MINI-REVIEW

A Review of Proteomics in Cancer Research

Prabhudas S Patel[1], Shaila D Telang[2], Rakesh M Rawal, Manisha H Shah

Abstract

The most functional compartment encoded by the genome is proteome. Therefore study of proteome i.e. proteomics is the promising approach in identification, separation and quantitation of functional changes. It aims to gain a comprehensive understanding of the expressions, modifications, interactions, and regulation of proteins in cells. The power of two-dimensional electrophoresis and advances in mass spectrometric techniques, combined with sequence data base correlation, has enabled speed and accuracy in identification of proteins in complex mixtures. Therefore, proteomics may provide a better understanding of the molecular basis of cancer growth, with identification of potential pathological markers and therapeutic targets. Tobacco related cancers are the major health hazard in Asian countries; the proteomics approach should be employed for understanding the underlying disease processes and hopefully reveal important clues for identifying high risk individuals and early changes during malignant transformation.

-Asian Pacific J Cancer Prev, **6**, 113-117

Introduction

The human genome project has provided a wealth of information about the sequences of more than 30,000 genes. The biggest challenge for biomedical researchers is to assign molecular and cellular functions to the thousands of newly predicted gene products and to explain how these products regulate complex physiological processes. To address these problems, the field of proteomics has emerged with the goals of developing and implementing tools for global analysis of the proteome. Therefore, in the post-genome era the focus of biological research is moving towards proteomics i.e. identifying the structure, function and interaction of the proteins produced by these genes and their role in specific disease processes, including cancer. Recent reports have documented noteworthy clinical significance of proteomics studies in various malignancies (Table-1). Coupled with transcript profiling, proteomics may herald the advent of molecular therapy tailored to the individual patient's neoplasm. Protein microarrays, emerging class of proteomic technologies, have broad applications for discovery and quantitative analysis.

Onco-proteomics can be mainly divided into:

*1. Structural/Expression Proteomics***:** This is basically associated with monitoring the global expression of large number of proteins that are differentially displayed within a given cell type (Anderson and Seilhammer, 1997) tissue or body fluids and the patterns of expression are quantitatively identified in different disease states.

2. Functional/Cell Mapping Proteomics: This approach is basically associated with the analysis of protein-protein or protein-DNA / RNA interactions and post-translation modifications to build a picture of the complex networks that constitute intracellular signaling pathways (Humphery - Smith et al., 1997; Gygi etal., 1999 ; Black Stock and Weir, 1999)

Techniques of Proteomics

Analytical techniques for proteome analysis use a combination of sophisticated techniques including 2D Gel electrophoresis, image analysis, mass spectrometry, amino acid sequencing and bio-informatic tools to quantify and to characterize complex proteins.

Sample Preparation: Appropriate sample preparation is the first absolutely essential step for obtaining good 2-D results.

Two Dimensional Poly-Acrylamide Gel Electrophoresis (2-D PAGE): This technique sorts proteins according to two independent properties in two discrete steps: the first dimensional step i.e. Iso-Electric Focussing (IEF) separates proteins according to their Iso-electric points (pI) or charge

[1]Biochemistry Research Division, The Gujarat Cancer & Research Institute, [2]Biochemistry Department M. S. University of Baroda, Vadodara-390 002. Reprint Requests and Correspondence to: Dr. Prabhudas S. Patel, Head, Biochemistry Research Division, The Gujarat Cancer & Research Institute, Asarwa, Ahmedabad – 380 016, INDIA Phone: +91 79 22681451 (Extn303) Fax: +91 79 22685490, e-mail : prabhudas_p@hotmail.com

