Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography B, 787 (2003) 207–222
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Review

Protein expression profiling in human lung, breast, bladder, renal, colorectal and ovarian cancers

Ayodele A. Alaiya[a,*], Uwe J. Roblick[a,b], Bo Franzén[c], Hans-Peter Bruch[b], Gert Auer[a]

[a]*Unit of Cancer Proteomics, Department of Oncology and Pathology, Cancer Center Karolinska, (CCK R8: 04), Karolinska Institutet and Hospital, S-171 76 Stockholm, Sweden*
[b]*Department of Surgery, Medical University Lübeck, 23538 Lübeck, Germany*
[c]*AstraZeneca R&D Södertälje, Södertälje, Sweden*

Keywords: Reviews; Proteomics, cancer

Contents

1. Introduction

Proteins are the active molecules that carry most cellular functions while DNA sequence information per se reveal little or nothing about the level of expression of encoded proteins. It is apparent that the paradigm of one gene encoding a single protein is no longer tenable because of processes such as alternative mRNA splicing, RNA editing and post-translational protein modification. Thus, the functional complexity of normal or neoplastic tissues exceeds that indicated by its genome sequence alone. This problem can only be addressed through direct gene expression analysis either at the mRNA or protein

*Corresponding author. Tel.: +46-8-5177-5249; fax: +46-8-321-047.
E-mail address: alaiya.ayodele@cck.ki.se (A.A. Alaiya).

1570-0232/02/$ – see front matter
doi:10.1016/S1570-0232(02)00559-7

level. Powerful techniques have been developed for the rapid screening of mRNA expression, but it has become apparent that there is no predictive correlation between the level of RNA transcription and the quantity of the corresponding functional protein present within a cell. Additionally, protein degradation can significantly influence the intracellular concentration of active protein molecules. It has been estimated, that the average number of proteins per gene is one or two in bacteria, three in yeast and from three to more than six in humans [1].

The extent of complexity resulting from modifications and degradation of proteins can only be understood through qualitative and quantitative studies of gene expression at the level of functional proteins themselves. Therefore the direct measurement of protein expression is essential to analyze biological processes in normal and disease conditions. This interface between protein biochemistry and molecular biology for global analysis of gene expression is termed ''proteomics'' [2]. The core elements of the classical proteomics research combines the separation of polypeptides using the ''workhorse'' two-dimensional gel electrophoresis (2-DE) together with mass spectrometry (MS) or tandem MS (MS–MS) protein identification as schematically outlined in Fig. 1. Based on these techniques, complete proteome analyses have been undertaken only for relatively simple organisms such as *M. genitalium* [3], *H. influenzae* [4], *S. melliferum* [5], *E. coli* [6] and yeast [7]. Characterizing the complete proteome of more complex organisms, including human beings, is a challenging but perhaps impossible task using the currently available technology. However, proteomic analysis can be used in a narrow context to define patterns of protein expres-

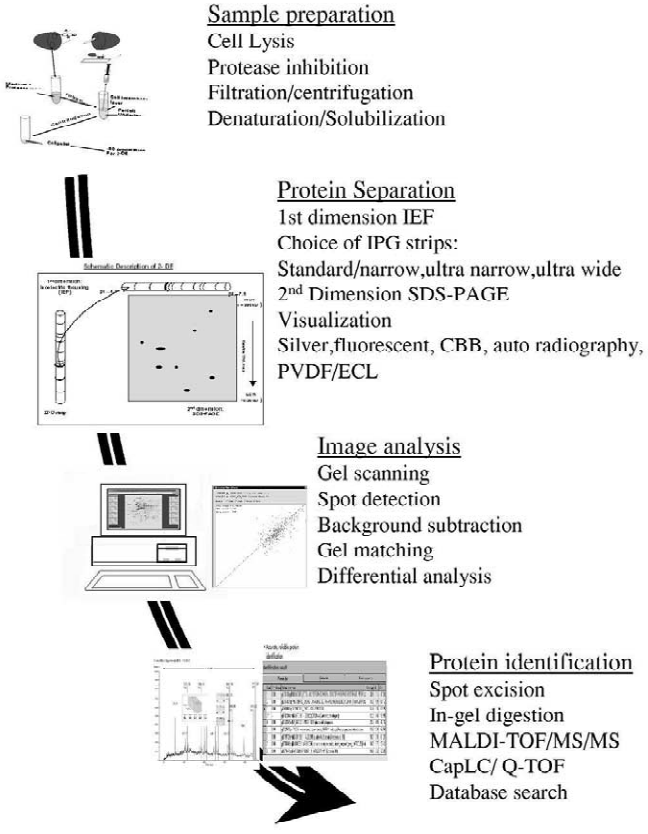

Fig. 1. Schematic illustration of the basic steps in proteomics technique including sample preparation, protein separation, image analysis and protein identification.

sion in particular cells and tissues. In cancer research, this information can then be used to identify functional cellular processes, involved in the characteristic behavior of normal cells and their malignant counterparts.

Regarding neoplasm, early detection and/or the use of treatment modalities rendering total dismissal of the tumor mass are the only procedures promising fundamental improvement of patients survival. In spite of diagnostic difficulties, empirical clinical observations clearly show, that human tumors of the same histogenetic origin, e.g., adenocarcinomas of the breast, ovary and colon may include subtypes, extremely divergent in their aggressiveness and treatment sensitivity. Comprehensive data accumulated during the last years indicate, that this complexity in tumor behavior can hardly be defined by a few alterations of gene expression. There is need for a more extensive mapping of cell composition and function. Therefore, the discovery of biological markers or marker sets, the prediction of treatment and patient outcome, and the deeper insight in malignant cell signaling pathways are in the main focus of cancer proteomics research.

2. Molecular markers

An important property of a marker is its sensitivity and specificity irrespective of the source of measurement (tissue, blood, body fluids). The specificity of the majority of the existing markers is questionable. One example, the prostate specific antigen (PSA), is a serum marker for diagnosis of prostate cancer showing wide variations in its sensitivity and specificity. Another serum marker, the α-fetoprotein (AFP) is used in clinical diagnosis of hepatocellular carcinoma (HCC) [8]. High levels of this marker can be detected, e.g., in patients suffering from ovarian or testicular cancers and patients with liver cirrhosis [9]. In addition, HCC often co-exists with liver cirrhosis and therefore the power of this marker is diminished. A summary of some of the marker proteins being used in routine tumor diagnosis is presented in Table 1.

