

TECHNOLOGIES DRIVING PROTEOMICS IN EVOLUTIONARY STUDIES: A CRITICAL OVERVIEW

Saurabh Manish Kapoor

Department of Prosthodontics, Faculty of Dental Sciences, King George's Medical University, Lucknow, India

DOI: <https://doi.org/10.5281/zenodo.17379598>

Abstract

Proteomics is the analysis of the protein composition of the cell, organ, or organism at any given time. The nascent and subtle developments in proteomics have ushered a new attractive and useful gateway to study the whole protein content of a biological sample in one set of experiments. In the past few years, proteomics has rapidly progressed and major advances in this sphere have led to radical researches into new applications like analysis of proteins for clinical diagnostic purposes. Quantifying the protein expression is a crucial part of proteomics research and employs the methods that provide accurate analytical values in protein and peptide mixtures in biological fluids such as serum or plasma. Core technologies applied for the separation of protein and peptide mixtures are one- and two-dimensional gel electrophoresis, chromatography and mass spectrometry coupled with MALDI-TOF and ESI-MS. In this review, we have attempted to describe various approaches established for the analyses of proteins and the highlight the current technologies used for proteome fractionation and separation of biological samples. Exploration of protein analysis has opened new dimensions in the clinical biochemistry and medical field.

Keywords: Chromatography, Data analysis, Electrophoresis, MALDI-TOF, Mass Spectrometry, Proteomics

I. Introduction

Proteins are indeed essence of life. They are initiator, regulator, controller and terminator of virtually all metabolic and physiologic activities. Expectedly, therefore, analysis of proteins has been an integral part of the field of clinical chemistry for decades. In post genomic era, the technological boost towards scientific research has opened the path to study the protein component of the cell and introduced the term "Proteome and Proteomics". Proteome is defined as the total protein component of the cell and the study of the structure and function of these proteins is called Proteomics. In the recent years, the fields of proteomics have been broadly expanded by the concept of "classical proteomics methodology" based on the development in three areas: 2-D Electrophoresis, Mass Spectrometry and Bioinformatics tools. Mounting interest of biochemists has led to prompt developments in the field of proteomics in recent years. Current advancement in proteomics proposes the potential to analyze hundreds of components at a time and to identify proteins and their structural modifications in minuscule specimens. Technologic advances have greatly extended the potential in diagnostic analysis of proteins.

II. Genesis And Background

The study of protein began in 1975 with the mapping of protein from E.coli, mouse and guinea pigs¹. Subsequently, a project entitled "The Human Protein Index" was proposed to identify all the human proteins but due to lack of funding and technical limitations the projects remain uncompleted². The term „proteomics" was first introduced in 1994 and was defined as the large scale characterization of the entire protein component of a cell, tissue or organism³. This field however roots back to the history of clinical chemistry where first laboratory test for a protein cancer marker, the Bence Jones protein in urine, was described in 1847.

Despite the fact that the definitions of proteomics vary from the classical definition (study of proteins only) to a more exclusive definition (study of proteins along with genomic analysis)², the purpose remains the same i.e. to study all the proteins of a cell rather than each one individually. Human genome was mapped long back by 2001 but human proteome is still being deciphered⁴.

III. Need Of Proteomics

Exploration of the protein content of the genome of a biological system has become need of the Biological research because of various reasons. The completion of the "Human Genome Project" suggested that studying protein directly was not necessary as the sequence of amino-acid of a protein can be deduced from its DNA sequence. But the recent advances in proteomics have proved that this is not factual⁵. The dogma that one gene encodes for one protein have been demolished now suggesting the need of proteomics.

There are an estimated 25,000 genes in the human genome, and each gene can yield multiple protein products because of variable splicing and posttranslational processing. Though DNA in each cell type in the body is same but different set of genes are expressed in different tissues and hence the protein component of a cell varies from cell to cell type. Functions of proteins can be inferred from their 3-D structure after the complete identification of all the proteins in a cell. Mislocalization of protein drastically affects their cellular functions. It is now clear that a single gene can encode multiple proteins as a result of post-transcriptional modifications (alternative splicing, polyadenylation and mRNA editing), regulation at the level of translation and post-translational modifications⁶. Hence "one gene one protein is an oversimplification" and this appreciates the need to directly analyze the protein component of the cell⁷.