Table 1. Recent Studies of Cancer Biomarkers Using Proteomics Approaches

No.	Aim (Reference.)	Proteomic techniques	Utiilty
1	Identification of proteins crucial for initiation of apoptosis in NCOL-I pre-neoplastic colonocytes (Angelika et al., 2004)	Proteome analysis of cells exposed to flavonoids, flavone and quercetin, using 2D-PAGE and peptide mass fingerprinting differs in their ability to induce apoptosis	To identify markers involved in initiation of apoptosis, to develop new strategies for cancer prevention
2	Protein expression profile (Dwek and Alaiya, 2003)	Discrimination between normal, benign breast tissue and breast cancer tissue using 2D-PAGE cluster analysis	Early diagnosis and classification of breast cancer
3	Determination of protein expression profiles in kidney cancer (Thomas and Manfred, 2004)	Hierarchical cluster image analysis of protein patterns between drug-sensitive and drug-resistant kidney carcinoma cell culture	Diagnosis, therapy and treatment outcome
4	Breast cancer-specific antigenic marker (Ryuji, 2001)	2D PAGE and Western blot analysis of sera of breast cancer patients for identifying circulating tumour antigens.	Rapid identification of proteins involved in cancers.
5	Protein expression in gastric cancer (Jin-Woo et al., 2003)	Identification of functioning proteins using 2D-PAGE and MALDI-TOF	More effective therapy to gastric cancer patients
6	Detection of prostate cancer using serum proteomics pattern in histologically confirmed population (Carroll, 2004)	Study of protein markers using protein chip array and MALDI-TOF	Serum proteomics patterns may potentially aid in the early detection
7	Protein profiles in Kidney cancer tissue and their primary cell cultures (Dong et al., 2003)	Alterations of protein expression studied using 2D PAGE.	As an alternative method when high quality 2-D PAGE from tissue samples is difficult.
8	Review of proteomics for haematologic malignancies. (Rees-Unwin et al., 2004)	2D-PAGE and MALDI-TOF.	This technique may facilitate the development of a rapid diagnostic test applicable in haematologic malignancies.
9	Determination of metabolic protein expression in different tumor cell lines (Gruber- olipitz et al., 2004)	Tumor cell line profiling studied using 2D-PAGE followed by MALDI-TOF.	Understanding of metabolic pathways for tumor manifestation, progression, metastasis, drug and radiation resistance.
10	Recent advances in understanding carcinogenicity of oral squamous cell carcinoma: from basic molecular biology to latest genomic and proteomic findings (Mehrotra et al., 2004)	2D-PAGE and MALDI-TOF/SALDI-TOF	Identification of potential molecular biomarkers for prevention, diagnosis and therapeutic monitoring and novel molecular targets involved in signal pathway dominance.
11	Proteomics analysis for early detection and rational treatment of cancer – realistic hope? (Posadas et al., 2005)	Laser capture microdisection, tissue lysate arrays and mass spectrometry.	Identification of newer biomarkers for the early detection of cancer and promise hope of new serological screening methods for diagnosis.

(O' Farrell, 1975; Klose, 1975) The second dimensional step, SDS-Polyacrylamide gel electrophoresis (SDS-PAGE), separates proteins according to their molecular weight (Gorg et al., 1985; Gorg et al., 1988). Each spot resulting on the two dimensional array corresponds to a single protein species in the sample. Thousands of different proteins can thus be separated and information such as the protein pI, the apparent molecular weight and the amount of each protein is obtained. 2-D PAGE ability to separate thousands of proteins and to detect post and co-translational modifications, analysis of cell differentiation, detection of disease markers, monitoring therapies, drug discovery, purity checks and micro scale

protein purification in cancer research.

Protein Visualization: Proteins are commonly visualized in gels by staining them with dyes such as colloidal Coomassie brilliant blue (CBB) G, silver stain and fluorescent stain (Gygi et al., 2000). Coomassie brilliant blue stain, detects in greater than 100ng of proteins. For greater sensitivity silver staining is used, which can detect proteins up to 2.5ng range. However, care must be taken with silver staining to avoid cross-linking reagents such as glutaraldehyde (Corthals et al., 2000). Because the use of aldehydes makes the proteins less susceptible to protease digestion and reduce the efficiency of subsequent mass spectrometry. The recent developments of new fluorescent dyes such as the SYPRO series of protein dyes have improved the sensitivity of protein detection without compromising mass spectrometric analysis. These dyes are effective for staining 2-D gels and offer a similar level of sensitivity to silver staining (Shevehenko et al., 1996; Steinberg et al., 1996a).

Database Search

The protein expression patterns generated by separation techniques are usually complex and are best analysed by computer based image analysis. In this way, it is possible to create 2-D databases. 2-D databases of a number of specific human tissue types and tumours have been developed. One of the best established databases is the SWISS-2D database (Steinberg et al., 1996b). A key part of many Proteomic studies, involves the comparison of protein expression patterns between normal and abnormal tissues. To facilitate such comparisons, it is necessary to use specialized image analysis software, like ELISE, HERMeS, MELANIE-I, II, III, etc., which vary in the way their databases are constructed and organized. MELANIE - II, is one of the latest and most user friendly 2-D gel analysis system and can be linked via Internet to Proteome databases available on the worldwide web (Appel et al., 1993).