Due to the multi-factorial genesis of cancer and the complexity of treatment responses, it is very likely that a combination of several markers will be necessary to effectively predict the biological behavior of different tumors. In this context it is useful to distinguish between diagnostic, prognostic and treatment predictive markers. *Diagnostic markers* are used to aid histopathological classification. Lung cancer can serve as an example. These neoplasms are classified into squamous, small and large cell carcinomas as well as adenocarcinomas. Furthermore, lung malignancies have to be divided into primary lung carcinomas and metastases originating from extrapulmonary malignancies. Markers need to be developed to aid the classification of these different tumors. This is not only of academic interest but is necessary to make optimal treatment choices. We observed that TA02 (Napsin A), a new type of aspartic protease is expressed in primary lung adenocarcinoma but not in metastases from, e.g., colorectal malignancies [10]. Thus the analysis of this aspartic protease in lung adenocarcinomas can be decisive for diagnosis and treatment.

Prognostic markers such as hormone receptors, proliferation markers, proteases, markers of angiogenesis, etc., are used in routine diagnosis of cancer (Table 1). These markers provide information about the malignant potential of the tumor, and hence of the patient's prognosis. Such information is of

Table 1
Examples of tumor marker proteins currently being used in routine tumor diagnosis

Abbreviation	Name	Detection method	Target cell/tumor type
ER	Estrogen receptor	Immunohistochemistry/ELISA	Breast carcinomas
PSA	Prostate specific antigen	Immunohistochemistry/ELISA	Prostate carcinomas
CA 125	Cancer antigen 125	Immunohistochemistry/ELISA	Ovarian carcinomas
CEA	Carcinoembryonic antigen	Immunohistochemistry/ELISA	Gastrointestinal cancers
CK7,8,15,18,19,20	Cytokeratins	Immunohistochemistry/2-DE	Adenocarcinomas of epithelial origin
PCNA	Proliferating cell nuclear Antigen	Immunohistochemistry/2-DE	Cell proliferation marker
AFP	α-Fetoprotein	Immunohistochemistry/ELISA	Hepatocellular carcinomas

utmost importance for clinicians to develop adequate therapeutic strategy. In patients with tumors of low malignant potential and an excellent prognosis, highly cytotoxic adjuvant chemotherapy with severe side effects can be avoided. On the other hand, patients with highly malignant tumors may benefit from a more aggressive intervention and surveillance. Such individually tailored treatment modalities will significantly improve the quality of life as well as the clinical outcome of cancer patients.

Finally, treatment *predictive markers* are used to choose between different therapy modalities. Patients who have estrogen receptor-positive breast tumors are generally treated with anti-estrogens such as tamoxifen, whereas patients with estrogen receptor-negative tumors are treated with chemotherapy. In general, whereas a fairly large number of diagnostic and prognostic markers have been described, there are only a few markers, which can predict the treatment outcome in each individual patient.

3. Proteomic technology

An adequate sample preparation procedure is instrumental to high quality 2-DE results. The procedure includes preparation/purification of cells, cell lysis and solubilization; a schematic illustration is given in Fig. 1. Sample selection and handling are of paramount importance, for example the use of fresh tumor material for 2-DE analysis was reported to be superior to working with frozen sample [11]. Another key problem in cancer proteome analysis is tissue heterogeneity. It is decisive that the analyzed cells are ''pure and relevant'' (free of stroma, blood, serum) and represent the cell population to be investigated. In this context the impact of cancer cell lines derived from tissue samples cannot be overstated. However, the representativity of the cell line model and the correlation and transfer of results to ''real life'' are still under debate.

Different methods have been developed to deplete ''non-relevant'' cell types in tissue samples. Reymond et al. described a sample preparation method for human colorectal tissue [12]. Crypts were isolated by mechanical preparation and epithelial cells were selected using Dynabeads. Significant changes in protein expression between normal mucosa and colorectal cancer were observed in the resulting 2-DE gels. Sarto et al. used an antibody-based strategy to purify human kidney cells in whole tissue samples from contaminating lymphocytes [13]. Other investigators have worked with pieces of tumor tissue without purification of the tumor cells [14]. The rationale was, that stromal components are part of the tumor and that many genes are differentially expressed in tumor connected stroma compared to the normal counterpart. Examples are proteolytic enzymes that are often expressed by tumor surrounding fibroblasts but not by the tumor cells [15,16]. However, solubilizing tumor tissue without enrichment for tumor cells leads to very complex patterns due to the presence of several cell types, and makes an interpretation of such studies problematic.

Recently, different groups showed, that laser capture microdissection (LCM) could be used to prepare defined cell populations from normal and malignant tissue samples [17–19]. This method of sample procurement could be quite suitable in some tumor types (e.g., prostate) with high degree of heterogeneity. Banks et al. showed a comparative analysis of 2-DE derived from LCM and whole tissue [20]. The result indicated a high degree of similarity between the different samples, but with more enrichment of some proteins in the LCM sample. Similar results were reported by other authors [21,22]. This demonstrates the value of LCM in classical proteomics using the conventional large format 2-DE gels to study protein expression in complex tumor samples. Even though this technique proved to be generally applicable and offers several benefits for protein profiling studies with small numbers of samples, the time factor (low-throughput) as well as sample alterations due to tissue staining remain drawbacks. Furthermore, a high quality of the dissected material can only be guaranteed if the LCM is performed by or under direct supervision of a specialized pathologist.

Finally, subcellular fractionation of cellular organelles can be used to enrich cells for analysis. This improves the resolution and in addition higher number of protein spots can be separated. In addition, the procedure increases the power and success rate of protein identification and reduces the ambiguity in protein mixtures. Furthermore, proteins that are specifically associated with a particular

organelle may be easily identified. However, fractionation requires larger quantities of biopsy material and this requirement can only be fulfilled for some tumour groups.

Our group has developed a non-enzymatic cell extraction method using fresh clinical tumor material of different entities. Depending on the size and the histomorphological characteristics of the tumor type investigated (richness of tumor cells↔vs. amount of connective tissue↔vs. fragility of cells), the method involves mechanical cell extraction with either fine needle aspiration (e.g., thyroid and breast tumours), surface scraping of tumor tissue (e.g., prostate tumours) or mincing of tumor tissue (e.g., non-Hodgkin lymphomas). As in cytological fine needle biopsies the underlying rational is that tumor cells are less attached to connective tissue and will become preferentially released by mechanical force. This method is rapid and generally results in pure tumor cell populations, free from contaminating serum proteins, red blood cells, connective tissues and necrotic tissue materials. The validity of the method has been checked for intrasample variability and ability to detect expected changes in the expression of proliferation associated genes. After preparation, the representativity of each individual sample has to be controled by a pathologist who compares the routine histology of a subsequent H&E section with a Giemsa-stained smear of the filtered cells.

3.1. Advances and limitations in 2-DE-based technology

3.1.1. 2-DE, protein identification and sequencing

Two-dimensional electrophoresis (2-DE) is the most powerful and commonly used method currently available to resolve complex protein mixtures from cells, tissues, or other biological samples. The separation method is based on two independent parameters; the first-dimension isoelectric focusing separates proteins exclusively according to their charges (isoelectric point p*I*). Proteins are then further separated in a second dimension, according to their relative molecular mass (M_r). A representative 2-DE gel derived from a malignant ovarian tumor is shown in Fig. 2. Some of the identified gel separated proteins are marked on the gel. Many of these proteins have been described to show significant expression profiles across different human malignancies.

