PROTEOMICS IN HUMAN DISEASE: AWARENESS OF NEW BIOMEDICAL OPPORTUNITIES

Moslem Bahadori MD, FCCP

National Research Institute for Tuberculosis and Lung Disease, Tehran, Iran

Abstract

Recent advances in biomedical technology have facilitated global analysis of cellular proteins and are termed as “proteomics”. Proteomics has become the key area of research in biomedicine and microbiology which uses meticulous and sophisticated techniques including two-dimensional acrylamid gel electrophoresis, mass spectrometry and bio-informatic databases. Proteomics is rapidly growing into post-genomic era in biomedical investigations. Its base lies on structural, physical, functional and other bio-characteristics of proteins and its application provides great opportunity to clarify the action of various pathogenetic agents. It will identify new diagnostic methods, new diagnostic markers of disease, new therapeutic agents, as well as protein candidates for vaccines. Application of this phenomenon expands from basic research, to clinical applications such as drug development, control of infectious diseases, cancer, neuropathology and cardiovascular diseases.

Keywords • Proteomics • genomic • drug development • cancer research • neuropathology • protein mapping • infectious disease

Introduction

Proteomics is the quantitative and physical mapping of cellular proteins and is widely used for studying the pathogenetic basis of diseases as well as in microbiological researches to analyze global protein synthesis as an indicator of gene exposure.1,2

Since the human genome project provided a large amount of information about the sequence of individual genes, it has become evident that much more questions exist about the outcome of their products; the proteomes. Today, many researchers do not accept the theory that “one gene produces only one protein” and they believe that more than one protein may be produced by one gene. Recent studies suggest that a fair number of proteins are produced per gene. In bacteria, for example, it is estimated to be one to two proteins per gene and in human cells, three to more than six proteins per gene.2

The focus of research is moving to identify the structures, functions and interactions of proteins produced by individual genes and their effects in various and specific disease processes. The global analysis of cellular proteins has recently been termed proteomics, which is a major topic of developing research in the post-genome period. Proteomics is a rapidly growing research area that encompasses both genetic and environmental features. The protein composition represents the functional status of a biological component.1,6

The basic need for proteomic approach in medical and biological studies

Although genomic studies give much information about the outcome of their products, but many features influence the pathway from genome to protein products. The way in which gene and its protein expression can be regulated or modified from transcription to post-translation is not a single step and differs greatly at various stages. Anderson, et al compared the amount of protein with selected mRNA in liver cells and found that the level of mRNA expression does not represent the amount of active protein in the cells. The correlation is small and transcripts can be spliced in various ways to yield different protein forms.3

*Correspondence: M. Bahadori MD FCCP, National Research Institute for Tuberculosis and Lung Disease, Masih-Daneshvari Hospital, Tehran, Iran. E-mail: bahadori@ams.ac.ir.
Modification during or after translation may show extensive variations and gene sequences do not correlate with these alterations. Phosphorylation leading to multiple protein products from a single gene is an example in this regard and is a very important process in protein function and activity.4

The study of the genome per se does not describe dynamic cellular processes and it can not represent the physical structure or the functional type of a presumed protein, because such results are under the influence of many modifications including splicing, phosphorylation, glycosylation, methylation, cleavages and complex formations within the cell.1,5

Finally, all protein products of a gene are not necessarily similar or identical in function and activity. Proteome mapping plays a significant role in providing an integrated (genomic and proteomic) view of every disease process at the protein level.

Techniques of proteomics1,2,6-8

The concept of mapping the human proteome goes back to 1982, but the opportunity of fast developing proteomics depends on new technical strategies using two-dimensional gel electrophoresis with high resolution. The main steps for displaying the proteome (proteomics) are the separation, visualization and identification of complex protein mixtures by using two-dimensional (2-D) gel electrophoresis plus mass spectrometry and bio-informatic database. Several improvements have been made in this method during the past few years but still there are new inventions in contemporary medical research work in this regard.6,7

The main goal of the techniques, which are highly sophisticated, is to visualize and identify protein spots in the cell, tissue or bacteria. The major steps are (a) solubilization, then separation of proteins8, (b) identification and characterization of proteins, and (c) application of protein bio-informatic database. This is a highly sensitive technology and can give molecular weight information at the attomole level (10^-14g) for a 10kD molecule. The protein expression patterns are analyzed by computer-based image analysis with special software, which is available via internet network which links to many sites provided from the Expasy proteomics server (www.expasy.ch/www/tools.html). In addition, there is a list of useful proteomic resources and databases available on the World Wide Web (see reference 2, page 283 for this site, and the same reference for detail of 2-D gel electrophoresis technique).

