Applications Of Proteomics

M Abhilash

Citation

M Abhilash. *Applications Of Proteomics*. The Internet Journal of Genomics and Proteomics. 2008 Volume 4 Number 1.

Abstract

Proteomics is the large-scale study of proteins, particularly their structures and functions. Proteins are vital parts of living organisms, as they are the main components of the physiological metabolic pathways of cells. The term "proteomics" was coined to make an analogy with genomics, the study of the genes. The word "proteome" is a blend of "protein" and "genome". The proteome is the entire complement of proteins, including the modifications made to a particular set of proteins, produced by an organism or system. This will vary with time and distinct requirements, or stresses, that a cell or organism undergoes.

INTRODUCTION

The word “proteome” is derived from PROTEins expressed by a genOME, and it refers to all the proteins produced by an organism, much like the genome is the entire set of genes.Proteomics is the large-scale study of protein, particularly their structures and functions. This term was coined to make an analogy with genomics, and while it is often viewed as the “next step”, proteomics is much more complicated than genomics. Most importantly, while the genome is a rather constant entity, the proteome differs from cell to cell and is constantly changing through its biochemical interactions with the genome and the environment. One organism has radically different protein expression in different parts of its body, in different stages of its life cycle and in different environmental conditions.

The entirety of proteins in existence in an organism throughout its life cycle, or on a smaller scale the entirety of proteins found in a particular cell type under a particular type of stimulation, are referred to as the proteome of the organism or cell type respectively. Since proteins play a central role in the life of an organism, proteomics is instrumental in discovery of biomarkers, such as markers that indicate a particular disease.

With completion of a rough draft of the human genome, many researchers are now looking at how genes and proteins interact to form other proteins. A surprising finding of the Human Genome Project is that there are far fewer protein-coding genes in the human genome than there are proteins in the human proteome (~20,000 to 25,000 genes vs. ~1,000,000 proteins). The large increase in protein diversity is thought to be due to alternative splicing and post-translational modification of proteins. This discrepancy implies that protein diversity cannot be fully characterized by gene expression analysis alone, making proteomics a useful tool for characterizing cells and tissues of interest.

BRANCHES OF PROTEOMICS

Protein separation. All proteomic technologies rely on the ability to separate a complex mixture so that individual proteins are more easily processed with other techniques.

Protein identification. Well-known methods include low-throughput sequencing through Edman degradation. Higher-throughput proteomic techniques are based on mass spectrometry, commonly peptide mass fingerprinting on simpler instruments, or De novo repeat detection sequencing on instruments capable of more than one round of mass spectrometry. Antibody-based assays can also be used, but are unique to one sequence motif.

Protein quantification. Gel-based methods are used, including differential staining of gels with fluorescent dyes (difference gel electrophoresis). Gel-free methods include various tagging or chemical modification methods, such as isotope-coded affinity tags (ICATs), metal coded affinity tags (MeCATs) or combined fractional diagonal chromatography (COFRADIC). Modern day gel electrophoresis research often leverages software-based image analysis tools primarily to analyze bio-markers by quantifying individual, as well as showing the separation between one or more protein “spots” on a scanned image of a 2-DE product. Additionally, these tools match spots between gels of similar samples to show, for example, proteomic differences between early and advanced stages of

an illness.

Protein sequence analysis. This is more of a bioinformatic branch, dedicated to searching databases for possible protein or peptide matches, but also functional assignment of domains, prediction of function from sequence, and evolutionary relationships of proteins.

Structural proteomics. This concerns the high-throughput determination of protein structures in three-dimensional space. Common methods are x-ray crystallography and NMR spectroscopy.

Interaction proteomics. This concerns the investigation of protein interactions on the atomic, molecular and cellular levels. see related article on Protein-protein interaction prediction.

Protein modification. Almost all proteins are modified from their pure translated amino-acid sequence, so-called post-translational modification. Specialized methods have been developed to study phosporylation (phosphoproteomics) and glycosylation (glycoproteomics).