Earlier proteomics referred to simply analyzing proteins. Over the years proteomics has expanded into profiling, structural and functional analysis including use of a wide range of technologies. Profiling proteomics endeavours to profile the proteins differently expressed between two different samples e.g. comparing samples from a healthy state and a diseased state. Functional proteomics refers to study protein functions on posttranslational modified proteins, or to the interaction of proteins with substrates and small molecules. Finally, structural proteomics highlights the tertiary structure of proteins and protein complexes with small molecules and other proteins.

IV. Goals Of The Proteomics

Proteomics has surfaced as a technology of post-genomic era to accomplish various aspirations. Identification of all proteins in cell along with a 3-D map of the cell indicating the localisation of the protein is the main goal of proteomics. Also various biomarker proteins for detection and diagnosis of diseases besides the design of drug therapy have been targeted through this novel technology⁸.

V. Proteomics Tools

The complexity of protein and peptide mixtures in biological fluids such as serum or plasma has been a big challenge where separation of different components is concerned^{9, 10, 11}. Even the highest resolution techniques for analysis of proteins can separate a maximum of a few thousand components. Cracking the maximum number of protein components depends on multidimensional chromatography or other techniques using multiple fractionation steps 10, 12, and 13,14,15,16.

For the typical proteomics analysis, blood samples necessarily be collected and processed according to a standardized tightly regulated protocol for achieving reproducible records¹⁷. Proteomic analysis should be done on serum or plasma has been a point of debate. Plasma presents many analytical challenges, such as predominance of a few major components, along with a large range of complex components, but the number of proteins that can be detected in plasma has expanded dramatically from hundreds to thousands. It has been suggested in a study that many highly abundant peptides are detectable only in serum samples, and not in plasma¹⁸. The available methods and technology for the identification and quantification of proteins in biological samples can be broadly divided into three categories as described below.

A. Solubilization, Separation and Isolation of Proteins from Cell

(i) Solubilization of protein

The extraction of protein from a cell lysate is a critical step for establishing a stable and reproducible proteome data¹⁷. In proteomic analysis because of in the vast range technology, detection of proteins that are low in abundance necessitates exhaustion of highly abundant proteins like albumin, globulins and others like transferrin, fibrinogen, Apo lipoproteins from the blood sample¹⁹. Yielding a soluble protein needs lysing the cells and collecting the supernatant. There are number of methods available which cover up the majority of the requirements in protein extraction from cells. Hydrophobicity and possibility of developing structural changes due to prolong storage at room temperature tend to offer certain limitations in the current protein extraction method²⁰. Also heterogeneous nature of tissue samples can result in large variation in protein content within the same tissue as well as tissue of different sources. Cultured cells have been the most popular source of proteins for studies of the proteomics as the experimental conditions can be controlled²⁰.

Unlike the breakthrough in the field of molecular biology, protein technology has shown little progress over the years. Protein Electrophoresis still remains the most effective way to resolve a complex mixture of protein. It is at this stage where there is restricted access, as it is a slow and wearisome procedure which cannot be easily automated.

(ii) Separation of the proteins by electrophoresis

Protein separation is the foundation of analysis of proteome. Numerous proteomics techniques are employed to resolve the large range of proteome. Basic technologies involved are one- and two-dimensional gel electrophoresis (1-DE, 2-DE) and chromatography. Coupled together with mass spectrometry (MS), these separation techniques have turned up complexities associated with the protein separation. The protocols used for these techniques have to be gently tuned to get the desirable results. The relevant aspects are discussed below

1. One-dimensional gel electrophoresis (1-DE):

It is method of choice to resolve protein mixture in sodium dodecyl sulphate (SDS) according to its molecular mass²¹. 1-DE is a simple and basic methodology that can resolve protein mixture with molecular weight ranging from 10-300kDa. Limited resolving power is the fundamental drawback involved with this method. To overcome this concern whenever a complex protein mixture is encountered, then two dimensional gel electrophoresis is preferred².

