REVIEW ARTICLE

Proteomics and its role in medicine

Petrák J.
Institute of Hematology and Blood Transfusion, Prague, Czech Republic

SUMMARY

Proteomics is a novel methodological and conceptual approach to the study of living systems, seeking to map the whole proteome, the genome protein complement. Proteomics is seeking a quantitative and qualitative description of all proteins present in a cell, tissue, or a body at a time. The basic tools of proteomics include separation techniques based on the principles of electrophoresis and chromatography serving to simplify complex protein mixtures and, most importantly, mass spectrometry used to identify individual proteins. Proteomics is a very young branch of science experiencing a turbulent development, which has shown a major potential for the study of many physiological and pathological molecular mechanisms as well as direct use in establishing the diagnosis of severe disease.
Key words: proteomics, proteome, two-dimensional electrophoresis, chromatography, mass spectrometry.

Čas. Lék. čes., 2005, 144, pp. 365-370.

The period of frantic genome sequencing is unavoidably a thing of the past and scientists are becoming increasingly aware of the fact that mere knowledge of sequences of billions of bases is by far not enough to clarify the biological functions of the body. We are increasingly asking questions related directly to gene products, that is, proteins. Proteomics involves a methodological systemic approach to the study of living systems seeking to chart the proteome, that is, quantitative and qualitative mapping of all proteins present in an organelle, cell, tissue, or body at a given moment. Proteomics also investigates how protein expression in a cell or tissue varies in response to various external and internal stimuli, and asks questions regarding interaction of individual proteins.

WHY PROTEOMICS WHEN THERE IS ALREADY GENOMICS?

Genomic studies have to date furnished a wealth of data regarding changes in the expression of individual genes during the course of many physiological or pathological states. However, this information only gives the levels of mRNA produced, not the amounts, structure, stability, and activity of proteins, which are the ultimate agents or creators of the phenotype. In fact, it has been found there is no direct and reliable correlation between mRNA levels and the amount of the corresponding protein in the cell (1, 2). While the phenotype of a body is gene-mediated, it is made up eventually by proteins. Hence, answers to questions regarding the molecular principles of biology must be sought for in the world of proteins, their modifications and complexes.

The human genome carries about 25,000 to 30,000 genes. However, most of these genes produce several mRNA variants, which may give rise to several different proteins, often with most diverse function. As a result, it has been suggested that the human proteome includes hundreds of thousands of various proteins. The structure, function and activity of proteins are modulated by more than 200 types of recognized post-translation modifications (phosphorylation, glycosylation, acetylation, etc.), constituting another level of information also in the focus of proteomics. Only exceptionally do proteins perform their function as individual molecules, as they mostly make part of very intricate functional protein complexes, which are also targets of proteomic investigations.

The molecular mechanisms of physiological processes cannot be described using merely knowledge of gene structure and expression. The structure and expression of genes only suggest "what is possible". Proteomics provides an insight into how actually the phenotype is made up, into "what really is". In addition to answering basic questions of molecular physiology, proteomics offers medical science a host of specific approaches. For example, it allows to identify differences between healthy and diseased tissue, to track changes occurring in cells exposed to the action of drugs, as well as providing an invaluable opportunity to identify biomarkers of various diseases.

TOOLS OF PROTEOMICS, 2D-PAGE

It follows from the definition of proteomics as a branch of science seeking to chart the whole proteome that the starting material of proteomic investigations are complex protein mixtures (body fluids, tissue and tumor homogenates, etc.) containing thousands of various proteins. The principle of most proteomic experiments is to separate such a complex mixture using electrophoresis or chromatography so as to be able to identify individual proteins. The most often used technique of separation is so-called two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) (3). A protein sample (serum, tissue homogenate, etc.) is first separated by isoeletric focusing on a thin strip of acrylamide gel with a pH gradient, with proteins separating by their charge (more exactly, by their isoelectric point, pI) (Fig. 1A). The isoelectric point of most proteins is within the range of pH 4-8. The gel strip with separate proteins is subsequently transferred to buffer with SDS detergent and placed on the upper margin of common acrylamide gel. When

Address for correspondence:
Jiří Petrák, MD.
128 08 Prague 2, U Nemocnice 1
Czech Republic
E-mail: petra@uhkt.cz

the electrical field perpendicular to the original direction of separation is applied, proteins migrate towards the anode and, in the presence of SDS detergent, they separate by their molecular weight (Fig. 1B). After protein-specific staining, each spot on the gel represents a specific protein found in the gel in coordinates determined by its isoelectric point (axis x) and molecular weight (axis y). The spot intensities indicate the relative expression of the respective proteins (Fig. 1C). Typical 2D-PAGE will separate and visualize up to several thousands of proteins in one gel at a time. The resulting pattern (specific separation of spots and their intensities) shows the specific profile of protein expression in a given tissue. The high reproducibility of the method allows comparing the patterns of spot separation including their intensity. For example, it is thus possible to compare protein expression in the tissue of a healthy individual with that of a diseased one, or to search for changes in protein expression produced by a drug. Proteins with differential expression or, more exactly, spots, are cut from the gel to be identified using mass spectrometry (see below).