Cellular proteomics. A new branch of proteomics whose goal is to map the location of proteins and protein-protein interactions in whole cells during key cell events. Centers around the use of techniques such as X-ray Tomography and optical fluorescence microscopy.

Experimental bioinformatics. A branch of bioinformatics, as it is applied in proteomics, coined by Mathias Mann. It involves the mutual design of experimental and bioinformatics methods to create (extract) new types of information from proteomics experiments.

KEY TECHNOLOGIES USED IN PROTEOMICS

One- and two-dimensional gel electrophoresis are used to identify the relative mass of a protein and its isoelectric point.

X-ray crystallography and nuclear magnetic resonance are used to characterize the three-dimensional structure of peptides and proteins. However, low-resolution techniques such as circular dichroism, Fourier transform infrared spectroscopy and small angle x-ray scattering can be used to study the secondary structure of proteins.

Tandem mass spectrometry combined with reverse phase chromatography or 2-D electrophoresis is used to identify (by de novo peptide sequencing) and quantify all the levels of proteins found in cells.

Mass spectrometry (no-tandem), often MALDI-TOF, is used to identify proteins by peptide mass fingerprinting. Less commonly this approach is used with chromatography and/or high resolution mass spectrometry.

This technique is becoming less used and the scientific world no longer accepts absolute identification of a protein based solely on peptide mass fingerprint data.

Affinity chromatography, yeast two hybrid techniques, fluorescence resonance energy transfer (FRET), and Surface Plasmon Resonance (SPR) are used to identify protein-protein and protein-DNA binding reactions.

X-ray Tomography used to determine the location of labelled proteins or protein complexes in an intact cell. Frequently correlated with images of cells from light based microscopes.

Software based image analysis is utilized to automate the quantification and detection of spots within and among gels samples. While this technology is widely utilized, the intelligence has not been perfected yet. For example, the leading software tools in this area tend to agree on the analysis of well-definedm well-separated protein spots, but they deliver different results and tendencies with less-defined less-separated spots - thus necessitating manual verification of results.

APPLICATION OF PROTEOMICS FOR DISCOVERY OF PROTEIN BIOMARKERS

Biomarkers of drug efficacy and toxicity are becoming a key need in the drug development process. Mass spectral-based proteomic technologies are ideally suited for the discovery of protein biomarkers in the absence of any prior knowledge of quantitative changes in protein levels. The success of any biomarker discovery effort will depend upon the quality of samples analysed, the ability to generate quantitative information on relative protein levels and the ability to readily interpret the data generated. This review will focus on the strengths and weaknesses of technologies currently utilized to address these issues.

APPLICATION OF PROTEOMICS IN THE STUDY OF TUMOR METASTASIS.

Tumor metastasis is the dominant cause of death in cancer patients. However, the molecular and cellular mechanisms underlying tumor metastasis are still elusive. The identification of protein molecules with their expressions correlated to the metastatic process would help to understand

the metastatic mechanisms and thus facilitate the development of strategies for the therapeutic interventions and clinical management of cancer. Proteomics is a systematic research approach aiming to provide the global characterization of protein expression and function under given conditions. Proteomic technology has been widely used in biomarker discovery and pathogenetic studies including tumor metastasis. This article provides a brief review of the application of proteomics in identifying molecular factors in tumor metastasis process. The combination of proteomics with other experimental approaches in biochemistry, cell biology, molecular genetics and chemistry, together with the development of new technologies and improvements in existing methodologies will continue to extend its application in studying cancer metastasis.

APPLICATION OF PROTEOMICS TECHNOLOGY TO THE FIELD OF NEUROTRAUMA

Near-completion of the Human Genome Project has stimulated scientists to begin looking for the next step in unraveling normal and abnormal functions within biological systems. Consequently, there is new focus on the role of proteins in these processes. Proteomics is a burgeoning field that may provide a valuable approach to evaluate the post-traumatic central nervous system (CNS). Although we cannot provide a comprehensive assessment of all methods for protein analysis, this report summarizes some of the newer proteomic technologies that have propelled this field into the limelight and that are available to most researchers in neurotrauma.

