture, and we also have cell fractioning methods. Carefully obtained cell suspensions are mixed with particles on the surface of which epithelium-specific antibodies were bound. With centrifugation, or some other methods, these particles are then regained from the suspension and the epithelial cells bound on their surfaces are detached. This way, epithelial cells from biopsy specimens can be obtained with 95% purity.<sup>28</sup> With laser capture microdissection we can even isolate single cells from a tissue. This method applies a thermoplastic film on the surface of the section. We can target laser beams at a microscopically

identified structure. For a second, the laser melts a point on the film thus binding it to the surface of the structure. When we lift off the film the selected structures are ripped off the section. With collecting these structures we can obtain a pure population,<sup>3</sup> which we can analyze with mass spectrometry.<sup>19</sup> Such methods are time consuming and can be applied only for research and not for diagnosis,<sup>18</sup> however, they allow studying phenomena that are beyond the capacity of light microscopy. With selected spectrum bars of lung biopsy specimens, not only the diagnosis of large cell carcinomas and their tissue classifications can be established, but it can also be determined whether it is a primary tumor or a metastasis. Certain spectrum bars can be associated with early metastasis formation; therefore, at the first biopsy we can already make assumptions concerning the prognosis.<sup>31</sup> There is hope that certain forms of therapy resistance can be predicted too,<sup>16</sup> which would enable clinicians to design individual therapies. To make these results applicable in clinical practice, the spectra of the total tissue samples need to be studied in correlations.

#### *Imaging mass spectrometry in the study of tissue proteome*

Mass spectra that also recorded the tissue origin were first made at the turn of the millennium,<sup>4</sup> but the development of this method still continues. Contrary to the methods described above, tissues do not need to be purified. Frozen sections have been placed on the sample holder of the spectrometer, but unfortunately contamination presented a problem, although in a different way, here too (for example, the microtome knife smeared the molecules from one lobule of the hypophysis to the other lobule). Therefore, samples have to be cut (and then treated) with exact orientation. Application of a matrix is not easy either, as the sample cannot be dissolved. Different films have been used under the frozen section so that the sample can be washed several times before and after the application of the matrix. These under-layer films were secured with two-sided adhesive tapes. Both the film and the tape had to be conductive to avoid local charge interfering with ionization; furthermore, there must not be air bubbles between them, etc.<sup>9</sup> However, washing caused the diffusion of molecules within the sample, since the section was not fixed. Formalin-containing fixation cannot be used, because it creates cross-bonds between different molecules thus preventing their desorption (i.e. sublimation in gas phase). Organic solutions and alcohol can be used for fixation here too. In the end, the matrix was applied in very fine drops on the sample with a special instrument, and microscope was used to check when the matrix reached the desired even thickness.<sup>13</sup> The sample was then dried or lyophilized. First the points to be examined have been marked on a dyed section adjacent to the sample, and then these marks have been transferred to the sample.

Later, glass slides have been successfully made conductive, and tissue stainings have been identified which, even applied on the sample, have not modified the spectrum (methylene blue). Consequently, no two sections were needed for the examination, with microscopy it was possible to mark the points (profiling) on the same section for which they wanted to obtain the spectrum.<sup>7</sup> In a further development, not only a few points of the section have been analyzed. The diameter of the laser beam has been decreased and a computer software has been developed that controlled the scanning and the ionization of the surface in raster points with identical and adjustable distances, thus providing a spectrum for the whole sample.<sup>8</sup> In a harmonized process, the detector software stores the spectrum for each point and it is able to recreate them in two different ways. On one hand, it can show all the spectra belonging to a given point of the tissue sample, which reveals the kind of molecules it consists of. On the other hand, it is capable of creating a 3-D image of the size of a spectrum bar for the whole sample, i.e. it reveals the quantity of a certain molecule in a morphological structure.<sup>5,6</sup> This is called imaging mass spectrometry (IMS).

#### *Criticism of biological mass spectrometry*

In the diagnosis of prostate cancer, several study groups have used the spectrum bars of the serum proteome selected with the method described above. The sensitivity of their method was above 90%, the specificity was between 75 and 100% (see table in ref. 27). However, among the spectrum bars used by the different groups only a few were the same, and the prostate-specific antigen (PSA), which until then had been traditionally used as a marker, was not among them. This remarkable fact has induced rightful criticism.<sup>11</sup> Although different chips have been used for the serum samples, the point in this criticism was that the proteins associated to these spectrum bars must have been present in the serum in a larger quantity than PSA, and thus obstructed its binding. For the detection of PSA with conventional methods, 20-100 mL serum is needed, whereas for mass spectrometric measurements only 2  $\mu$ L is used. The spectrum bars of proteins used in mass spectrometry are not known, and due to their larger quantity they cannot be as specific as PSA, probably they are only non-specific concomitants of prostate cancer. Although PSA is not specific to prostate carcinoma either, as it is a natural secretion of prostate epithelium and only gets into the serum because the tumor damages the basal membrane, the criticism has called the attention to the fact that the circumstances of biological mass spectrometric measurements are not standardized. Indeed, the same authors have not been able to recreate the specificity and sensitivity achieved in kidney cancer detection 10 months earlier.<sup>25</sup> Spectrometers of inferior quality, the calibration of the spectrometers, the drifting of the