Two-dimensional electrophoresis combined with mass spectrometry has been successfully used in numerous studies designed to unveil changes in the expression of proteins immediately related to various diseases. Examples include diseases of the heart (4), liver (5–7), nervous system (8–10), lungs (11–13), stomach (14, 15), kidneys (16, 17), blood (18–20), and many others including cancer (21–23).

PROTEIN IDENTIFICATION USING MASS SPECTROMETRY

Proteomics would be completely unthinkable if not allowing easy and inexpensive identification of individual proteins from a minimal sample. This is currently available using mass spectrometry in conjunction with bioinformatics. Mass spectrometry allows obtaining extremely accurate information about the weight and structure of molecules, in the case of proteomics specifically about the molecular weight of peptides and about their amino acid sequence. As a result, mere picomols to femtomols (in orders of 10^{-9} to 10^{-12} g) of protein are often enough to analyze and identify a protein. As whole protein molecules are too big for direct mass analysis, the starting material for the identification of protein is usually a mixture of peptides resulting from *in vitro* protein cleavage using a specific protease. The protease used most often is trypsin which, with a few exceptions, will only cleave peptide chain specifically after arginine and lysine thus generating peptides with an average length of about 10–30 amino acids.

A protein cut-out from the 2D-PAGE gel is thus first digested by trypsin and the resultant peptide mixture is extracted from the gel and purified. The mixture then serves as the starting material for protein identification. The first step in mass analysis is always peptide ionization. The most frequently employed method of ionization is matrix assisted laser desorption/ionization (MALDI). The purified mixture of peptides is mixed with so-called matrix (an organic substance absorbing laser energy), applied to a metallic target and dried. The matrix with the sample is subsequently hit by a pulse laser resulting in ionization of peptides, their desorption and transition into gaseous phase. Charged molecular peptide ions are directed into an analyzer. In the analyzer, there is vacuum with ions flying at velocities determined by their molecular weight, more exactly, by the mass to charge ratio (m/z). Very exact measurement of the time of flight of charged peptides through the analyzer makes it possible to calculate their exact molecular weight. These types of analyzer are therefore referred to as TOF (time of flight) analyzers. (Fig. 2A). In this manner, mass analysis will exactly (with a degree of accuracy of one to more decimal points) determine the molecular weights

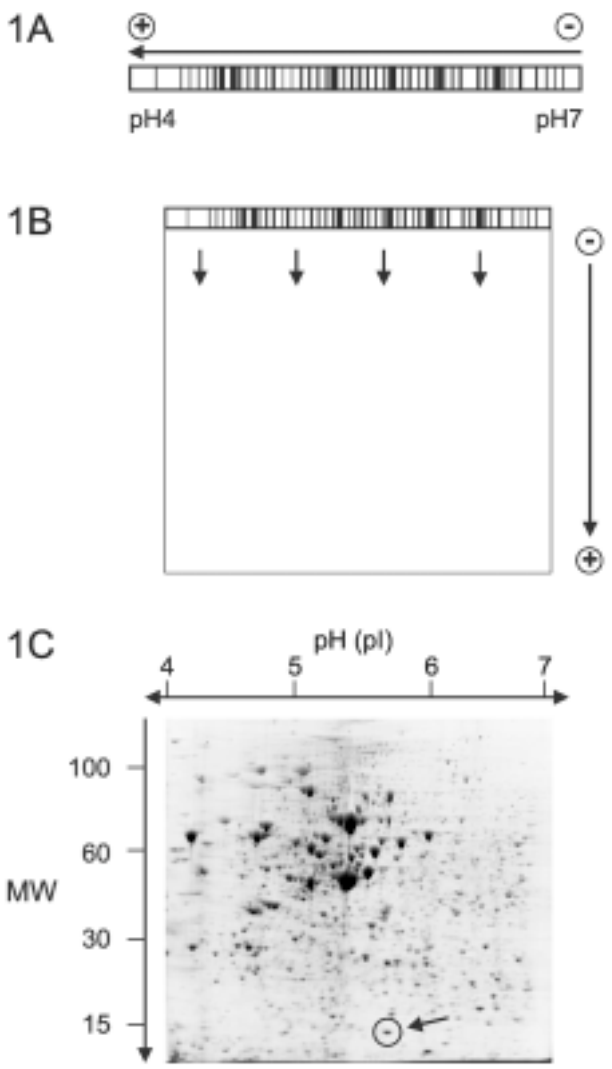


Fig. 1. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) of biological samples

1A. First dimension
The mixture is first separated by electrophoresis in a pH gradient. Proteins migrate by their actual total charge. Total charge is zero in a pH region corresponding to the isoelectric points of individual proteins. Proteins at isoelectric point thus lose mobility and are unable to migrate further. The result is separation of the whole protein mixture by pH gradient.