Three technical approaches (two-dimensional gel electrophoresis, direct analysis by mass spectrometry, including two-dimensional chromatography coupled to mass spectrometry and isotope coded affinity tags, and antibody technologies) are reviewed, and their advantages and disadvantages presented. A discussion of proteomic technology in the context of brain and spinal cord trauma follows, addressing current and future challenges. Proteomics will likely be very useful for developing diagnostic predictors after CNS injury and for mapping changes in proteins after injury in order to identify new therapeutic targets. Neurotrauma results in complex alterations to the biological systems within the nervous system, and these changes evolve over time. Exploration of the “new nervous system” that follows injury will require methods that can both fully assess and simplify this complexity.

APPLICATION OF PROTEOMICS IN RENAL DISEASE DIAGNOSIS.

Proteomics is widely envisioned as playing a significant role in the translation of genomics to clinically useful applications, especially in the areas of diagnostics and prognostics. In the diagnosis and treatment of kidney disease, a major priority is the identification of disease-associated biomarkers. Proteomics, with its high-throughput and unbiased approach to the analysis of variations in protein expression patterns (actual phenotypic expression of genetic variation), promises to be the most suitable platform for biomarker discovery. Combining such classic analytical techniques as two-dimensional gel electrophoresis with more sophisticated techniques, such as MS, has enabled considerable progress to be made in cataloguing and quantifying proteins present in urine and various kidney tissue compartments in both normal and diseased physiological states.

Despite these accomplishments, there remain a number of important challenges that will need to be addressed in order to pave the way for the universal acceptance of proteomics as a clinically relevant diagnostic tool. We discuss issues related to three such critical developmental tasks as follows: (i) completely defining the proteome in the various biological compartments (e.g.tissues, serum and urine) in both health and disease, which presents a major challenge given the dynamic range and complexity of such proteomes (ii) achieving the routine ability to accurately and reproducibly quantify proteomic expression profiles; and

(iii) developing diagnostic platforms that are readily applicable and technically feasible for use in the clinical setting that depend on the fruits of the preceding two tasks to profile multiple disease biomarkers.

THE APPLICATION OF PROTEOMICS IN NEUROLOGY

Rapidly progressing proteomics techniques have been widely adopted in most areas of biology and medicine. In neurology and neuroscience, many applications of proteomics have involved neurotoxicology and neurometabolism, as well as in the determination of specific proteomic aspects of individual brain areas and body fluids in neurodegeneration. Investigation of brain protein groups in neurodegeneration, such as enzymes, cytoskeleton proteins, chaperones, synaptosomal proteins and antioxidant proteins, is in progress as phenotype related proteomics. The concomitant detection of several hundred proteins on a gel

provides sufficiently comprehensive data to determine a pathophysiological protein network and its peripheral representatives. The rapid spread of proteomics technology, which principally consists of twodimensional gel electrophoresis (2-DE) with in-gel protein digestion of protein spots and identification by massspectrometry, has provided an explosive amount of results. An additional advantage is that hitherto unknown proteins have been identified as brain proteins. The current proteomics methods, however, have shortcomings and disadvantages. We would emphasize the failure to separate hydrophobic proteins as a major problem. So far, we have been unable to analyze the vast majority of these proteins in gels on 2-DE. There are several other analytical problems which also need to be overcome, and once solved, will allow for a more comprehensive analysis of the individual disease process. Here, we have reviewed the recent progress in proteomics research on neurodegeneration, with reference to its technological utility and problems in clinical application.