An overview of quantitative differences of the expression level of some proteins that discriminate between benign and malignant tumors in the breast, prostate and ovary is presented in Table 2.

Since the technique was introduced in 1975 by O'Farrell and Klose, a number of methodological improvements have been made. Great advances have been achieved in making 2-DE more reproducible through the development of immobilized pH gradients (IPG strips) of different ranges, thus allowing inter-laboratory comparison of results [23]. This improvement increases the loading capacity (milligram amounts) of total proteins [24,25]. High protein loads, especially in the absence of more sensitive detection methods will enhance detection of less abundant proteins and subsequently provide sufficient amounts for protein sequencing.

Since the introduction of IPG strips a wide spectrum extending from more acidic to more basic proteins can be separated with excellent resolution [26]. The resolution of more basic and hydrophobic proteins has also been increased by the introduction of narrower basic pH strips. The use of both wide (4–12) and narrow (10–12) pH ranges to resolve those basic proteins has been extensively described [27].

The sensitivity of some of the staining methods is unable to detect low copy number proteins. Other proteins have limited solubility in the sample lysis buffer being used in isoelectric focusing. Rabilloud and co-workers reported a number of new reagents with significant impact to improve sample solubility such as more efficient detergents, zwitterions and chaotropes. This particularly permits the solubilization of hydrophobic proteins. The consequence of too many spots may result in gel overcrowding. On the other hand, the use of sequential extraction protocols especially in complex protein mixtures from tissue samples will significantly increase the total number of resolved spots and avoid gel overload on the expense of the high abundant proteins [28–31].

One area within the field of proteomics that has witnessed considerable improvement and almost full automation in protein characterization is mass spectrometry (MS) [32,33]. Nowadays, MS is the method

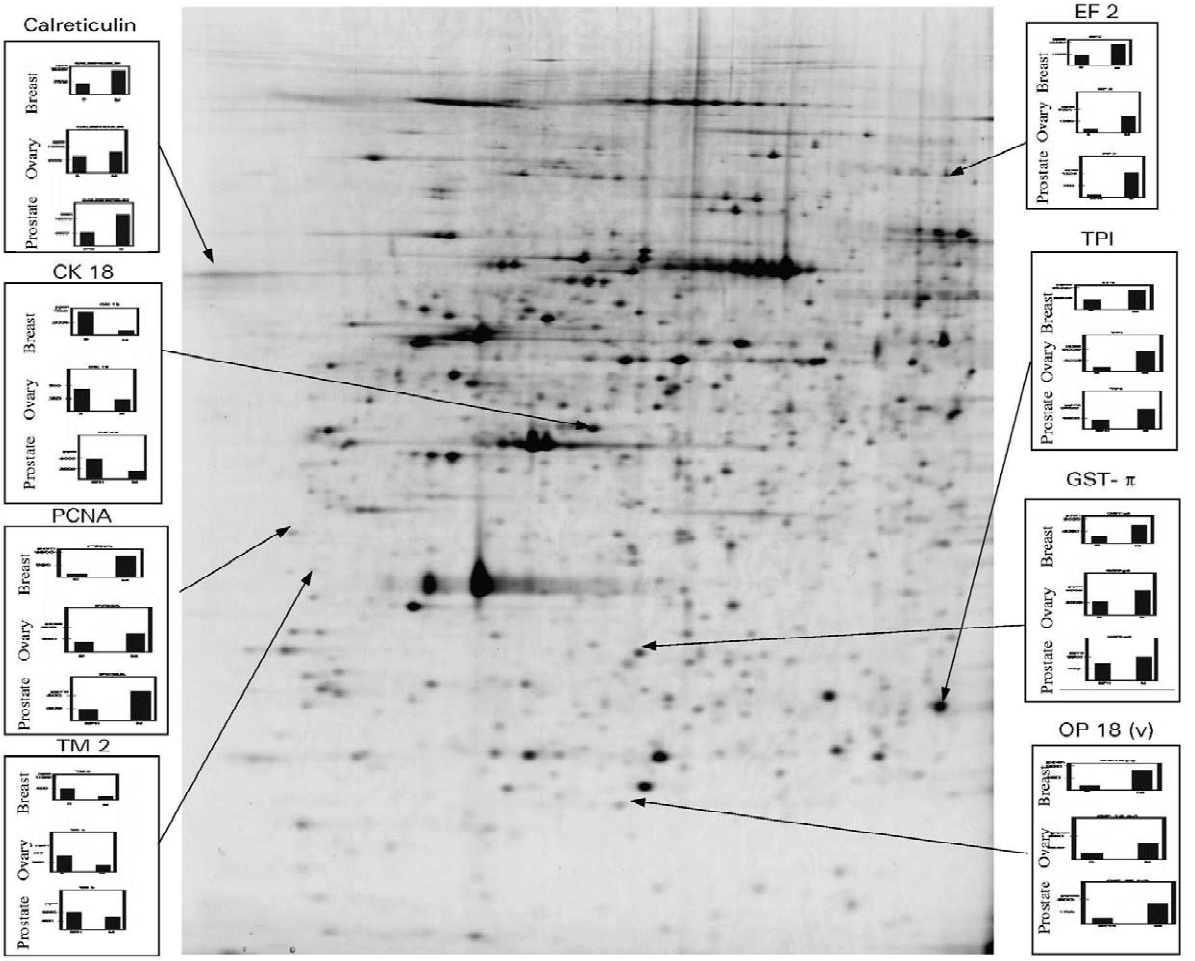

Fig. 2. Protein expression profiling in human malignancies: arrowed are some polypeptides with known identities that differ in expression between benign and malignant tumors of the breast, ovary and prostate. The first histogram bar represents average protein quantitation in benign tumors and the second bar average quantitation in malignant tumors.

of choice for protein identification and a potential tool for the characterization of post-translational modifications.

With full automation, the matrix-assisted laser desorption and ionization time of light mass spectrometry (MALDI-TOF-MS) technique is practically capable of analyzing several thousands of gel-separated proteins within a reasonable time period. The introduction of instrumentations such as spot cutter, robots directly guided by the image analysis software, development and incorporation of robotic prep stations (e.g., Proteomeworks™ system/WorksBase™) resulted in high-throughput identification of gel separated protein spots. This type of integrated system facilitates the logistics at the same time as it decreases the problem of keratin contamination.

Another great advantage is the improved sensitivity of the MS instrumentation. At present, subpicomole quantities of protein are sufficient to give non-equivocal protein identification [34].