1B. Second dimension
A strip of gel with proteins separated in the first dimension is placed on the top of acrylamide gel containing SDS detergent. In an electrical field perpendicular to direction of original separation, proteins transit into SDS gel where they migrate and are separated based on their molecular weights.

1C. Complete lysate from human hepatoma cells (HepG2 cells) separated by 2D electrophoresis in pH 4–7 gradient and 10% SDS-PAGE
Proteins were stained by colloidal Coomassie Blue. About 800 spots can be detected on the gel. The arrow marks a selected protein whose subsequent identification is explained in the legend to Figure 2.

of individual peptides previously obtained by tryptic digestion of a selected protein. The measured molecular weights of peptides serve as entry data for searching up and identifying the protein in available protein and gene databases. The search algorithm is based on comparing the measured molecular weights of peptides with a database of theoretical molecular weights of peptides which would occur by cleavage of all known proteins by trypsin.

As protein databases are by far not complete, the hypothetical cleavage by trypsin includes all known DNA sequences translated into amino acid sequences. This technique of protein identification is referred to as peptide mass fingerprinting (see Fig. 2A). The exactly measured weights of several peptides originating from a single human protein are usually enough to locate the protein in databases. Should peptide mass fingerprinting fail to identify the protein—whatever the reason—the only option is to obtain some information about the amino acid sequence of one or several proteins. The necessary information can be obtained using so-called tandem MS-MS (Fig. 2B). Using this approach, requiring more sophisticated and expensive technology, a selected peptide, once its molecular weight has been established, is retained in the mass spectrometer and fragmented by a series of collisions with an inert gas. The weakest point of a peptide is its peptide bond so collision with gas leads to peptide fragmentation down to individual amino acids. Optimally, the fragmentation gives rise to a mixture comprising a whole peptide and its fragments of various size differing by one amino acid in their molecular weight. All fragments are "weighed" by another MS analyzer (hence MS-MS) and recorded as a mass spectrum. Differences in the measured molecular weights of fragments (corresponding to individual peaks in the spectrum) and knowledge of the molecular weights of all existing amino acids can be used to read the amino acid sequence in the spectrum. The identified sequence will then be used for direct database searches (see Fig. 2B). This approach makes it theoretically possible to determine the whole sequence of the protein and to identify peptide mutations or modifications.

In addition to MALDI, as described above, combined with one or two TOF analyzers, other techniques of ionization, also in combination with various types of analyzers (quadruple analyzer, ion trap, or the ion cyclotrone) are commonly used today. Whatever the technique, the principle of analysis of a peptide invariably involves its ionization and subsequent determination of its exact molecular weight based on its behavior in the vacuum and electromagnetic field.

CHROMATOGRAPHIC METHODS IN PROTEOMICS

Although 2D-PAGE continues to be the most important tool for proteomics, it is not by far omnipotent. Its main drawback is that it is unable to separate large (over 150 kDa) or transmembrane (highly hydrophobic) proteins. Another limiting factor is its demand on technology complicating any attempts at automation and analysis of a large number of samples. Over the past five years, these limitations have led to the development of alternative techniques of separation of complex protein or peptide mixtures based on principles of chromatography.

With chromatographic-based separation used in proteomics, the starting material is again a complex protein mixture, e. g., a tumor tissue homogenate, which is, however, exposed to specific trypsin digestion before actual separation. The result is an extremely complex peptide mixture (cleavage of an average protein with trypsin gives rise to an approx. 15–20 peptides), subsequently separated using a combination of at least two techniques of chromatography (2D chromatography). The mixture of peptides is first separated using a ion exchange chromatography, with individual collected fractions further separated by means of so-called reverse phase chromatography. Individual peptides leaving the second chromatographic column may be transferred directly to a mass spectrometer and identified. In this case, identification is only possible using MS-MS, whereby the peptide is fragmented and its amino acid sequence read. This procedure has been used, e. g., to identify a major part of