instrument (the wearing out, or “getting tired”, of the detector and the ionization laser charge), moreover, slight alterations in the different series of the chips, and also the different treatments of the samples can all make the evaluation of the results difficult not only among laboratories but also within a laboratory. No doubt that the mass spectrometer is most sensitive in the examination of proteins (and peptides) with a molar mass below 30,000. It is also true that we do not fully know the normal variability of the proteome and the life span and dynamics of the different proteins. Certain proteins present in large quantities may prevent the detection of proteins in smaller quantities; therefore, the different proteomes (e.g. serum and urine) cannot be directly compared quantitatively.<sup>29</sup> In their answers the authors<sup>23</sup> have accepted the criticism and declared it necessary that the WHO, the Human Proteome Organization (HUPO), the Food and Drug Administration, the Red Cross, etc. work out international standards. However, they sustained the opinion that mass spectrometric studies of the proteome would be able to establish a series of clinically effective tumor-specific markers, without the need to produce them in a pure form first, know their structure, or produce immune sera against them.<sup>22</sup> Ovarian carcinoma marker bars have been reinvestigated with the above mentioned controls. It was established that the difficulties can be overcome by applying the more stable hybrid-quadrupole TOF spectrometer.<sup>10</sup>

#### *The future of the markers*

Searching for clinical markers is only one aspect of the work. Obviously, we need to fully understand the marker molecules associated to the spectrum bar: their structure, dynamics, tissue localization, as well as their biological role. Only then can our expectations be fulfilled to know better the functions of the genes, and also the pathological processes in the body. The first step to achieve this goal is that, after all, we need to produce pure molecules. Naturally, the expertise of physicians is not enough, we need the help of molecular biologists. The whole arsenal of proteomics: the omics technologies will be applied in this work.<sup>34</sup> In most cases, molecules have to be digested to peptides, and then the compositions and sequences of these have to be determined. Various mass spectrometric techniques are again of crucial importance in this work, because they are capable of carrying out the job,<sup>20</sup> as this is the area from where they came into clinical diagnostic use. If we know the molar mass of a peptide, we can turn to partly already existing international databases,<sup>24,28</sup> which register data for macromolecules of human origin. In the future we can also rely on further international databases that are developed in order to make practical use of proteomics.<sup>2</sup> Thus, in the end we will be able to identify the marker molecule in question. The identification of the marker of the HIP/PAP I pancreas carcinoma was a typical



and fortunate process. Right after measuring the molecular mass, it has been found in an international database. It turned out that it was identical to an already known molecule against which an immune serum already existed, too. Therefore, ELISA titration was used for measuring it in the serum and pancreatic fluid; mass spectrometric analysis became unnecessary. This protein protects the intact acinus cells from apoptosis; therefore, it is not tumor-specific either, only gets into the secretion of the pancreas because of the necrosis caused by the tumor.<sup>26</sup>

No doubt, human proteome studies will renew and expand the possibilities of clinical laboratory – more closely, the oncological – diagnostics. Whether the mass spectrometer becomes a routine diagnostic tool only future can tell. “Antibody-based approaches have dominated the clinical chemistry landscape mass spectroscopy platforms of the future, coupled to heuristic pattern recognition algorithms, may become superior to immunoassays”.<sup>23</sup>