2. Two- dimensional gel electrophoresis (2-DE): It was first introduced in early 1970"s ^{22, 23,24,25,26}. It has been a mainstay of biochemistry and has made important contributions to proteomic analyses. For the large scale purification of protein, two dimensional gel electrophoresis is preferred. Proteins are separated according to their net charge (pI) in first dimension and then according to their molecular weight in second dimension ^{8, 20, 22}. The separation in second dimension is done in a polyacrylamide matrix in a sodium dodecylsulfate (SDS) background. An acrylamide gradient of 10–20% is commonly used followed by visualization in 2-D gels by various detection methods. For protein staining many protocols are exploited²⁷, whereby in function Coomassie Brilliant Blue and silver staining have been extensively accepted. Initially it was very tedious and irreproducible technique but with the introduction of immobilized pH gradient gels the practical technical problems were resolved²⁶. This methodology is a well-accepted protein separation technique, but because of solubility and mass of the separated proteins, there are a few restrictions. Large hydrophobic proteins cannot penetrate the gel, thus limiting the identification of less abundant proteins when a total cell lysate is analyzed as rich proteins take a lead in the gel. Also post translational modifications of proteins can significantly affect the observed molecular weight and isoelectric point (pI) which can offer advantage in comparing protein expression of two samples both qualitatively and quantitatively.

(iii) Detection of proteins separated by 2-D gel electrophoresis

Over the years a number of methods have been developed to visualize proteins, separated by gel electrophoresis. Coomassie blue and silver staining have been the method of choice because of their ease of use and sensitivity. The use of fluorescence staining (e.g Cy dyes, LAVAPurple, Sypro dyes), radiolabeling and immunodetection provide a superior sensitivity and a broad dynamic range as compared to the standard methods^{20, 28}. The 2D gel can display upto 2000 spots from the blood or tissue extracts which signifies different proteins amounting to approximately 100–300 in number. The spot is liable to shift on

the pH gradient and molecular weight scale depending on post-translational modifications, sample composition etc. In the present time 2-D gel images are evaluated and analyzed using specialized software packages eg. Melanic package from Swiss Institute of Bioinformatics, Phoretix 2D software from Phoretix, Gellab II from scanalytics etc. which can store all of the relevant information on all of the spots of a 2-D gel in a database²⁰.

(iv) Isolation of protein

The information contained in a gel is not total hence the additional information has to be extracted via blot digestion of proteins and In-gel digestion of proteins.

Blot digestion of proteins (Electrotransfer / Electroblothing): After separation of proteins by 1-DE/2DE, proteins are separated from the gel by the technique electro blotting, which was specifically designed for this purpose. The proteins are transferred from gel to a binding membrane (nitrocellulose membrane) under the influence of electric field. The protein trapped in the gel migrates and encounters binding nitrocellulose membrane to which it gets attached. The blot can be visualized by staining. On the other hand In-gel digestion of proteins, first introduced by Wilm et al in 1996, is the method of choice for generation of peptides from proteins separated on the gel. The protein spot of interest is digested in gel with trypsin after excising it from the gel followed by extraction to obtain the peptides^{2, 20, 29}.

VI. Chromatographic techniques

Separation of proteins by chromatography has been a bastion of biochemistry and has contributed significantly towards proteomic analyses. Chromatography is one of the most potent methods of separation and employs one or more inherent characteristics of a protein-its mass, isoelectric point, hydrophobicity or biospecificity. The well recognized separation methods based on the physical characteristics of proteins, comprise of ion exchange, size exclusion, reversed-phase high-performance liquid chromatography (HPLC)^{30,31,32}, capillary IEF, and capillary zone electrophoresis^{33,34,35,36}. Other methods that are extremely selective include affinity chromatography that uses protein affinity for antibodies, specific target proteins, or chemical moieties^{37,38}, and liquid chromatography, coupled with mass spectrometry (MS). These techniques digest the sample first and separate the peptides for identification.

Capillary electrophoresis coupled to MS (CE/MS) provides high separation efficiency in small volumes^{39, 40, 41}. Proteins are resolved according to their elution time from the CE and the mass of the protein in the MS and have been exploited in the identification of polypeptide biomarkers in urine samples⁴². For interface with electrospray ionization (ESI) (discussed later), high-performance liquid chromatography (HPLC) is commonly applied for separation. A liquid interface allows for coupling the separation step directly to the ionization source of the mass spectrometer. In liquid chromatography/MS (LC/MS), digested peptides are separated first by liquid chromatography and then injected directly into the mass spectrometer. LC/MS can identify low-abundance and hydrophobic proteins which cannot be identified by 2DE. Use of isotopic tags has even made quantification of proteins possible with the use of LC/MS⁴³.