Mass-Spectrometry Analysis

Two types of Mass Spectrometers are used for most of the work in proteomics:

1. MALDI-TOF-MS (Matrix - Assisted Laser Desorption Ionization Time Of Flight Mass Spectrometers) In this the proteins or peptides are mixed with a chemical matrix, (e.g.: alpha-cyano-4-hydroxy cinnamic acid) and deposited on a sample plate.(Vestling and Feneslau, 1994)

2. ESI (ElectroSpray Ionization Tandem Mass Spectrometers) - It perform two stages (or multistages) mass analysis of ions. Several different types of mass analysers are used in ESI-tandem-MS, including TOF (most important), ion trap and quadrupole. ESI-tandem-MS is ideal for the analysis of both, larger bio-macromolecules and smaller molecules, because it allows fine structural information to be obtained directly from the biomolecules (Chalmers and Gaskell, 2000; Yates et al., 1999; Yates et al., 1997).

Protein Identification and validation

After mass spectral analysis, the protein is identified by matching the experimental peptide masses to the corresponding peptide mass fingerprints found in the protein or nucleotide sequence databases. Most commonly used protein sequence databases include the SWISSPROT, OWL and NCBI databases, which are publicly available. Several software programmes for protein identification are also available on-line which include:

MOWSE: http://www.srs.hgmp.mrc.ac.uk/cgi-bin/mowse
MASCOT: http://www.matrixscience.com
PeptIdent-2: http://www.prowl.rockefeller.edu/cgi-bin/profound
Profound: http://www.proteometrics.com/procol-cgi/profound.exe

Thus, beyond identification, confirmation and validation of the protein of interest are also important steps in proteomics approaches. There are several ways to confirm the identified proteins, e.g. (i) generating an additional fingerprint with a protease of a different specificity (ii) amino acid sequencing by Edmann's method (iii) amino acid composition (iv) western blot analysis and (v) immunohistochemical analysis.

Role in Tobacco Related Cancers

In Asian countries the burden of tobacco related cancer is mounting day by day. It will be particularly important to perform proteomics experiments with specific populations of cells. It will generate significant information about cell-specific protein expression profiles like structure, functions and regulations and its role in different stages of carcinogenesis. Both global-expression proteomics and in particular, cell-mapping proteomics will contribute significantly in early identification of high risk phenotype. Proteomics approaches to tumor marker identification in these tumour types hold the promise of identifying specific protein modifications in tumor tissues to assist in individualizing treatment and monitoring the treatment response (Figure-1).

Future Developments in Proteomics

Proteomics is still developing and will have a major influence in the future. Much research is being done to improve the efficiency of proteomics techniques, especially with regards to automation, sample throughput, sensitivity and methods for obtaining cell population. Newer approaches to develop protein chips, which can considered to be analogous to DNA micro-arrays are undertaken in a big way. The chips may however in future be replaced with nano-coded metal rods in solution. Proteomics in future will also show the improvement in the quality of the capture