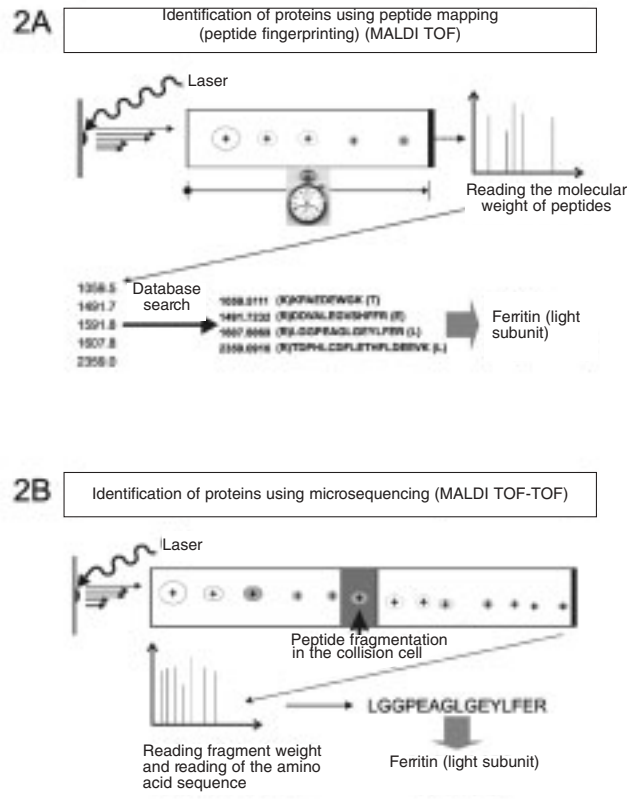


Fig. 2. Identification of proteins using mass spectrometry and bioinformatics

2A. Protein identification by peptide mapping using MALDI-TOF
The selected protein is cut from the gel and digested by trypsin. Resulting peptides are extracted from the gel, purified, and mixed with matrix (procedure performed specifically with spot marked with arrow in Fig. 1C). Laser pulses ionize peptides in the mixture, with the peptides subsequently entering a TOF (time of flight) analyzer. The TOFs of individual peptides through the analyzer are measured with a high degree of accuracy. The result is the mass spectrum of a peptide mixture. Times of flight serve to calculate the molecular weights to individual peptides (more exactly, the mass-to-charge ratio, m/z). The measured weights of peptides (originating from a single protein) are compared with all theoretical peptides (formed by trypsin-based digestion in protein and gene databases. Provided a protein represented in databases at least in the form of a DNA sequence is involved, it is identified. In this particular case, a light subunit of ferritin was analyzed and identified.

2B. Protein identification by microsequencing using tandem mass spectrometry (MALDI-TOF-TOF device)
A mixture of ionized peptides is subjected to "weighing" by means of TOF (time of flight) determination. The peptide whose amino acid sequences are to be determined is selected from the mass spectrum. Only molecules of the selected peptide are retained in the collision cell, not those of the other peptides. An inert gas is driven into the collision cell resulting in peptide fragmentation, usually at the site of the peptide bond. The resultant fragments are "weighed" by another TOF analyzer or an analyzer of another type. The resulting mass spectrum of the mixture of fragments can be used to read at least a partial amino acid sequence. This is employed to search for and identification of the starting protein in databases.

the yeast proteome (24) and close to 1,600 proteins of the human epithelium including hydrophobic proteins and proteins present in the cell in a very small number of copies (25). Among other things, chromatographic approaches in proteomics have been used with success for the study of tumor (26–28) as well as diseases of the liver (29, 30) and kidney (16, 31).

The drawbacks of highly sensitive and often high-pressure methods of chromatography are they require sophisticated and costly technology. This is offset by their ability to study even membrane