B. Protein Identification and Characterization

(i) Edman degradation

Edman and Begg developed a technique for the N-terminal sequencing of proteins in early 1960's⁴⁴. This technique allows chemical degradation (phenyl isothiocyanate) of protein into individual amino-acid in a cycle-dependant manner from the N-termini of the proteins, typically upto 20 amino-acids^{2, 4}. It is a very chaotic and extensively laborious approach to sequence a protein. Currently due to availability of protein and DNA database protein identification by using N-terminal amino acid sequencing became obsolete.

(ii) Mass Spectrometry (MS)

Mass spectrometers are instruments that can produce and separate ions according to their mass-to-charge ratio (m/z). MS techniques have significantly advanced the proteomics-based research in recent years. The idea of Mass spectrometry was established nearly a century ago and now it is widely applied as an analytical technique⁴⁵. MS is highly sensitive and the molecular weight (MW) or structural information such as peptide masses or amino acid sequences of a compound can be achieved in a short time. Though MS was found to be a tedious and slow process and required large volume of samples but over the years, the idea of using MS for identification of proteins has evolved with the improvement in instrumentation⁴⁶. MS instrumentation separates and detects the ions in gas phase and therefore, prior to any separation by MS, molecules must be ionized and converted into gas using different techniques. Development of two new ionization techniques Electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) in the late 1980s simplified the acquisition of mass spectra with minuscule quantities of peptides and proteins^{47, 48} thus making it possible to apply MS to the analysis of proteins. The introduction of Matrix assisted laser-desorption-ionization time-offlight (MALDI-TOF) and electron spray ionization tandem mass spectrometry (ESI-tandem MS) has revolutionized the field of mass spectrometry^{49, 50, 51}.

2-DE coupled with MS is widely used for protein analysis. The protein spots from 2-D gels are excised and subjected to in-gel digestion followed by identification of resulting peptide fragments by MS¹⁹. For identification of 2-D gel spots MALDI time of flight (TOF) MS is commonly used⁵².

MS has three essential parts (i) the ion source that will produce ions from the sample (ii) the mass analyzer to resolve ions based on mass/charge ratio (m/z) (iii) the detector to detect the ions resolved by the mass analyzer as shown in Fig.3. During separation electric or magnetic fields are generated within the instrument. Ion separation is necessarily done under vacuum to avoid the collision between accelerated ions and air molecules. The analyser and the detector also must be under vacuum, although ion production can also be done at atmospheric pressure depending on the type of ionization source.

By using MS protein structural information like peptide mass, amino-acid sequence can be obtained. The analyser is the part of the MS where separation of ions in gas phase takes place. Mass spectrometers known as tandem mass spectrometers may have two or more mass analysers coupled together. Analysers like TOF, quadruples and ion traps (ITs) are used either alone or in tandem in proteomics applications. In ESI, ions are formed at atmospheric pressure, while ions in MALDI they may be generated either at atmospheric

pressure⁵³ or under vacuum conditions, although the best performance is obtained when working at low pressure. The three techniques of MS presently in vogue are discussed as under

1. Matrix assisted laser-desorption-ionization time-of-flight (MALDI-TOF MS)

In the field of MS, MALDI-TOF allowed rapid and accurate process to measure mass of analytes. **MALDI** is used to change the peptide from a solid phase to the gaseous phase. It is a competitive process where the ionization of an analyte may be inhibited considerably by the existence of others⁵⁴. In MALDI, protein or peptide sample mix is allowed to combine with the matrix molecule in solution and small amounts of the mixture are deposited on a surface and allowed to dry. Here a matrix compound that is capable of absorbing ultraviolet (UV) light is employed⁴⁸. In this method of ionization, sample is incorporated into matrix molecules and then subjected to irradiation by a laser beam^{46, 55}. Matrix is typically made of small energy absorbing molecules such as 2, 5 dihydroxybenzoic acid or α -cyano-4-hydroxycinnamic acid and sinapinic⁴². Matrices like ice can also be used for peptides and proteins⁵⁷. The analytes along with matrix are mixed in the appropriate solvent and spotted on a sample probe that can measure upto dozens to hundreds of samples. The solvent is allowed to evaporate resulting in crystal formation. The plate is placed in MS at high vacuum and a laser beam of a specific wavelength is targeted onto a limited area of the spot. The matrix molecules absorb photons from the beam and become excited. The excess energy is then transferred to the peptides in the sample, which are then ejected into gas phase⁵⁴. Each peptide molecule takes up single proton hence peptide ions are singly charged^{2, 20}. A high voltage (+20V to +30V) is applied on MALDI plate to induce positively charged peptides to accelerate toward the orifice of the flight tube.