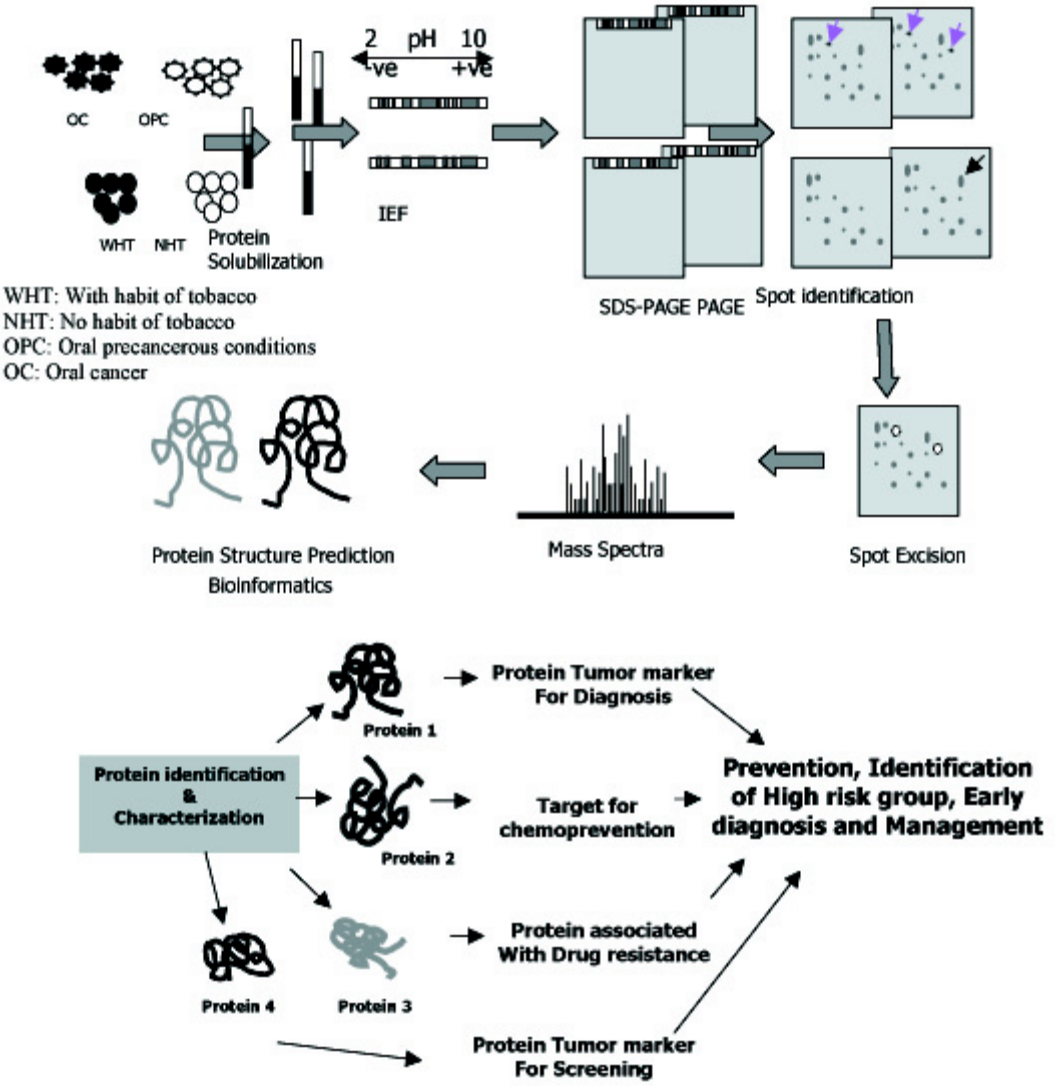

Figure 1. Schematic Representation for Onco Proteomics

reagents contained on the protein-arrays, as well as the methods used to detect bound proteins. Reagents that employ isotope-coded affinity tags (ICAT) are now commercially available for high-throughput proteome quantitation. Also in future, image analysis software for protein expression profiles is becoming ever more sophisticated, with more emphasis on computerized rather than human interpretation. Robots that can automatically cut proteins from gels are now commercially available. The proteins can then be digested and prepared for MS analysis under full automation. Development in the area of modern MS especially in its automation will help in high throughput and better resolution of proteins (Wolfkuhle et al., 2001). Further, a parallel development in the area of Bioinformatics will play a key role in database analysis, protein identification and development of protein interaction maps. Also in future, we will see the increase in development of new methods of cancer classification, which more accurately reflect clinical behaviour and provide the oncologist with the information to tailor therapy to individual patients. It also seems probable that Proteomics will get a central place in the research, understanding, diagnosis, monitoring and treatment of (pre) cancers of many different sites.

Conclusions

Proteomics will greatly contribute to our understanding of gene function in the post-genomic era. Genomics still remains an important approach but the value of proteomics lies in the fact that most of the diagnostics and drugs are targeted to proteins. It seems that the major outcome of the human genome sequencing has finally been to open the way to the exploration of the proteome, transferring the goals, the hopes and the difficulties or pitfalls to this. This is clearly a challenging but also a promising heritage.

In the field of cancer research proteomics has declared hopes in identification of tumour markers for early detection, classification and prognostication of human tumors as well

as pinpointing therapeutic targets for improved treatment outcomes. New high-throughput technologies, such as protein chips coupled with classic technologies such as 2D PAGE, laser capture micro dissection and mass spectrometry, provide cancer researchers with powerful tools for identifying high risk group, early detection and malignant phenotypes as well as identification of targets for therapy. The application of established and novel proteomic technologies to the molecular analysis of cancer has a boundless future.

References

Anderson I, Seilhammer J (1997). A comparison of selected mRNA and protein abundances in human Liver. *Electrophoresis*, **18**, 533-7.

Angelika Herzog, Sabine Kuntz, Hannelore Daniel, Uwe Wenzel (2004). Identification of biomarkers for the initiation of apoptosis in human preneoplastic colonocytes by proteome analysis. *Int J Cancer*, **109**, 220-9.

Appel RD, Sanchez JC, Bairoch A, et al (1993). SWISS - 2D PAGE: a database of two dimensional gel images. *Electrophoresis*, **14**, 1232-8.

Black Stock WP, Weir MP (1999). Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol*, **17**, 121-7.

Carroll PR (2004). Derection of prostate cancer using serum proteomics pattern in a histologically confirmed population. *Urol Oncol*, **22**, 497-8.

Chalmers MJ, Gaskell SJ (2000). Advances in mass spectrometry for proteome analysis. *Curr Opin Biotechnol*, **11**, 384 - 90.