## References

1. Anderson NG, Matheson A, Anderson NL: Back to the future: the human protein index (HPI) and the agenda for post-proteomic biology. *Proteomics* 1:3-12, 2001.
2. Apweiler R, Bairoch A, Wu CH, et al: UniProt: the universal protein knowledgebase. *Nucleic Acid Res* 32:D115-D119, 2004.
3. Bonner RF, Emmert-Buck M, Cole K, et al: Laser capture microdissection: molecular analysis of tissue. *Science* 278:1481-1483, 1997.
4. Caprioli RM, Farmer TB, Gile J: Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* 69:4751-4760, 1997.
5. Chauwand P, Schwartz SA, Rezyer ML, et al: Imaging mass spectrometry: principles and potentials. *Toxicol Pathol* 33:92-101, 2005.
6. Chauwand P, Sanders ME, Jensen RA, et al: Proteomics in diagnostic pathology. Profiling and imaging proteins directly in tissue sections. *Am J Pathol* 165:1057-1068, 2004.
7. Chauwand P, Schwartz SA, Bilheimer D, et al: Integrating histology and imaging mass spectroscopy. *Anal Chem* 76:1145-1155, 2004.
8. Chauwand P, Schwartz SA, Caprioli RM: Profiling and imaging proteins in tissue sections by mass spectrometry. *Anal Chem* 76:86A-94A, 2004.
9. Chauwand P, Stoekli M, Caprioli LM: Direct profiling of proteins in biological tissue sections by MALDI mass spectrometry. *Anal Chem* 71:5263-5270, 1999.
10. Conrads TP, Fusaro VA, Ross S, et al: High-resolution serum proteomic features for ovarian cancer detection. *Endocrine-Related Cancer* 11:163-178, 2004.
11. Diamandis EP: Proteomic patterns in biological fluids: do they represent the future of cancer diagnosis? *Clin Chem* 49:1272-1275, 2003.
12. Finkelburg W: Einführung in die Atomphysik. 9 und 10-te Auflage, Springer, Berlin-Göttingen-Heidelberg, 1964, pp 34-43, 48.
13. Garden RW, Sweedler JV: Heterogeneity within MALDI samples as revealed by mass spectrometric imaging. *Anal Chem* 72:30-36, 2000.
14. Issaq J, Conrads TP, Prieto DA, et al: SELDI TOF MS for diagnostic proteomics. *Anal Chem* 75:148A-155A, 2003.
15. Karas M, Hillenkamp F: Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60:2299-2301, 1988.
16. Lage H: Proteomics in cancer cell research: an analysis of therapy resistance. *Pathol Res Pract* 200:105-117, 2004.
17. Lander ES, Linton LM, Birren B, et al: Initial sequencing and analysis of the human genome. *Nature* 409:860-921, 2001.
18. Melle C, Kaufmann R, Hommann M, et al: Proteomic profiling in microdissected hepatocellular carcinoma using protein chip® technology. *Int J Oncol* 24:885-891, 2004.
19. Palmer-Toy DE, Sarracino DA, Sgroi D, et al: Direct acquisition of matrix-assisted laser desorption/ionization time-of-flight mass spectra from laser captured microdissected tissues. *Clin Chem* 46:1513-1516, 2000.
20. Pandey A, Mann M: Proteomics to study genes and genomes. *Nature* 405:837-846, 2000.
21. Petricoin EF, Ardekani AM, Hitt BA, et al: Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359:572-577, 2002.
22. Petricoin EF, Liotta LA: SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr Opin Biotechnol* 15:24-30, 2004.
23. Petricoin E 3rd, Liotta LA: The vision for a new diagnostic paradigm. *Clin Chem* 49:1276-1278, 2003.
24. Pieper R, Gatlin CL, Makusky AJ, et al: The human serum proteome: display of nearly 3700 chromatographically separated protein spots on two dimensional electrophoresis gels and identification of 325 distinct proteins. *Proteomics* 3:1345-1364, 2003.
25. Rogers MA, Clarke P, Noble J, et al: Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural network analysis: identification of key issues affecting potential clinical utility. *Cancer Res* 63:6971-6983, 2003.
26. Rosty C, Christa L, Kuzdzal S, et al: Identification of hepatocarcinoma-intestine-pancreas/pancreatitis associated protein I as a biomarker for pancreatic ductal adenocarcinoma by protein biochip technology. *Cancer Res* 62:1868-1875, 2002.
27. Röcken C, Ebert MPA, Roessner A: Proteomics in pathology, research and practice. *Pathol Res Pract* 200:69-82, 2004.
28. Reymond MA, Steinert R, Kähne T, et al: Expression and functional proteomics in colorectal cancer. *Pathol Res Pract* 200:119-127, 2004.
29. Seibert V, Wiesner A, Buschman T, et al: Surface-enhanced laser desorption ionization time-of-flight mass spectrometry (SELDI TOF-MS) and ProteinChip® technology in proteomic research. *Pathol Res Pract* 200:83-94, 2004.
30. Sheng Y, Jacobs JM, Comp DG, et al: Ultra-high-efficiency strong cation exchange LC/RPLC/MS/MS for high dynamic range characterization of the human plasma proteome. *Anal Chem* 76:1134-1144, 2004.
31. Yanagisawa K, Shyr Y, Xu BJ, et al: Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 362:433-439, 2003.
32. Venter JC, Adams MD, Meyers EW, et al: The sequence of the human genome. *Science* 291:1304-1351, 2001.
33. Villeneuve I, Philip I, Entenberg D, et al: Serum peptide profiling by magnetic particle assisted, automated sample processing and MALDI TOF mass spectrometry. *Anal Chem* 76:1560-1570, 2004.
34. Wetmore BA, Merrick BA: Toxicoproteomics: proteomics applied to toxicology and pathology. *Toxicol Pathol* 32:619-642, 2004.