TOF (Time-of-flight): It necessarily comprises of a flight tube in high vacuum. It is simplest mass analyzer which measure m/z ratio of an ion by determining the time required for it to traverse the length of flight tube². The vacuum in the flight tube is such that the chance of molecule collision with each other or with the flight tube is very low. As the peptides generated in MALDI are equally charged, they traverse flight tube with different velocities, which are inversely proportional to their masses⁵⁸ and hit the detector at different time intervals⁴. Lighter ions land at the detector faster than the higher ones separating the ions of different masses. To improve the resolution, reflectron a piece of hardware is placed in the path of ions⁵⁹. Principle application of MALDI-TOF-MS is peptide mass fingerprinting and large scale proteomics work. The technique is applicable not only for peptides and proteins, but for other bio-molecules like polysaccharides, lipids and polynucleotides^{60, 61, 62, 63}. It is a user friendly, fast, highly sensitive and accurate technique.

2. Surface-enhanced laser desorption/ionization (SELDI)

SELDI, an extremely versatile and convenient proteomic tool, is another method used for the analysis of protein mixtures⁶⁴. SELDI technology was developed by T. William Hutchens in 1993⁶⁵. It is an ionization method in mass spectrometry that is typically used with time-of-flight mass spectrometers to detect proteins in tissue samples, blood, urine, or other clinical samples for biomarker discovery⁶⁶. SELDI-TOF-MS is uses a target that is modified to achieve biochemical affinity with the analyte compound. SELDI is a

variation or rather an extension of MALDI where the protein solutions are applied to surface modified with a chemical functionality. It measures the intensity of peaks from the subset of proteins captured on the plate and the peak height differences between samples correlate with relative abundance in the sample. The proteins actively interact and some proteins bind to the surface according to their surface interaction potential, while the sample contaminants are removed by washing. After washing the spotted sample, the matrix is applied to the surface and target protein and matrix are allowed to cocrystallize. Binding to the SELDI surface acts as a separation step simplifying the analysis of the subset of proteins that bind to the surface.

The differences between SELDI and MALDI are in the assembly of the sample targets, the design of the analyzer and the software tools used to interpret the acquired data.

3. **Electron spray ionization tandem mass spectrometry (ESI-tandem MS)**

ESI is a soft ionization method which allows analyte to transfer from liquid phase to the gas phase at atmospheric pressure. In this method a high voltage is applied to a liquid flowing through a micro capillary tube into the orifice of MS. The electric field forces charged liquid to form a cone of droplets. As these droplets move toward mass analyzer they undergo a process of solvent evaporation^{46,47}. One of the advantages of electrospray is that enables ions to acquire multiple charges depending on their molecular mass and structure.

4. **Tandem mass spectrometry (MS/MS):** The identification of peptides by MS/MS has substituted almost completely the more time consuming and extensively painstaking method of Edman degradation. In the first stage of analysis, machine is operated in MS mode and all the ions of certain m/z ratio are transmitted to the third quadrupole (Q3) for mass analysis⁴⁶. In the second stage, machine is operated in MS/ MS mode in which a particular peptide ion is selectively passed into the collision chamber. Inside the collision chamber, peptide ions are fragmented by interaction with inert gas (collision induced dissociation)⁶⁷. The peptide fragments are then resolved on the basis of their m/z ratio by third quadrupole (Q3). Since two different mass spectra are used in this analysis, it is referred as tandem mass spectrometry^{68, 69}.