Corthals GL, Wasinger VC, Hochstrasser DF, Sanchez JC (2000). The dynamic range of protein expression; a challenge for proteomic research. *Electrophoresis*, **21**, 1104-15.

Dong Shi T, Liou F, Daun ZH, Novick AA, DiDonato JA (2003). Comparisons of protein profiles in Kidney cancer tissue and their primary cell cultures. *Proc Virt Conf Genom and Bioinf*, **3**, 2-4.

Dwek MV, Alaiya AA (2003). Proteome analysis enables separate clustering of normal breast, benigh breast and breast cancer tissues. *Br J Cancer* , **99**, 305-7.

Gorg A, Postel W, Gunther S, Weser J (1985). Improved horizontal two-dimensional electrophoresis with hybrid isoelectric focusing in immobilized PH gradients in the first-dimension and laying-on transfer to the second dimension. *Electrophoresis*, **6**, 599-604.

Gorg A, Postel W, Gunther S (1988). The current state of two dimensional electrophoresis with immobilized PH gradients. *Electrophoresis*, **9**, 531-46.

Gruber- olipitz M, Afjehi-sadat, L, Felizardo M, Slavc I, Lubec G (2004). Proteomic determination of metabolic protein expression in ten different tumor cell lines. *Cancer Genomics & Proteomics*, **1**, 311- 38.

Gygi SP, et al (1999). Corrrelation between protein and mRNA abundance in yeast. *Mol cell Biol*, **19**, 1720-30.

Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R (2000). Evaluation of two - dimensional gel electrophoresis - based proteome analysis technology. *Proc Natl Acad Sci USA*, **97**, 9390-5.

Humphery - Smith I, Cordwell SI, Back Steck WP (1997). Proteome research; complementarity and Limitations with respect to the RNA and DNA worlds. *Electrophoresis*, **18**, 1217-42.

Jin-Woo Ryu, Hyung-Jee Kim, Young- Sun Lee, et al (2003). The proteomics approach to find Biomarkers in gastric cancer. *J Korean Med Sci*, **18**, 505-9.

Klose J (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutation in mammals. *Humangenetik*, **26**, 231-43.

Mehrotra Ravi, Vasstrand Endre Normann, Ibrahim Salah Osman (2004). *Cancer Genomics & Proteomics*, **1**, 283-94.

O' Farrell PH (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*, **250**, 4007-21.

Posadas EM, Simpkins F, Liotta LA, Macdonald C, Kohn EC (2005). Proteomic analysis for the early detection and rational treatment of cancer-realistic hope? *Annals of Oncology*, **16**, 16-22.

Rees-Unwin KS, Morgan GJ, Davies FE (2004). Proteomics and the haematologist. *Clin Lab Haem*, **26**, 77-86.

Ryuji Kobayashi (2001). A proteomics approach to find a new breast cancer-specific antigenic marker. *Clinical Cancer Research*, **7**, 3325-7.

Shevehenko A, Wilm M, Vorm O, Mann M (1996). Mass spectrometric sequencing of proteins from silver- stained polyacrylamide gels. *Anal chem*, **68**, 850-8.

Steinberg T, Jones LJ, Haugland RP, Singer VL (1996a). SYPRO ruby and SYPRO red protein gel stains: One step fluorescent staining of denaturing gels for detection of nanogram levels of protein. *Anal Biochem*, **239**, 223-37.

Steinberg T, Jones LJ, Haugland RP, Singer VL(1996b). Applications of SYPRO orange and SYPRO ruby gel stains. *Anal Biochem*, **239**, 238-45.

Thomas Efferth, Manfred Volm (2004). Protein expression profiles indicative for drug resistance of kidney carcinoma. *Cancer Genomics & Proteomics*, 1, 17-22.

Vestling MM, Feneslau C (1994). Polyvinylidene difluoride (PVDF): an interfacece for gel electrophoresis and matrix assisted laser desorption / ionization mass spectrometry. *Biochem Soc Trans*, **22**, 547-51.

Wolfkuhle JD, Mc Lean KC, Paweletz CP, et al (2001). *Proteomics*, **1**, 1205-15.

Yates JR 3rd, Carmack E, Hays L, Link AJ, Eng JK (1999). Automated Protein identification using microcolumn liquid chromatography - tandem mass spectrometry. *Meth Mol Biol*, **112**, 553-69.

Yates JR 3rd, Mc Cormack AL, Schiettz D, Carmack E, Link A (1997). Direct analysis of Protein mixtures by tandem mass spectrometry. *J Protein Chem*, **16**, 495-7.