Biomedical application of proteomics

Although there is tremendous interest in the application of proteomics in biomedical research and disease study, but the maximum effects of proteomic-based approach still is in its primitive phase. Nevertheless, many new applications of this technique can be seen through the medical media and Medline and is continuously expanding. The followings are a few examples of many efforts, including infectious and bacterial diseases, cancer research, cardiovascular diseases, neuropathology and drug development.1,2,7

Proteomics application in infectious disease and microbiology9-21

There is rapid progress in microbial proteomics, due to wide availability of whole genome sequences for many bacteria. This was facilitated because bacteria have a small number of genomes. Many current applications of proteomics focus on pathogenic and nonpathogenic bacteria. Since 1983, the proteome of E. coli has been investigated and 2-D gel database of E. coli, containing 1600 different protein spots. Information about many of these proteins and how they are expressed and changed according to different bacterial growth conditions are available in database. E. coli has 4000 genes, 350 of which have been matched to approximately 400 protein spots.9,10

Similar investigations have been applied on Salmonella typhimorium, Streptococcus pneumoniae, Helicobacter pylori, yeast, and in particular Mycobacterium tuberculosis (MTB) and Mycobacterium bovis. The main objectives for most of these bacterial studies have been search for (a) new diagnostic markers, (b) new antigens for vaccine candidates, (c)markers for virulence, (d) determinants for drug/antibiotic resistant strains and (e) proteins responsible for mycobacterial dormancy which is resistant to anti-TB therapy.11,18

For some bacteria, these data are either completed or near complete and 2-D electrophoresis database are available on the Internet (www.expasy.ch/). The biogenesis and maturation of phagosomes are areas of study, which have employed aspects of proteomic analyses. This study has been employed for MTB.
Proteomics in Human Disease

Proteomics has been used in the study of immune responses in infections with different bacteria. Study with *H. pylori* identified 20 proteins, which were reactive with the serum of infected patients. This has been done in the hope to find a protein candidate for generation of vaccines. In the case of *MTB*, although BCG vaccination has been used for many years but the results are controversial from zero percent effective in the South India to 70% benefits in England. A novel protein candidate through proteomics technology has been introduced by some investigators, which has been incorporated into trial vaccines in animal models. Study on genetic differences between virulent *MTB* and avirulent *M. bovis*, identified that three distinct regions were present in the virulent strain but absent in the avirulent strain. When these three were introduced to the avirulent *M. bovis* strain, they became virulent. Mycobacterium has more complex technology. Among the 4000 genes in *MTB*, the target proteins vary. Researchers are searching for (1) new drug invention, (2) virulence protein factors, (3) drug resistance factors, and (4) protein candidates for vaccine.19-21

In summary, the use of proteomics technology provides insight for investigating the epidemiology and taxonomy of human infectious agents and identification of novel mechanisms in pathogenesis, biogenesis, sensitivity and resistance of many bacteria.

Drug development22-25

Genomics and proteomics are well established fields in drug discovery and in combination with combinatorial chemistry and high-throughput screening, are helping to bring forward an unprecedented number of potential lead compounds.22 Genomic data, i.e. DNA/RNA sequences identified by subtractive expression pattern analyses, are not enough per se for a clear identification of a therapeutic target, mainly because proteins and not DNA/RNA are the loci for drug mode-of-action. Most drugs exert their effects on proteins. The effects seen on the gene expression level are only a response to drug effects on protein level. As mentioned, there is not always a direct correlation between gene expression and protein expression. Proteomics has demonstrated proof-of-concept in toxicology as shown by a number of successful applications in mechanistic toxicology and lead selection. Proteomics technology is now starting to impact the clinical phase of drug development. Generally drug development is focused to upregulate or downregulate a specific protein activity implicated in the pathogenesis of disease. The goal of pharmaceutical proteomics is to provide23,24:

- a drug to counteract or correct protein products of specific genetic-based lesions.
- a drug for neutralizing antibodies or inhibiting a receptor, for example, receptor for BCR-ABL, tyrosine kinase in chronic myeloid leukemia.
- mechanistic-based grouping of lead compounds is reflected in its effects on proteins.
- another important use of proteomics technology for drug-target discovery is to identify side effect toxicity of a presumed drug.
- another approach in this field is to present database of molecular pharmacology of cancer, which has been shown by Myers, et al.25

In summary, a substantial amount of research, which is published in the medical literature and is available in the public domain, exhibits the potential of this new drug development in the preclinical phases. Sandra Steiner and her colleague predicted that over the next few years, it is anticipated that functional genomics and proteomics will have major impact on the clinical phases of drug development. Many pharmaceutical companies now eagerly investigate in this field.

Proteomics in cancer research26-34

Studies of proteomics in human tumors have led to the identification of a variety of proteins, potentially useful as novel antigens or markers, for diagnostic, prognostic and therapeutic purposes.8,26 Molecules implicated in carcinogenesis are increasingly being investigated. Many changes in gene-exposure, between benign and malignant human tumors are due to post-translational modification not detected by DNA/RNA analyses. Indeed, proteomics-based studies of many tumor types are now underway and more complementary techniques to improve this procedure are being used. Among these tumors, tumors of the urinary tract, including prostate cancer, have been searched for most, and 2-D database for cancer of the bladder, that includes protein profiles of both transitional and squamous cell carcinoma is already available.27-29 Urological malignancies account for approximately 16% of all cancers and their five-year survival rate is poor, because many of these malignancies are diagnosed after metastasis, hence the early detection of these tumors is critical. Several protein markers from
early change of metaplasia to invasive cancer have been proposed. In addition, these markers will allow the prediction of prognosis and response to therapy, and they may be used as therapeutic targets. Proteome analyses have also yielded information about tumor heterogeneity and the degree of relation between primary tumors and their metastases. Celis and colleagues, developed strategies to identify squamous cell carcinoma metaplastic lesions of the bladder, as well as biomarkers in the urine for follow-up studies of patients with squamous cell carcinoma.28,29

In another study, Ostergaard and co-workers, with 2-D electrophoresis, found six cases of squamous cell carcinoma among the 150 cases with bladder carcinoma, which were supposed to be transitional, on the basis of characteristic protein-expression patterns.30 Psoriasin, keratin 10 and 14, PA-FABP, galectin 7 and stratifin were particularly valuable in assessing the degree of differentiation. Emmert-Buck and his colleagues in the National Cancer Institute, by proteomic strategy, analyzed two examples of esophageal cancer from two patients.30 They microdissected 50000 cells and resolved approximately 675 distinct proteins or isoforms with molecular weight ranging between 10 to 200 kD and isoelectric points of pH 3-10. Seventeen proteins showed tumor-specific alterations; 10 were uniquely present in the tumor and seven were observed only in normal epithelium.30

Study for the differentiation between benign and malignant tumors was another aspect of proteomics techniques. Other cancers such as, lymphomas, leukemias and carcinoma of the liver and lung, were another field of investigation.31,32

Proteomics technology is also being used in other areas of cancer research including tumor development, tumor progression and tumor regression and classification.33 One of the most important areas is the study of apoptosis.34 This has been particularly used in the study of malignant lymphoma, in particular nuclear protein expression in Burkitt’s lymphoma BL60 cell. In addition, there are many other cancers such as cancer of the liver, colon, rectum and bronchial tumors, which are under study by this technology.