C. **Database Utilization**

All the mass spectra obtained by mass spectrometer are written to a file or loaded into a database and are then further analyzed, using the bioinformatics tools⁷⁰. The goal of analysis of proteins/peptides by database searching is to identify a large number of proteins in a short span of time. Three types of database searching are: **(i) Peptide mass fingerprinting database searching:**

Peptide mass fingerprinting is a type of protein identification method where the analysis and database search is fully automated^{71, 72, 73}. In this method, mass of each peptide obtained from the proteolytic digestion of an unknown protein is matched against the predicted mass of the peptide from the theoretical digestion of proteins obtained from the in silico digestion at the same enzyme cleavage sites of all protein amino acid sequences in a database⁷⁴. If they overlap, protein identification can be done. The proteins in

the database are positioned depending upon their peptide masses matching their sequence within a given mass error tolerance. Programs used for protein identification are PepSea⁷¹, MS-Fit⁷⁵, PeptIdent/MultiIdent⁷⁶ and ProFound⁷⁷. The biggest advantage of database searching is rapidity and that it can be fully automated. However there are a few limitations also associated with this technique. First is the ambiguity in protein identification i.e. a peptide with ten amino acid will carry same mass by simple reorganization of its constitutive amino acids. Secondly it is effective for analysis of protein from organism whose genome is small and completely sequenced⁷⁸. Another important factor is mass accuracy i.e. if an unknown protein is extensively modified by post-translational modifications, the peptides produced from the proteins will not match with the unmodified proteins in database.

(ii) **Peptide amino acid sequencing database searching:** This is the most specific type of database searching for protein identification. By determining the amino acid sequence of a peptide, the information about protein from which it is originated can be obtained by search database. One such method based on this theory is peptide mass tag searching where peptide mass together with a short sequence produced by partial interpretation of the spectrum is used for database searching^{2, 79, 80}. For protein identification, peptide mass tag searching is more specific and can be used for protein mixtures in contrast to peptide mass fingerprinting^{81, 82}. The major disadvantage of this procedure is that the process is not easily automated hence becomes time consuming.

(iii) **De novo sequencing**

De novo peptide sequencing is independent of any information present in databases. When protein identification remains intangible with other methods, de novo peptide sequencing is the only option to acquire information about the analyte. The sequence of a peptide can be reconstituted from its fragmentation spectrum⁸³.

VII. Clinical Applications of Proteomics

The proteome research aims to recognize multifarious responses to various stimuli. . Clinical proteomics is defined as the application of proteomics in the field of medicine. It plays an important role in diagnosis and the management of the patients^{19, 84,85,86,87}. Proteomics research has opened the path to investigate various body fluids or cells to aid in diagnosing various diseases by identifying and quantifying the protein module. Testing large quantity of proteins in a single analysis by proteomic proficiency has enabled the researchers to recognise newer biomarkers for wide spectra of diseases.

The extensive intricacy along with immense range of proteins involved in all biological fluids, particularly serum and plasma, has led to development of technologies that have achieved high-resolution and high-sensitivity proteomics analysis.

Proteomic mass spectrometry coupled with bioinformatic tools provides important means for detection of new biomarkers in serum for a variety of human disorders. With newer development of modalities of improved protein analysis, detection of specific disease-induced modified forms of proteins like cardiac troponins is paving ways to better and more specific diagnoses and therapies^{88,89}. Analyses of serum or

plasma concentrations of other proteins and their different molecular forms such as myoglobin and creatine kinase MB isoenzyme might lead to identifying additional proteins that may serve as biomarkers for heart disease. In search of prospective novel biomarkers useful in diagnosis and prognosis of stroke patients, nucleoside diphosphate kinase A and PARK7 have been recognised as proteins released by brain injury in the cerebrospinal fluid (CSF) of stroke patients. Investigation of postmortem CSF proteins has led to discovery of potential new biomarkers of ischemic brain injury⁹⁰.

Protein expression profiling on patients of breast cancer identified role of breast cancer biomarkers (BC3 and BC2) in breast cancer [91]. Proteomic signatures of sera in cases of sleep apnea have been identified through proteomic strategies. This has might aid in development of new serum-based diagnostic biomarkers of anoxia⁹². Application of proteomics has revolutionised the field of inborn errors of metabolism. Tandem mass spectrometry (MS/MS) has become an important tool in population-based screening for IEMs^{93, 94}. Many screening programs may adopt MS/MS screening as a routine method for population-based screening for aminoacidopathies such as PKU. It is rapid, highly sensitive and specific, and amenable to automation and can be adapted to the Guthrie collection protocols in use by most screening programs ^{95, 96, 97, 98, 99}.