APPLICATION OF PROTEOMICS TO FETAL AND MATERNAL MEDICINE

The recent elucidation of the human genome sequence has provided a wealth of useful information but does not provide information on diseases caused by changes at the protein level. Proteomics includes the characterization and functional analysis of all proteins that are expressed by the genome at a certain moment, under certain conditions. Since expression levels of many proteins strongly depend on complex, but well-balanced regulatory systems, the proteome, unlike the genome, is highly dynamic. This variation depends on the biological function of a cell, but also on signals from its environment. In (bio) medical research it has become increasingly apparent that cellular processes, in particular in disease, are determined by multiple proteins. Hence it is important not to focus on one single gene product (one protein), but to study the complete set of gene products (the proteome). In this way the multi-factorial relations underlying certain diseases may be unraveled potentially identifying therapeutical targets. For many diseases characterization of the functional proteome is crucial for elucidating alterations in protein expression and modifications. When proteins undergo non-genetically determined alterations such as alternative splicing, or post-translational modifications, e.g. phosphorylation or glycosylation, it may affect their function. Although abnormalities in splicing or post-translational modifications can cause a disease process, they can also be a consequence. An example is that patients with diabetes have high blood glucose which glycosylates hundreds or even thousands of proteins, including HbA1c which is used to monitor diabetic control.

PROTEOMICS IN UROLOGICAL CANCER RESEARCH

Proteomic analysis allows the comparison of the proteins present in a diseased tissue sample with the proteins present in a normal tissue sample. Any proteins, which have been altered either quantitatively or qualitatively between the normal and diseased sample are likely to be associated with the disease process. These proteins can be identified and may be useful as diagnostic markers for the early detection of the disease or prognostic markers to predict the outcome of the disease or they may be used as drug targets for the development of new therapeutic agents. The purpose of this review is to outline the principle technologies involved in proteome analysis and indicate current and potential future applications of proteomic analysis in urological cancer research.

APPLICATION OF PROTEOMICS IN AUTOANTIBODY PROFILING FOR THE STUDY AND TREATMENT OF AUTOIMMUNE DISEASE.

Proteomics technologies enable profiling of autoantibody responses using biological fluids derived from patients with autoimmune disease. They provide a powerful tool to characterize autoreactive B-cell responses in diseases including rheumatoid arthritis, multiple sclerosis, autoimmune diabetes, and systemic lupus erythematosus. Autoantibody profiling may serve purposes including classification of individual patients and subsets of patients based on their 'autoantibody fingerprint', examination of epitope spreading and antibody isotype usage, discovery and characterization of candidate autoantigens, and tailoring antigen-specific therapy. In the coming decades, proteomics technologies will broaden our understanding of the underlying mechanisms of and will further our ability to diagnose, prognosticate and treat autoimmune disease.

APPLICATION OF PROTEOMICS IN CARDIOVASCULAR RESEARCH.

The development of proteomics is a timely one for cardiovascular research. Analyses at the organ, sub cellular, and molecular levels have revealed dynamic, complex, and subtle intracellular processes associated with heart and vascular disease. The power and flexibility of proteomic analyses, which facilitate protein separation, identification,

and characterization, should hasten our understanding of these processes at the protein level. Properly applied, proteomics provides researchers with cellular protein "inventories" at specific moments in time, making it ideal for documenting protein modification due to a particular disease, condition, or treatment. This is accomplished through the establishment of species- and tissue-specific protein databases, providing a foundation for subsequent proteomic studies.

Evolution of proteomic techniques has permitted more thorough investigation into molecular mechanisms underlying cardiovascular disease, facilitating identification not only of modified proteins but also of the nature of their modification. Continued development should lead to functional proteomic studies, in which identification of protein modification, in conjunction with functional data from established biochemical and physiological methods, has the ability to further our understanding of the interplay between proteome change and cardiovascular disease.

APPLICATION OF PROTEOMICS TO DIABETES RESEARCH.

Proteomics is the investigation of all the proteins and their various modifications making up a system, be that a cell, tissue or organism. The techniques involved in proteomics allow the global screening of complex samples of proteins and provide qualitative and quantitative evidence of altered protein expression. This lends itself to the investigation of the molecular mechanisms underpinning disease processes and the effects of treatment. This review describes the main techniques of proteomics and how they have begun to be applied to diabetes research.

ROLE OF PROTEOMICS IN NUTRITION RESEARCH

There are about 100,000 proteins in humans with various physiological functions. The complement of proteins in the organism as well as their interactions is defined as the proteome. Its analysis (proteomics) by highly specific, sensitive, and accurate MS has been made possible with matrix-assisted laser desorption ionization or electrospray ionization of proteins and large peptides. Currently, the most commonly used proteomics technologies involve either specific digestion of proteins (the bottom-up approach using 2-dimensional polyacrylamide gel electrophoresis and multidimensional

Protein identification technology) or direct analysis of intact proteins after their chromatographic separation (the top-down approach and surface-enhanced laser desorption ionization).