**Proteomics in neuropathology**35-43

The normal functioning of the central nervous system (CNS) requires complex interaction among numerous biological compounds. The pathophysiology of perturbation in this system is as complex as that of neurological diseases. Many methods examine the biological output of dysfunctional cells in a diseased system, but most are not sufficient for complete diagnosis.35,36

A 2-D electrophoresis database of normal brain proteins has been created from samples of total brain tissue from the parietal lobe cortex. In the database, 400 identical spots correspond to 180 different brain proteins, mostly of cytoplasmic or mitochondrial origin. This should serve as a useful reference against which gels prepared from brain extracts of abnormal individuals can be compared. Hence, proteomics has been used to investigate certain neuropathological disorders including Creutzfeld-Jakob disease (CJD), its variant (vCJD) or bovine spongiform encephalopathy (BSE), Alzheimer’s disease (AD), Down syndrome (DS) and demyelination disease. These studies have been performed in both in animal models and human patients. Noneka, et al used proteomics to detect differences in protein expression between the brain of controls and patients with Down syndrome and Alzheimer’s disease. They showed that the expression of the dehydropryminase-related protein 2 (DRP-2) in DS and AD brain, is downregulated at the mRNA level and dysregulated at the protein level. This may help to explain deranged migration and histogenesis of DS brain and wiring of AD brain.38 The major protein component of the neurofibrillary tangle in AD is tau proteins. Tau is an a microtubule-associated protein and exists as six alternatively spliced isoforms. Many studies through 2-D electrophoresis have been done on this protein as well as on the amyloid-associated protein in this disease.39

For discrimination of CJD and other types of dementia, Zerr and coworkers demonstrated two proteins (p130 and p131) by the proteomics method. These proteins are a member of the 14-3-3 family, which is important in this disease. Several other studies with 2-D gel-electrophoresis confirmed the value of the 14-3-3 protein as a discriminatory maker for CJD in the brain of patients with dementia. In the case of BSE (similar to vCJD), the infectious particle (prion) supposed to be a conformational isomer of a normal cellular glycoprotein (PrPc), and nominated PrPcs, this is unusual and resistant to proteolytic degradation, or partial degrades with protease-treatment, leading to increased electrophoretic mobility of the three different from CJD particles.41 This particle can also be detected from the tonsils of patients with
vCJD isoforms. The proteomics approach has also been used to investigate demyelination disease in animal models.

**Proteomics in heart disease**

The pathogenesis of cardiac dysfunction is still largely unknown. So research in this field is mandatory. Recently, the application of proteomics found priority in the study of cardiovascular disease. Both animal models as well as human patients became the target of investigation. At present, the focus is on dilated cardiomyopathy, which is multifactorial. Jungblut and coworkers compared the patterns of 2-DE for dilated cardiomyopathy proteomics with those of controls. They discovered 25 statistically significant intensity differences, 12 of which were identified by amino acid analysis through matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS). They constructed a human myocardial 2-DE database containing 3300 protein spots and 150 identified protein species. Similarly, Li and colleagues separated more than 3000 myocardial protein species from the Wistar Kyoto rat, which is an important animal model using the HR2-DE technique. Indeed, these studies, as well as other studies performed in this field, have shown that the expression of approximately 100 cardiac proteins is significantly different from normal in dilated cardiomyopathy, and most proteins are fewer in the diseased than in the normal heart. Weekes, et al in cattle, and Field and coworkers in bovine dilated cardiomyopathy, found a seven-fold increase in the enzyme ubiquitin C-terminal hydrolase and supported the suggestion that inappropriate ubiquitination and subsequent protein degradation by the proteasome may contribute to the development of heart failure. Finally, there is federated 2-DE database of human cardiac proteins (www.expasy.ch/ch2d/2d-index.html).

**CONCLUSION**

Proteomics is an exciting new method, opportunity, and biomedical approach in the examination of pathological processes and is of value in facilitating the understanding of many conditions. Proteomics complements genomics-based approaches and provides additional information on the molecular basis of pathogenetic pathways. This is a dynamic procedure and has more sophisticated technology, but needs further investigation for the improvement of techniques. While it offers new opportunities in biomedical approach to diseases, but still some limitations exist. Formalin-fixed archival tissue for instance, cannot be used by 2-DE. Another problem is that it is time-consuming. Lefkowitz, et al presented the analytical limitations in obtaining information on polypeptides in 2-D gel spot. Although limitations exist, proteomics, being much more than a basic biomedical research, has been used in clinical practice for the prevention, diagnosis, treatment and follow up of disease.

**References**

4. Humphrey-Smith I, Corwell SJ, Blackstock WP. Proteome research; complementary and limitations with respect to the RNA and DNA worlds. Electrophoresis. 1997; 18: 1217-42.