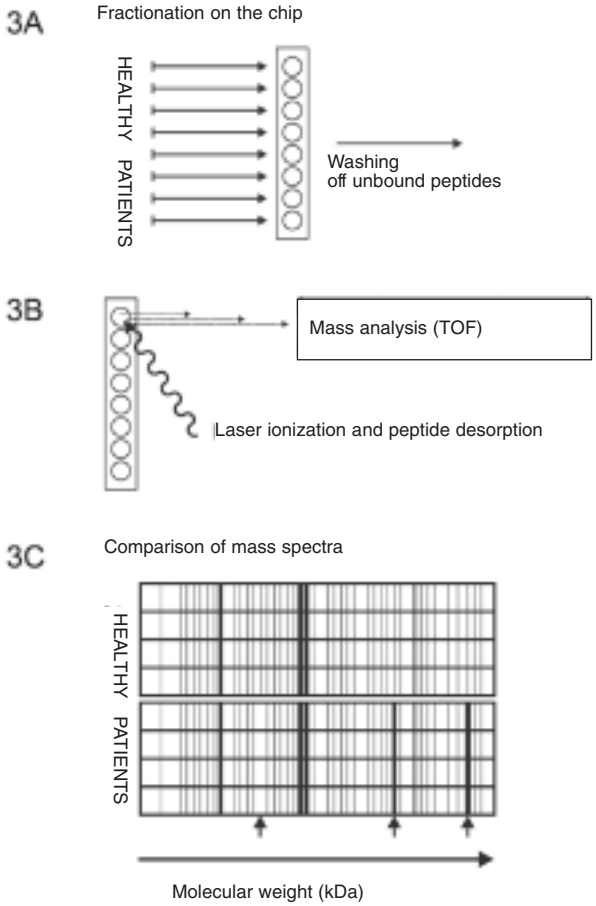


Fig. 3. Protein profiling using chip and SELDI
A protein sample (a patient's serum) is applied to a chip with a chromatographic surface **(3A)**.
The chip contains several separation sites so several samples can be analyzed at a time. The fraction of the peptides and proteins present in the serum is retained by chromatography, with the remainder washed away. A matrix is applied to the chromatographic surface with retained peptides and, when hit by a pulse laser, the peptides become ionized and their molecular weights measured in a TOF mass analyzer (**3B**). The result is a complex spectrum of the whole analyzed mixture of peptides and smaller proteins. For parallel analysis of several samples (e. g., when comparing the sera of 50 healthy individuals and 50 patients) the spectra are transformed into graphic format resembling the bar code (**3C**). Each bar represents a peptide of a specific molecular weight. Specific and statistically significant differences in the sera of healthy individuals and patients represent potential diagnostic biomarkers.

proteins (receptors, transporters, etc.) which are most attractive from the point of view of pharmacology as potential targets of various drugs. A combination of principles of chromatography with mass spectrometry also makes the basis for SELDI, a most promising technique.

SELDI, CLINPROT AND PROTEIN AND PEPTIDE PROFILING

Surface enhanced laser description ionization – time of flight (SELDI-TOF), also referred to as "protein chip array technology" is a novel, most promising approach to proteomic analysis of complex peptide and protein mixtures using a combination of chip-based microchromatography and mass spectrometry. The technique is intended mainly for analysis of peptides and small proteins present in biological samples, e. g., in the serum. Analysis using SELDI-TOF starts with selective extraction of small proteins and peptides from a sample based on the interaction with the chromatographic surface of a chip. Retained peptides are subsequently ionized directly on the chip and analyzed in a TOF mass spectrometer. The material studied using SELDI most often to date is the serum. Blood serum is very rich in proteins and peptides (it has been estimated the serum to contain thousands of various proteins and peptides) making it an almost inexhaustible source of information on the status of a body. On its transit through organ systems, the blood "collects" traces of all ongoing physiological and pathological processes. This information is carried in the form of metabolites, proteins and peptides released from cells and tissues. As a result, study of a blood plasma or serum proteome or peptidome provides a high chance to identify specific protein or peptide biomarkers applicable in establishing timely and exact diagnosis of serious diseases.

Cellular extracts or biological fluids are usually too complex for direct mass spectrometric analysis making it imperative to "simplify" the mixture by fractionation, in this particular case by chip-based microchromatography. Only a fraction of proteins of the study sample is retained on the chip by interaction with some of the types of chromatographic matrices (anex, katex, hydrophilic, hydrophobic, metal-affinity and bio-affinity) based on their (bio)chemical properties. A combination of several chromatographic surfaces and various washing conditions results in adequate fractionation of the original sample directly on the chip. The starting biological material is merely several microliters of the sample (e. g., serum). Analysis using SELDI provides a comprehensive mass spectrum—a protein profile—whereby individual peaks correspond to peptides or proteins of accurately measured molecular weight (parameters m/z – mass/charge). Several dozens to hundreds of peptides can be detected in a single spectrum. After normalization and obligatory adjustment, the spectra can be compared (e. g., protein profiles of serum samples of healthy individuals and patients with a specific type of tumors) to search for statistically significant differences in spectra indicating potential biomarkers. To make comparison of a large number of spectra easier, the spectra can be transformed into graphic format resembling the bar code. Each bar of the code represents a peptide and its position is determined by the peptide's molecular weight (Fig. 3). Potential biomarkers can be identified, again using mass spectrometry.

SELDI-TOF has recently been used to identify novel highly specific and sensitive biomarkers for the diagnosis of cancer of the urinary bladder (32), prostate (33), ovaries (34), breast (35) and, also, malignancies of the liver (36), lungs (37), pancreas (38) as well as several non-cancer conditions.

A technique analogous to SELDI developed by a company called Bruker was presented in 2004. Its principle called Clinprot involves chromatographic microfractionation of proteins and peptides, however, not on a chip but on magnetic microbeads. The main difference from SELDI is that bound peptides are released from the microbeads following fractionation and only part of the eluate mixture is used to determine the mass spectrum (protein profile). The remaining part of released peptides can be used for additional analyses, identification of peptides, immunodetection, and so on. An advantage of this approach is its flexibility as it allows combining magnetic separation with common as well as high-performance mass spectrometers enabling fragmentation and direct identification of peptides based on their amino acid sequence. As Clinprot has been established only recently, no more than two papers documenting its successful application in medical research have been published to date. Specific changes have been identified in the serum of patients with glioblastoma (39) and in asthma patients (40).