Proteomics holds great promise for discoveries in nutrition research, including profiles and characteristics of dietary and body proteins; digestion, absorption, and metabolism of nutrients; functions of nutrients and other dietary factors in growth, reproduction, and health; biomarkers of the nutritional status and disease; and individualized requirements of nutrients. The proteome analysis is expected to play an important role in solving major nutrition-associated problems in humans and animals, such as obesity, diabetes, cardiovascular disease, cancer, aging, and intrauterine fetal retardation.

CHALLENGES IN PROTEOMICS WITH RESPECT TO PLANT PHYSIOLOGY

Although significant advances in the comprehensive profiling, functional analysis, and regulation of proteins has occurred in model organisms such as yeast (Saccharomyces cerevisiae) and in humans, proteomics research in plants has not advanced at the same pace. The availability of the complete Arabidopsis (Arabidopsis thaliana) genome, which is small compared to that of other plants, along with an increasingly comprehensive catalog of protein-coding information from large-scale cDNA sequencing and transcript mapping experiments, set it apart as a complex but accessible model organism to study plant proteomics. The application of proteomic approaches to plants entails three major challenges:

1. Comprehensive identification of proteins, their isoforms, and their prevalence in each tissue.
2. Characterizing the biochemical and cellular functions of each protein.
3. The analysis of protein regulation and its relation to other regulatory networks.

ACKNOWLEDGEMENT

I am are very much thankful to Mr S.Narasa Raju, Chairman of Children's education trust,Mr. Narasimha raju, Executive director of Children's education trust ,Prof K Basavaraju, Dr.T.Krishnan Principal,Dr Kusum Paul of The oxford college of Engineering,Bangalore for their support and encouragement.

References

r-0. Belhajjame, K. et al. Proteome Data Integration:

Characteristics and Challenges. Proceedings of the UK e-Science All Hands Meeting, ISBN 1-904425-53-4, September 2005, Nottingham, UK.
r-1. Twyman, R. M. 2004. Principles of proteomics. BIOS Scientific Publishers, New York. ISBN 1-85996-273-4.(covers almost all branches of proteomics)
r-2. Westermeier, R. and T. Naven. 2002. Proteomics in practice: a laboratory manual of proteome analysis. Wiley-VCH, Weinheim. ISBN 3-527-30354-5.(focused on 2D-gels, good on detail)
r-3. Liebler, D. C. 2002. Introduction to proteomics: tools for the new biology. Humana Press, Totowa, NJ. ISBN 0-585-41879-9 (electronic, on Netlibrary?), ISBN 0-89603-991-9 hardback, ISBN 0-89603-992-7 paperback.
r-4. Wilkins MR, Williams KL, Appel RD, Hochstrasser DF. Proteome research: new frontiers in functional genomics. Berlin Heidelberg, Springer Verlag; 1997, ISBN 3-540-62753-7.
r-5. Arora, Pankaj S., et al. (2005). “Comparative evaluation of two two-dimensional gel electrophoresis image analysis software applications using synovial fluids from patients with joint disease”. Journal of Orthopaedic Science 10 (2): 160–166. doi:10.1007/s00776-004-0878-0.
r-6. Rediscovering Biology Online Textbook. Unit 2 Proteins and Proteomics. 1997-2006.
r-7. Weaver. R.F. Molecular Biology. Third Edition. The McGraw-Hill Companies Inc. 2005. pgs 840-849.
r-8. Campbell and Reece. Biology. Sixth Edition. Pearson Education Inc. 2002. pg 392-393.
r-9. Hye A, Lynham S, Thambisetty M, et al. “ Proteome-based plasma biomarkers for Alzheimer's disease.” Brain 129: 3042-3050, (2006).

Author Information

M. Abhilash, M. B.E, M.Tech
Department of Biotechnology, The Oxford college of Engineering