There are several publicly available EST databases which are regularly up-dated and which have significantly improved the percentage of spots identified following MALDI-TOF-MS analysis. This has led to

Table 2
Protein expression profiling in human malignancies: list of identified proteins in which the combined expression pattern allows to discriminate between benign and malignant tissues from the breast, prostate and ovary

	Abbreviation	Name	Rel. mol. mass/p*I*	Food change benign:malignant		
				Breast	Prostate	Ovary
1	PCNA	Proliferating cell nuclear antigen	34.6/4.86	3.7	2.5	2.0
2	OP 18 (v)	Oncoprotein 18 variant	16.7/5.5	3.3	3.7	2.7
3	GST-π	Glutathione *S*-transferase-π	25.1/5.44	2.85	1.5	2.0
4	EF 2	Elongation factor 2	103.0/6.9	2.2	6.5	4.0
5	TPI	Triose phosphate isomerase	27.1/6.53	2.2	2.2	4.3
6	CALR	Calreticulin	68.0/4.0	2.3	2.3	1.3
7	SOD	Superoxide dismutase	22.2/6.72	2.5	2.75	1.7
8	HSP 90	Heat shock protein 90	90.9/5.21	2.4	4.0	1.7
9	CK 18	Cytokeratin 18	47.9/5.40	0.22	0.44	0.53
10	TM 2	Tropomyosin 2	36.3/4.88	0.22	0.58	0.38

a considerably reduced number of samples that have to be subjected to full sequence analysis.

3.1.2. Automation and high-throughput analysis

One major limitation of classical protein separation using 2-DE is the lack of high-throughput expression analysis. The art of running 2-DE is still labor intensive and it takes approximately 1 week from sample preparation to image analysis. Improved instrumentation, more sophisticated data analysis software and bioinformatics are needed for rapid characterization of large amounts of gene products. The ''in-house'' robotic systems for handling standard format 2-DE gels are more commonly used by mainly commercial proteomics groups [35]. These systems allow nearly full automation in handling hundreds of 2-DE gels with only few interventions in running, staining and analysis.

Fluorescence-two dimensional differential gel electrophoresis (2-D DIGE) is a new development in proteomics. Two samples containing complex protein mixtures can be labeled covalently by two fluorescent dyes prior to isoelectric focusing; the mixture of the samples is then run in a single 2-DE gel [36,37]. This method allows differential expression analysis and significantly reduces the problem of gel matching, since the two samples are run under the same condition.

Getting sufficient amounts of representative clinical samples for classical proteome studies is often limited. New methods such as LCM and micro fluidic approaches are emerging technologies that may help to solve this problem.

High throughput methodologies and novel instrumentation are rapidly emerging in proteome studies e.g., affinity chromatography, protein arrays/assays, lab on chip technique (Agilent-Chemical Life Sciences). Surface-enhanced laser desorption/ionization (SELDI, Ciphergen Biosystems), is a method based on chromatographic protein separation coupled with MALDI-TOF-MS. This technique may provide an alternative rapid protein profiling and has the power and capacity to analyze hundreds of samples simultaneously. The analysis is rapid, requires only micro volumes of sample and minimal sample preparation procedures have to be performed. One obvious drawback of this technology is the fact, that SELDI is only efficient for profiling low-molecular mass proteins ($<$20 kDa), as this can only cover a small subset of the entire proteome.

3.1.3. Information technology and efficient data management

The developments in computer-based image analysis systems (e.g., PDQUEST and MELANIE) have allowed the efficient evaluation of thousands of spots on 2-DE maps [38–40]. However, the accuracy of gel matching declines with increasing heterogeneity between the samples being compared, especially if 2-DE gels were produced in different experiments or with varying pH ranges. Even though some of the available 2-DE software has improved the reliability and the accuracy of the automated processes of spot

detection, spot counting, quantitation, matching and statistical analysis, a limited level of manual editing and re-evaluation of auto tasks will probably always be required.

There are several 2-DE, protein and genomic databases available on public domains. They offer access to different proteins, support information and include multiple query options. These databases can be reached from Web sites like World 2D-PAGE at the ExPASy server in Geneva (http://www.expasy.ch/ch2d/2d-index.html). Lemkin described a dedicated gel comparison method that allows the user to match his local gel with any gel image from any of the www-based 2D gel databases [41]. The program is available on the web site (http://www.lecb.ncifcrf.gov/flicker). Such inventions make inter-laboratory gel comparisons possible and enhance protein spot identification by gel matching. Despite improvements in 2-DE methodology, resulting in high gel reproducibility, exchange of 2-DE gel patterns for inter-laboratory comparisons should be treated with caution. This is because of some minor variations in sample handling prior to 2-DE and differences in the types of samples. Protocols, which may give rise to protein modifications should be avoided, as these changes in sample preparation protocols and handling may result in specific loss or gain of protein species.

The web-based gel matching limits were experienced by Jungblut et al. when they unsuccessfully attempted to match their colon cancer gel with a reference map of liver tissue and a colon carcinoma cell line [42]. In contrast, positive results of web-based gel matching were achieved by Bini et al. when more than 30 protein spots from breast cancer samples could be matched to different reference maps of the SWISS-2D-PAGE database [14].

Automation in mass spectrometry has facilitated rapid protein identification. Some of these databases offer features to create ‘‘clickable’’ spot images in which annotations to protein spots can be made (e.g., ProteomeWorks™/WorkBase™; Bio-Rad/Micromass). Another interesting development in proteome data acquisition and database development is the creation of a laboratory information processing system (LIPS) by Ali et al. [43]. Such a comprehensive data inventory contains information like sample entry, disease diagnosis, analyzed images, spot quantitation data, etc.

More of such user-friendly software packages with features for integration of clinical and pathological data, types of samples and 2-DE expression patterns will increase the possibilities for multivariate analysis of data and prevent loss of vital biological information. In the future, integrated networks may serve as a ‘‘*proteome scanner*’’, capable of assisting and/or complementing clinicopathological judgement. Hopefully, this will lead to a more accurate diagnosis, prognosis and treatment prediction in clinical practice.

4. Applied proteomics in cancer research

In view of the present state of the art in genetic cancer research, clinicians are still faced with difficulty in early detection and diagnosis of pre-clinical lesions. Furthermore, it is difficult to predict disease-specific biological behaviour and whether or not a tumour will metastasise or resist cytostatic and radiation therapies. The goals of cancer proteomics are to improve molecular classification of tumors and to discover more sensitive biomarkers for disease prognosis and treatment sensitivity assessment. An overview of cancer proteomic strategies is presented in Fig. 3.

The increasing interest of both the academic institutions and the biopharmaceutical industry reflects the potential of proteome approach in the discovery of proteins as targets for drug development. Drug development in the treatment of disease conditions primarily acts via protein functions and changes in their expression levels as applicable to different pathological states may help to design individual tailored treatment modalities. Brief summaries of interesting results and potentials of the proteomics approach in the study of different human cancers are given below.