The benefits of protein profiling, either using magnetic beads (Clinprot) or the chip (SELDI) include very low consumption of the sample, potential for robotization and processing of a high number of samples with subsequent statistical analysis. As a result, both

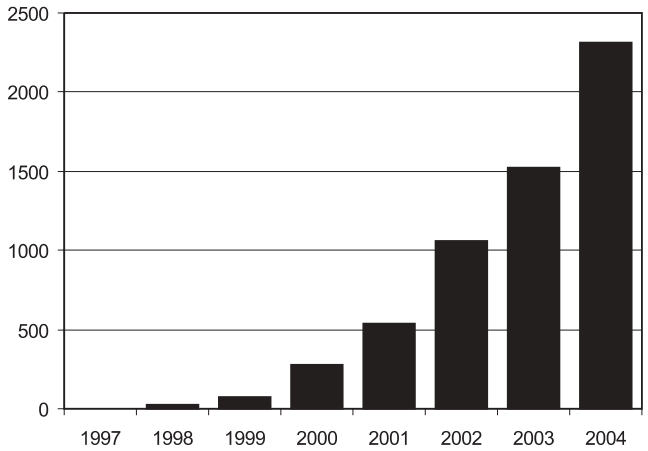
approaches have a considerable potential for routine diagnosis in the foreseeable future.

CONCLUSION

Proteomics is a young innovative branch of science experiencing turbulent development. This is supported, inter alia, by the exponential growth of scientific papers that report on studies using proteomic approaches to tackle biological issues (Fig. 4). There is almost no doubt proteomics will help, within a couple of years, to our understanding of the molecular principles of life as well as contribute to advances in the diagnosis and treatment of serious diseases (Graph 1).

Abbreviations

2D-PAGE	- two-dimensional polyacrylamide gel electrophoresis
MALDI	- matrix assisted laser desorption/ionization
MS	- mass spectrometry
MS-MS	- tandem mass spectrometry
mRNA	- messenger ribonucleic acid
SDS	- sodium dodecyl sulphate
SELDI	- surface enhanced laser desorption/ionization
SELDI-TOF	- surface enhanced laser desorption ionization – time of flight
TOF	- time of flight



Graph 1. Development of the number of scientific publications involving proteomic-based techniques searched in Medline database