VIII. Corollary

Proteomics has surfaced as a postgenomic technology which has made possible to study proteins on large scale with wide basic and clinical applications. Over the past decades, protein identification and quantification of complex biological samples has improved significantly. Specialized instruments with increased resolution and scan speed provide enhanced analytical capabilities and better diagnostic approach for certain types of analyses. The two main approaches being exploited to fullest are MALDI-time-of-flight (TOF)-MS and liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) procedures to obtain peptide fingerprinting profiles. Proteomics has enabled the researchers and clinicians to a vast variety of aspects like discovery of molecular markers, novel biologically active molecules, drugs and study of basic cell function and molecular organization. Evaluation for both quantitative and qualitative analysis may open new vistas of territory in the clinical laboratory and medical field. Even today, the study of proteomics is a big challenge because of various reasons. For example the methods used to isolate and characterize the proteome are still not universally established due to certain limitations and restrictions. 2-DE still lacks the capabilities for identifying hydrophobic proteins, basic proteins and low copy proteins. Secondly, the methods involved in protein purification, quantification of proteins from a complex mixture need to be improved. Third, as a result of post translational modifications functions of many proteins are altered without change in their concentration. Comprehensive attention is needed to explicate the biologic behaviour and conduct of these post-translationally modified proteins for complete and concrete understanding of the precise mechanism of their action in cellular milieu. Fourth, analysis of low-abundance protein is difficult as there is no equivalent of PCR for proteins. Fifth, problems about sophisticated software used for data analysis still continue. For

integration of results from different sections and interpretation of large data sets simple and uncomplicated software are required. These disadvantages need to be overcome in the field of proteomics to improve the sensitivity of techniques. Thus, these areas provide platform on which future studies should be formulated and executed.

References

- J.M Walker . Protein structure, purification, characterization and function analysis In Principles and techniques of Biochemistry and Molecular Biology 6th ed. (Cambridge University press) 2005, 392-94.
- R.G Krishna ,F. Wold . Post-translational modifications of proteins, Adv Enzymol Relat Areas Mol Biol 1993; 67: 265-98.
- A.Pandey . Proteomics to study genes and genomes, Nature, 405, 2000, 837-46.
- S.B Giorgianni . Proteome analysis by two-dimensional gel electrophoresis and mass spectrometry: strengths and limitations., Tr Anal Chem, 22, 2003, 273-81.
- N.L Anderson , N.G Anderson .The human plasma proteome: history, character, and diagnostic prospects,Mol Cell Proteomics ,1,2002, 845-67.
- P. Lescuyer ,D.F Hochstrasser ,J.C Sanchez . Comprehensive proteome analysis by chromatographic protein prefractionation, Electrophoresis, 25, 2004, 1125-35.
- G.S Omenn ,D.J States ,M. Adamski ,T.W Blackwell ,R. Menon ,H. Hermjako, et al. Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database, Proteomics,5, 2005,3226-45.
- Z.Yang ,W.S Hancock ,T.R Chew ,L. Bonilla . A study of glycoproteins in human serum and plasma reference standards (HUPO) using multilectin affinity chromatography couple with RPLC-MS/MS, Proteomics,5, 2005,3353-66.
- N.L Anderson, M. Polanski, R. Pieper , T.Gatlin ,R.S Tirumalai ,T.P Conrads , et al. The human plasma proteome: a nonredundant list developed by combination of four separate sources, Mol Cell Proteomics,3, 2004,311326.
- S Sheng , D Chen, J.E Van Eyk . Multidimensional liquid chromatography separation of intact proteins by chromatographic focusing and reversed phase of the human serum proteome: optimization and protein database, Mol Cell Proteomics,5, 2006,2634.

H.Wang ,S.G Clouthier ,V. Galchev ,D.E Misek ,U.Duffner ,C.K Min , et al. Intact-protein-based high-resolution threedimensional quantitative analysis system for proteome profiling of biological fluids, Mol Cell Proteomics,4,2005, 618-25.

A.M. Rahbar, C. Fenselau . Unbiased examination of changes in plasma membrane proteins in drug resistant cancer cells, J Proteome Res, 4, 2005, 2148-53.

Q.Fu, J.E. Van Eyk . Proteomics and heart disease: identifying biomarkers of clinical utility: Expert Rev, Proteomics, 3, 2006, 237-49.