4.1. Bladder cancer

With a steady increase in the incidence of bladder cancer there is need to develop valid biomarkers to characterize the malignant potential of individual tumors and thus determine the prognosis. This rationale motivated Celis, Ostergaard and co-workers to perform 2-DE-based proteomic analysis of hundreds of urine samples and bladder tumors. Expres-

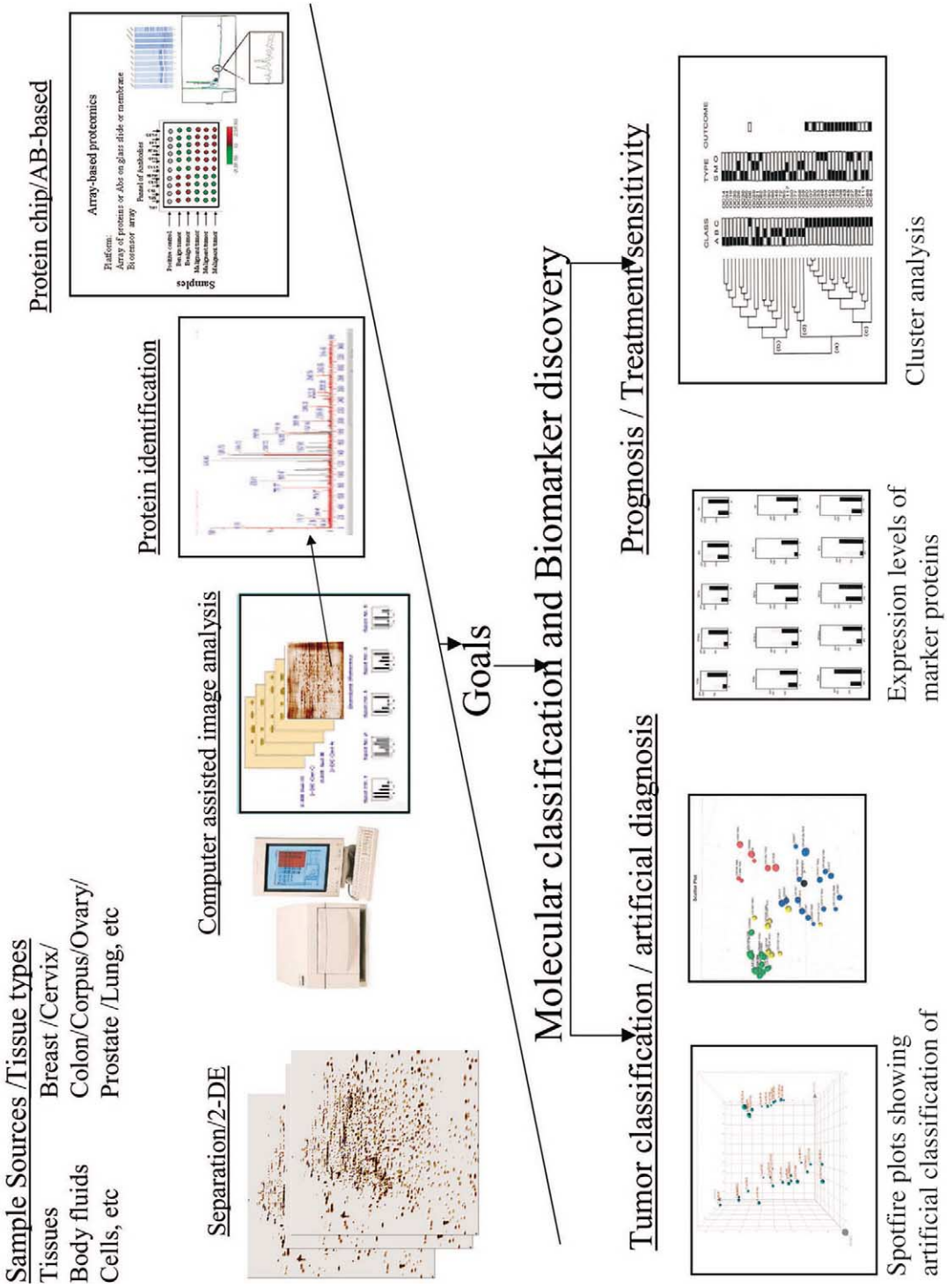

Fig. 3. An overview of cancer proteomics strategies.

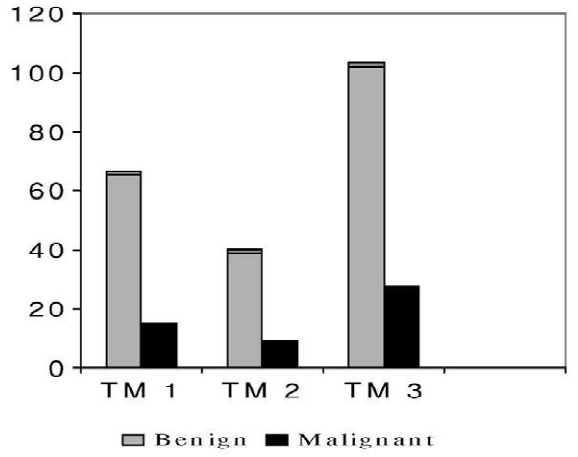

Fig. 4. Relative expression levels of three of the tropomyosin isoforms in benign and malignant breast tumors. Grey box, benign; black box, malignant.

sion of psoriasin, stratifin and gelatin could be related to a decreasing degree of differentiation. Psoriasin expression in the urine samples of patients suffering from squamous cell carcinoma was reported. This may be a potential diagnostic marker for early detection of premalignant bladder lesions [44–46]. An ELISA test may soon be available to validate Psoriasin as an early marker of this disease. An expression database for squamous cell carcinoma (SCC) as well as transitional cell carcinoma (TCC) is available at http://biobase.dk/cgi-bin/Celis [47].

4.2. Breast cancer

Tumor morphological and histopathological characteristics are still the only widely accepted forms of routine diagnosis and classification of breast tumors. Especially the accurate diagnosis of precursor breast lesions remains difficult. For example, features of lobular and ductal atypical hyperplasia are quite similar to those of lobular carcinoma in situ and ductal carcinoma in situ [48]. Such biological differences will be better understood using more objective molecular profiling methods.