REFERENCES

1. **Gygi, S. P., Rochon, Y., Franza, B. R., Aebersold, R.:** Correlation between protein and mRNA abundance in yeast. Mol. Cell. Biol., 1999, 19, pp. 1720-1730.
2. **Chen, G., Gharib, T. G., Huang, C. C. et al.:** Discordant protein and mRNA expression in lung adenocarcinomas. Mol. Cell. Proteomics, 2002, 1(4), pp. 304-13.
3. **O'Farrell, P. Z., Goodman, H. M., O'Farrell, P. H.:** High resolution two-dimensional electro-phoresis of basic as well as acidic proteins. Cell, 12, pp. 1133-1142.
4. **Van Eyk, J. E.:** Proteomics: unraveling the complexity of heart disease and striving to change cardiology. Curr. Opin. Mol. Ther., 2001, 3, pp. 546-553.
5. **Henkel, C., Roderfeld, M., Weiskirchen, R. et al.:** Identification of Fibrosis - Relevant Proteins Using DIGE (Difference in Gel Electrophoresis) in Different Models of Hepatic Fibrosis. Z. Gastroenterol., 2005, 43, pp. 23-29.
6. **Lane, C. S., Nisar, S., Griffiths, W. J. et al.:** Identification of cytochrome P450 enzymes in human colorectal metastases and the surrounding liver: a proteomic approach. Eur. J. Cancer., 2004, 40, pp. 2127-2134.
7. **Yokoo, H., Kondo, T., Fujii, K. et al.:** Proteomic signature corresponding to alpha fetoprotein expression in liver cancer cells. Hepatology, 2004, 40, pp. 609-617.
8. **Conti, A., Sanchez-Ruiz, Y., Bachi, A. et al.:** Proteome study of human cerebrospinal fluid following traumatic brain injury indicates fibrin(ogen) degradation products as trauma-associated markers. J. Neurotrauma, 2004, 21, pp. 854-863.
9. **Zhuang, Z., Lee, Y. S., Zeng, W. et al.:** Molecular genetic and proteomic analysis of synchronous malignant gliomas. Neurology, 2004, 62, pp. 2316-2319.
10. **Eun, J. P., Choi, H. Y., Kwak, Y. G.:** Proteomic analysis of human cerebral cortex in epileptic patients. Exp. Mol. Med., 2004, 36, pp. 185-191.
11. **Liu, Y., Chen, Q., Zhang, J. T.:** Tumor suppressor gene 14-3-3sigma is down-regulated whereas the proto-oncogene translation elongation factor 1delta is up-regulated in non-small cell lung cancers as identified by proteomic profiling. J. Proteome Res., 2004, 3, pp. 728-735.
12. **Granville, C. A., Dennis, P. A.:** An overview of lung cancer genomics and proteomics. Am. J. Respir. Cell. Mol. Biol., 2005, 32, pp. 169-176.
13. **Chen, G., Gharib, T. G., Huang, C. C. et al.:** Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. Clin. Cancer Res., 2002, 8, pp. 2298-2305.
14. **Ha, G. H., Lee, S. U., Kang, D. G. et al.:** Proteome analysis of human stomach tissue: separation of soluble proteins by two-dimensional polyacrylamide gel electrophoresis and identification by mass spectrometry. Electrophoresis, 2002, 23, pp. 2513-2524.
15. **Ryu, J. W., Kim, H. J., Lee, Y. S. et al.:** The proteomics approach to find biomarkers in gastric cancer. J. Korean. Med. Sci., 2003, 18, pp. 505-509.
16. **Cutillas, P. R., Chalkley, R. J., Hansen, K. C. et al.:** The urinary proteome in Fanconi syndrome implies specificity in the reabsorption of proteins by renal proximal tubule cells. Am. J. Physiol. Renal. Physiol., 2004, 287, pp. 353-364.
17. **Hoffert, J. D., van Balkom, B. W., Chou, C. L., Knepper, M. A.:** Application of difference gel electrophoresis to the identification of inner medullary collecting duct proteins. Am. J. Physiol. Renal. Physiol., 2004, 286, pp. 170-179.
18. **Gelfi, C., Vigano, A., Ripamonti, M. et al.:** A proteomic analysis of changes in prothrombin and plasma proteins associated with the G20210A mutation. Proteomics, 2004, 4, pp. 2151-2159.
19. **Harris, M. N., Ozpolat, B., Abdi, F. et al.:** Comparative proteomic analysis of all-trans-retinoic acid treatment reveals systematic post-transcriptional control mechanisms in acute promyelocytic leukemia. Blood, 2004, 104, pp. 1314-1323.
20. **Verrills, N. M., Walsh, B. J., Cobon, G. S. et al.:** Proteome analysis of vinca alkaloid response and resistance in acute lymphoblastic leukemia reveals novel cytoskeletal alterations. J. Biol. Chem., 2003, 278, pp. 45082-45093.
21. **He, Q. Y., Chen, J., Kung, H. F. et al.:** Identification of tumor-associated proteins in oral tongue squamous cell carcinoma by proteomics. Proteomics, 2004, 4, pp. 271-278.
22. **Somiari, R. I., Sullivan, A., Russell, S. et al.:** High-throughput proteomic analysis of human infiltrating ductal carcinoma of the breast. Proteomics, 2003, 3, pp. 1863-1873.
23. **Hiratsuka, M., Inoue, T., Toda, T. et al.:** Proteomics-based identification of differentially expressed genes in human gliomas: down-regulation of SIRT2 gene. Biochem. Biophys. Res. Commun., 2003, 309, pp. 558-566.
24. **Washburn, M. P., Wolters, D., Yates, J. R. 3rd.:** Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat. Biotechnol., 2001, 19, pp. 242-247.
25. **Jacobs, J. M., Motta, H. M., Yu, L. R. et al.:** Multidimensional Proteome Analysis of Human Mammary Epithelial Cells. J. Proteome Res., 2004, 3, pp. 68-75.
26. **Hamler, R. L., Zhu, K., Buchanan, N.S. et al.:** A two-dimensional liquid-phase separation method coupled with mass spectrometry for proteomic studies of breast cancer and biomarker identification. Proteomics, 2004, 4, pp. 562-577.
27. **Baker, H., Patel, V., Molinolo, A. A. et al.:** Proteome-wide analysis of head and neck squamous cell carcinomas using laser-capture microdissection and tandem mass spectrometry. Oral. Oncol., 2005, 41, pp. 183-199.