Franzén et al. have systematically investigated protein profiles derived from fresh clinical tissue samples of different grades and histological types [49,50]. These studies described several proteins (including calreticulin, HSP 60, HSP 90 and PCNA) with significantly increased expression in invasive carcinomas compared to benign breast lesions. In contrast, the high-molecular mass tropomyosin isoforms were observed to be down regulated in carcinomas. Fig. 4 shows the relative expression of tropomyosin 1–3 in benign and malignant breast tumours. A decreased expression of these tropomyosins was similarly observed in a primary breast cancer cell line [51]. Several studies have described the expression of the cytoskeletal protein family in breast epithelial cells. A large number of cytokeratins including cytokeratin 5, 6, 7, 8, 17, 18 and 19 were reported to be differentially expressed in cultured normal breast epithelial cells and the corresponding tumor cells [52]. The expression of cytokeratins 8 and 18 was found to be elevated in fibroadenomas when compared with breast cancers [49]. Irrespective the entity, similar patterns of high expression of cytokeratin 8 and 18 were reported in benign ovarian tumors as well as benign prostatic hyperplasias [53]. In an attempt to define more sensitive breast cancer markers for clinical use, Hondermarck et al. reported that a molecular chaperone-14-3-3 σ was pathognomonic of normal breast epithelial cells and rarely detectable in breast cancer [54]. A 2-DE database providing expression profiles from several human breast samples was developed by Giometti et al. and can be reached on: http://www.nl.gov/CMB/PMG/projects/index-_hbreast.html [55].

4.3. Colo-rectal cancer

Both Jungblut et al. and Stulik et al. successively defined the differential expression patterns of normal colonic mucosa, colonic polyps and adenocarcinomas. Several proteins that were up regulated in the neoplastic lesions were described, including HSP 70 and some members of the S-100 family [42,56,57]. The validity of these gel-separated proteins was supported by immunohistochemical analysis. In another study, Stulik et al. went further to show, that a polypeptide with relative molecular mass of 13 kDa that was subsequently identified as calgranulin B, was expressed in patients suffering from colon cancer and ulcerative colitis [57]. Both calgranulin A and calgranulin B are components of calprotectin, a protein that was earlier described to be found in stool

samples of patients suffering from tumors of the gastrointestinal tract [58]. This protein could be a useful potential marker for colon cancer screening. In a recent study, Stulik et al. investigated 27 pairs of colonic tumors, normal mucosa and 13 colonic polyps. A total of 18 proteins exhibiting significant quantitative changes between normal, premalignant and carcinoma samples were identified [59]. All the identified polypeptides are suspected to play a role in the genesis of colonic tumors. Recently, our group succeeded using a proteomics approach in defining a poorly differentiated pelvic mass involving the left ovary and the recto-sigmoid colon. Standard histological and immunohistochemical investigations failed to clearly define the origin of this tumor. Computer tomography scan and nuclear magnetic resonance imaging could rule out all other possible origins except for ovarian and colorectal cancer. Ten T3/T4Nx sigmoid cancer samples and 10 high stage ovarian carcinomas were run with 2-DE and matched against the pelvic tumor. The image analysis demonstrated, that the pelvic tumor was of colorectal origin (submitted).

4.4. Lung cancer

Proteome studies of lung tumors have shown that different histopathological tumor types exhibit significantly varied protein expression patterns. Schmid et al. described several protein spots resolved from 2-DE gels that correlated with the different histopathological types of lung cancer [60]. Later, Hirano et al. observed similar results. In their analyses of clinical lung cancer samples they found a pair of polypeptides (TAO1, TAO2), that were significantly overexpressed in primary lung adenocarcinomas but absent in secondary adenocarcinomas metastasized from the colo-rectum [61]. One of these spots (TAO2) was later identified to be a novel protease and very similar to Napsin A, a member of the aspartic protease family [10]. The role of this protein is still unclear, but it is suggested to be involved in the cleavage of pro-surfactant proteins produced by the type II pneumocytes.

More recently, a group of gel-separated proteins were described and found to have differential expression profiles in epithelial tumors of the lung, breast and ovary [62]. Oh et al. have developed a comprehensive lung cancer database consisting of both transcriptomic and proteomic information for different types and stages of lung cancer [63].

4.5. Ovarian cancer

Studies from Alaiya and co-workers have defined benign, borderline and malignant ovarian tumors [64]. Ovarian carcinomas showed high abundance of proliferating cell nuclear antigen (PCNA), oncoprotein 18, phosphorylated heat shock protein 60 (pHSP 60), HSP 90 and elongation factor 2. The expressions of tropomyosin 1 and 2 were decreased in the carcinomas compared to the benign tissues. Using a set of nine proteins, it was possible to discriminate between benign and malignant samples. Bergman et al. identified potential marker proteins that are upregulated in ovarian malignancies, including retinoic acid-binding protein II, galectin-1, apolipoprotein A-1 and annexin IV [62]. Our group has used the differential expression of proteins resolved by 2-DE to classify ovarian tumors using the combined score of 223 polypeptides. It was possible to distinctly discriminate between malignant and benign ovarian tumors, whereas borderline tumors were classified as intermediate or benign using principal component and partial least square discriminant analysis (PCA/PLS-DA) [65]. When the data was further analyzed using hierarchical cluster analysis, the classification of the borderline tumors was improved as shown in Fig. 5.

We recently investigated whether prognostic information regarding patient outcome could be achieved using hierarchical cluster analysis of protein expression profiles. Twelve of 19 patients with ovarian carcinomas had a disease related fatal clinical outcome (2–6 years clinical follow-up), while none of the benign and borderline patients died of the disease. There was no clear clustering of tumors with a similar outcome when the dataset comprising all 40 patients was evaluated (Alaiya, in press, Int. J. Cancer). However, when only the 19 patients with ovarian carcinoma were evaluated, tumors from patients with similar outcome distinctly clustered together (Fig. 6). These results are promising. Even though, analysis of larger numbers of patients with tumors of similar subtype and clinical stage will be