28. **Bergen, H. R. 3rd, Vasmatzis, G., Cliby, W. A. et al.:** Discovery of ovarian cancer biomarkers in serum using NanoLC electrospray ionization TOF and FT-ICR mass spectrometry. Dis. Markers., 2003–2004,19, pp. 239-249.
29. **Li, C., Hong, Y., Tan, Y. X. et al.:** Accurate qualitative and quantitative proteomic analysis of clinical hepatocellular carcinoma using laser capture microdissection coupled with isotope-coded affinity tag and two-dimensional liquid chromatography mass spectrometry. Mol. Cell. Proteomics, 2004, 3, pp. 399-409.
30. **Lane, C. S., Nisar, S., Griffiths, W. J. et al.:** Identification of cytochrome P450 enzymes in human colorectal metastases and the surrounding liver: a proteomic approach. Eur. J. Cancer, 2004, 40, pp. 2127-2134.
31. **Pisitkun, T., Shen, R. F., Knepper, M. A.:** Identification and proteomic profiling of exosomes in human urine. Proc. Natl. Acad. Sci. USA, 2004, 101, pp. 13368-13373.
32. **Zhang, Y. F., Wu, D. L., Guan, M. et al.:** Tree analysis of mass spectral urine profiles discriminates transitional cell carcinoma of the bladder from noncancer patient. Clin. Biochem., 2004, 37, pp. 772-779.
33. **Qu, Y. S., Adam, B. L., Yasui, Y. et al.:** Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients. Clin. Chem., 2002, 48, pp. 1835-1843.
34. **Petricoin, E. F., Ardekani, A. M., Hitt, B. A. et al.:** Use of proteomic patterns in serum to identify ovarian cancer. Lancet, 2002, 359, pp. 572-577.
35. **Li, J., Zhang, Z., Rosenzweig, J. et al.:** Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. Clin. Chem., 2002, 48, pp. 1296-1304.
36. **Poon, T. C., Yip, T. T., Chan, A. T. et al.:** Comprehensive proteomic profiling identifies serum proteomic signatures for detection of hepatocellular carcinoma and its subtypes. Clin. Chem., 2003, 49, pp. 752-760.
37. **Zhukov, T. A., Johanson, R. A., Cantor, A. B. et al.:** Discovery of distinct protein profiles specific for lung tumors and pre-malignant lung lesions by SELDI mass spectrometry. Lung Cancer, 2003, 40, pp. 267-279.
38. **Koopmann, J., Zhang, Z., White, N. et al.:** Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. Clin. Cancer. Res., 2004, 10, pp. 860-868.
39. **Villanueva, J., Philip, J., Entenberg, D. et al.:** Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. Anal. Chem., 2004, 76, pp. 1560-1570.
40. **Zhang, X., Leung, S. M., Morris, C. R., Shigenaga, M. K.:** Evaluation of a novel, integrated approach using functionalized magnetic beads, bench-top MALDI-TOF-MS with prestructured sample supports, and pattern recognition software for profiling potential biomarkers in human plasma. J. Biomol. Tech., 2004, 15, pp. 167-175.

The study was supported by Grant 303/04/0003 by the Grant Agency of the Czech Republic.

Translation: René Prahl

COMMENTARY

Comments by Dr Jiráček on "Proteomics and its role in medicine"

The remarkable advances in molecular biology in the last decade have made it possible to characterize the genomes of a number of animal species, topped with the completion of sequencing of the human genome reported in the spring of 2003. Current efforts should be aimed at explaining and understanding the function of genes. In the body, genetic information is expressed in the form of proteins, which are the actual functional molecules. Almost any chemical process in the body is mediated by proteins, and even a simple process such as carbon dioxide hydration has "its own" enzyme, carbonic anhydrase, accelerating the reaction rate by a factor of more than a million. Recent research has shown that the human body, containing about 100,000 genes, expresses a total of some 25,000 proteins, while it is estimated that a single "average" human cell expresses about 10,000 proteins. However, the total number of proteins and peptides occurring in the body is many times higher due to post-translational modifications, alternative splicing, existence of protein complexes, etc. Protein expression is exactly regulated, be that at the level of transcription or translation. However, there is usually bad correlation between the levels of DNA, mRNA and proteins. For these reasons, it is clear that mere knowledge of the sequence of the genome will not be enough to unravel its biological functions. Hence the renewed interest shown by researchers; a new branch of science called proteomics was born.

Proteomics can be defined as quantitative analysis of proteins present in the body at a time and under exactly defined conditions (1). In fact, it is a set of techniques and methods allowing protein analysis. The term proteome was coined in 1994 as an analogy to the protein equivalent of the genome. The proteome represents the set of all proteins expressed and subsequently modified by the genome throughout the life of a cell. The term is also used in a less general sense to express the protein composition of the body, an organ, tissue or body fluid at a specified time and under exactly defined conditions. As the proteome reflects the current metabolic status of a given cell or body, it is a highly dynamic system characteristically influenced by changes in conditions in the environment.

Address for correspondence:
Jiří Jiráček
Institute of Organic Chemistry and Biochemistry
Academy of Sciences of the Czech Republic
166 10 Prague 6, Flemingovo náměstí 2
Czech Republic
E-mail: jiracek@uochb.cas.cz

In my view, analysis of the identity and function of the proteome(s) represents a major challenge for biological and chemical sciences for at least the next decade. I am confident the ramifications of current and future discoveries in proteomics will extremely benefit medicine in particular. It is for this reason that I hail review articles helping the medical community to get familiar with proteomics, its applications and results.

The review article by Jiří Petrák offers a relatively comprehensive outline of the basic literature addressing various aspects of proteomics while concisely and clearly explaining the principles of individual techniques such as that of conventional two-dimensional gel electrophoresis (which, in my view, will long continue to be the most useful technique of protein separation), mass spectrometry or the recent chip-based techniques. It should only be regretted that more space has not been given to a more detailed description of specific proteomic results applicable in medicine, e.g., identification of cancer markers, etc.

REFERENCES

1. **Lottspeich, F.:** Proteome Analysis: A Pathway to the Functional Analysis of Proteins. Angew. Chem. Int. Ed. Engl., 1999, 38, 2476.

Translation: René Prahl