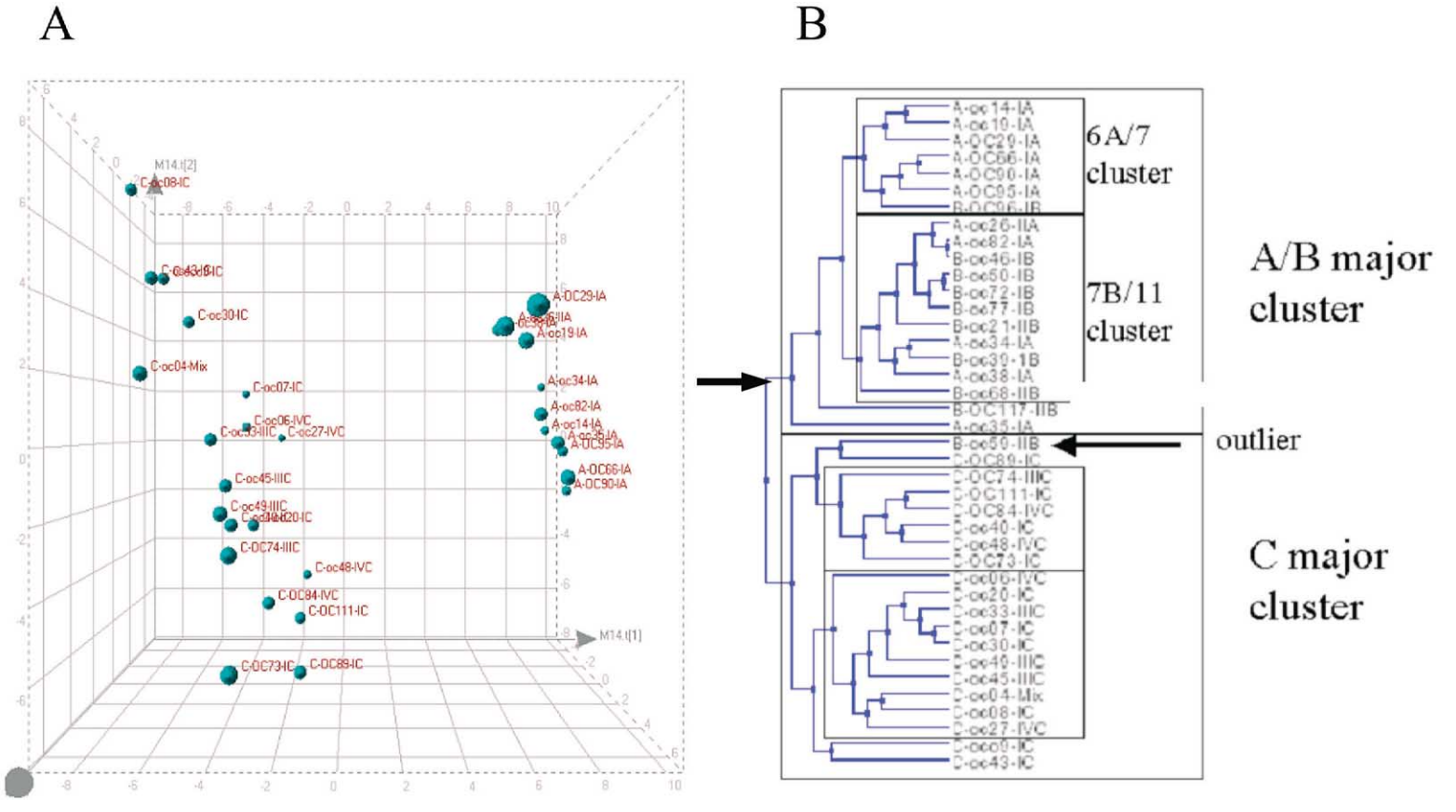

Fig. 5. (A) Expression data of 223 protein spots from 40 ovarian tumors was used to classify the tumors into benign and malignant using the spot fire software. (B) This data set was then applied to accurately predict the classification of the borderline tumors using hierarchical cluster analysis. Classification: (A) benign; (B) borderline; (C) malignant.

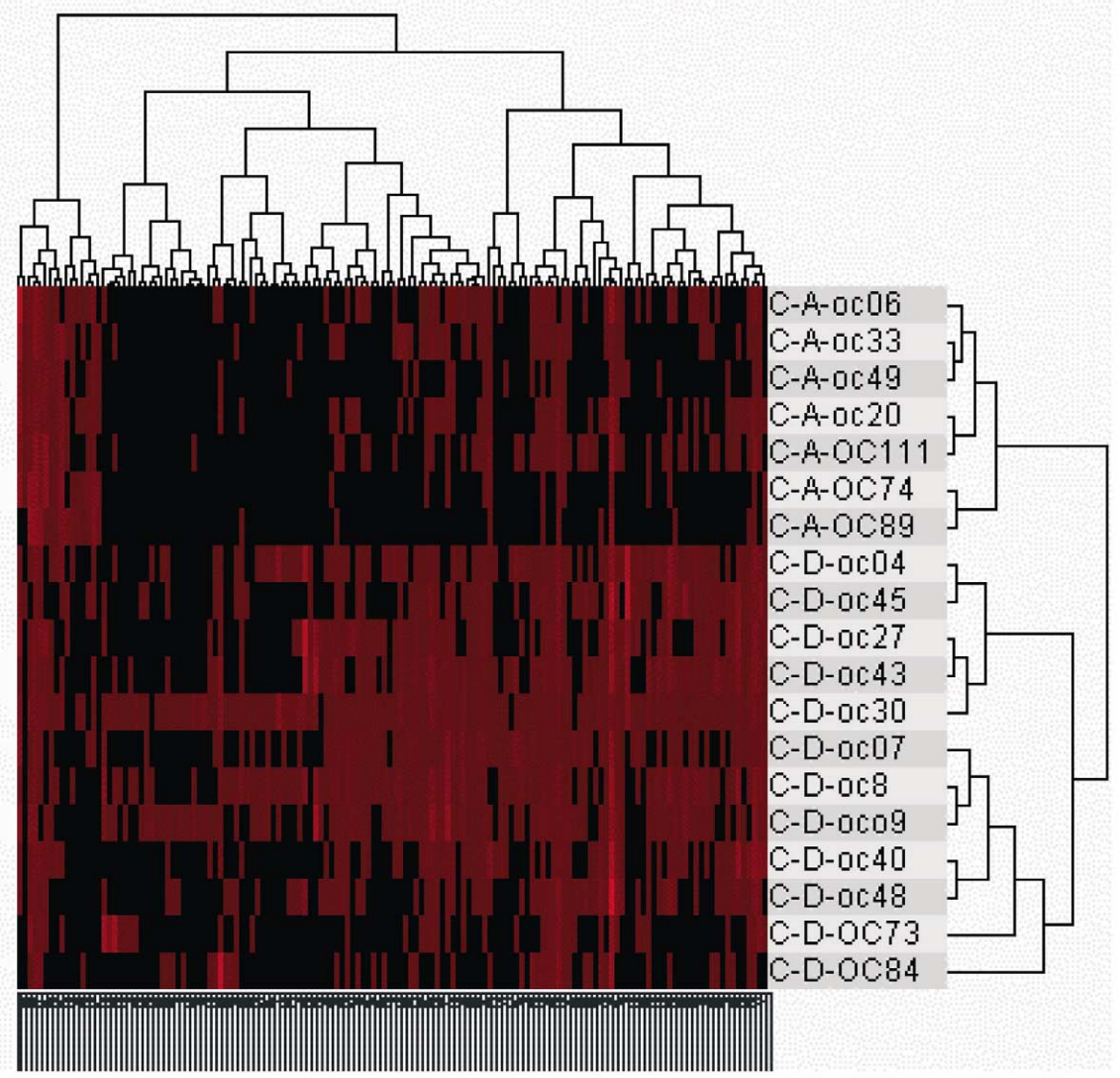

Fig. 6. Dendrogram from Manhattan Average linkage correlation hierarchical cluster analysis of 19 ovarian carcinomas using the J-Express software. The expression levels of 142 protein spots were used to classify the samples in relation to survival. Strikingly, exclusively all tumors from surviving patients could be clustered. (C-A, carcinoma alive; D, carcinoma Dead). Analysis details are described in Alaiya et al. (in press, Int. J Cancer).

necessary to validate, whether cluster analysis can be used as a prognostic tool.

4.6. Renal cancer

Several approaches including proteome analysis have been directed to study kidney cancers in order to gain better insight in the development and management of renal cell carcinomas (RCC). For this entity is still neither an acceptable screening method nor an early diagnostic marker available.

Differential protein expression profiles of normal and malignant kidney specimens were investigated by Sarto et al. [13]. Among those proteins identified as potential markers for RCC were ubiquinol cytochrome reductase and mitochondrial NADH-ubiquin-

one oxido-reductase complex I. These proteins were expressed in normal kidney tissue but absent in RCC. In a more recent study, a pair of manganese superoxide dismutase (Mn-SOD) isoforms was identified and reported to be exclusively found in renal cell carcinomas and not in normal kidney tissue [66].

4.7. Future perspectives

2-DE is still a major component of proteomics and its high-resolution power makes it challenging to other existing separation techniques. However, it is labor intensive and requires skillful hands to generate highly reproducible results. Despite interesting data derived from 2-DE separation, especially in the area of cancer research, the obvious drawbacks of this method preclude the wide acceptance in clinical routine. There is need to develop equally sensitive separation methods that unify both, high sample throughput and reproducibility.

An isotope-coded affinity tag, followed by tandem mass spectrometry has recently been reviewed as a promising non-2-DE-based proteomic approach with potential for full automation and high-throughput analysis. Samples are isotopically labeled with heavy and light chain based on biotine and cysteine reactivity [67,68]. The limitation of affinity mass tagging is the lack of absolute quantitation, since the ratio of mass peaks generated from the different labeled samples gives only a measure of relative protein abundance in each sample.

Protein profiling by direct mass spectrometric analysis of protein mixtures can be performed [69]. Complete proteolytic digests of such complex mixtures are analyzed by liquid chromatography coupled with tandem mass spectrometry. This combined method is highly sensitive and allows identification of low abundant proteins that might elude detection using conventional 2-D gel electrophoresis.

Two-dimensional high-performance liquid chromatography is an alternative protein separation method to resolve complex protein mixtures. The HPLC uses ion-exchange (first dimension) coupled with reversed-phase liquid chromatography (second dimension). It is preferentially suited for sub cellular fractionation and protein purification. HPLC seems efficient in separating proteins within a wide molecular mass range and it is claimed that this technique is amenable to automation [70].

The future generation of expression profiling, especially in the clinical field may be seen in protein arrays that are based on different platform supports. Zhou et al. presented an interesting review on protein arrays based on solid supports such as glass, membranes or micro-fluidics [71]. A combination of liquid-phase proteomics with microarrays has been developed and described by Madoz-Gurpide et al. [70]. This separation technique results in a partial protein purification thus solving the problem of complexity of protein mixtures.

Acknowledgements

The Unit of Cancer Proteomics is supported by the Swedish Cancer Society (Cancerfonden) and the Cancer Society Stockholm (Cancerföreningen). U.J. Roblick receives a research grant from the Medical University of Lübeck (N14-2001) and a scholarship from the Karolinska Institutet, Stockholm.

References

[1] M.R. Wilkins, J.C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphery-Smith, D.F. Hochstrasser, K.L. Williams, Biotechnol. Genet. Eng. Rev. 13 (1996) 19.
[2] M.R. Wilkins, J.C. Sanchez, K.L. Williams, D.F. Hochstrasser, Electrophoresis 17 (1996) 830.
[3] V.C. Wasinger, S.J. Cordwell, A. Cerpa-Poljak, J.X. Yan, A.A. Gooley, M.R. Wilkins, M.W. Duncan, R. Harris, K.L. Williams, I. Humphery-Smith, Electrophoresis 16 (1995) 1090.
[4] A.J. Link, L.G. Hays, E.B. Carmack, J.R. Yates 3rd, Electrophoresis 18 (1997) 1314.
[5] S.J. Cordwell, D.J. Basseal, I. Humphery-Smith, Electrophoresis 18 (1997) 1335.
[6] R.A. VanBogelen, K.Z. Abshire, B. Moldover, E.R. Olson, F.C. Neidhardt, Electrophoresis 18 (1997) 1243.
[7] J.I. Garrels, C.S. McLaughlin, J.R. Warner, B. Futcher, G.I. Latter, R. Kobayashi, B. Schwender, T. Volpe, D.S. Anderson, R. Mesquita-Fuentes, W.E. Payne, Electrophoresis 18 (1997) 1347.
[8] T.K. Seow, R.C. Liang, C.K. Leow, M.C. Chung, Proteomics 1 (2001) 1249.
[9] P.J. Johnson, J. Gastroenterol. Hepatol 14 (Suppl.) (1999) 32.
[10] Y. Chuman, A. Bergman, T. Ueno, S. Saito, K. Sakaguchi, A.A. Alaiya, B. Franzen, T. Bergman, D. Arnott, G. Auer, E. Appella, H. Jornvall, S. Linder, FEBS Lett. 462 (1999) 129.

[11] B. Franzen, S. Linder, K. Okuzawa, H. Kato, G. Auer, Electrophoresis 14 (1993) 1045.
[12] M.A. Reymond, J.C. Sanchez, G.J. Hughes, K. Gunther, J. Riese, S. Tortola, M.A. Peinado, T. Kirchner, W. Hohenberger, D.F. Hochstrasser, F. Kockerling, Electrophoresis 18 (1997) 2842.
[13] C. Sarto, A. Marocchi, J.C. Sanchez, D. Giannone, S. Frutiger, O. Golaz, M.R. Wilkins, G. Doro, F. Cappellano, G. Hughes, D.F. Hochstrasser, P. Mocarelli, Electrophoresis 18 (1997) 599.
[14] L. Bini, B. Magi, B. Marzocchi, F. Arcuri, S. Tripodi, M. Cintorino, J.C. Sanchez, S. Frutiger, G. Hughes, V. Pallini, D.F. Hochstrasser, P. Tosi, Electrophoresis 18 (1997) 2832.
[15] P. Basset, C. Wolf, P. Chambon, Breast Cancer Res. Treat. 24 (1993) 185.
[16] A. Okada, J.P. Bellocq, N. Rouyer, M.P. Chenard, M.C. Rio, P. Chambon, P. Basset, Proc. Natl. Acad. Sci. USA 92 (1995) 2730.
[17] M.R. Emmert-Buck, R.F. Bonner, P.D. Smith, R.F. Chuaqui, Z. Zhuang, S.R. Goldstein, R.A. Weiss, L.A. Liotta, Science 274 (1996) 998.
[18] L. Liotta, E. Petricoin, Nat. Rev. Genet. 1 (2000) 48.
[19] C.P. Paweletz, L.A. Liotta, E.F. Petricoin 3rd, Urology 57 (2001) 160.
[20] R.E. Banks, M.J. Dunn, M.A. Forbes, A. Stanley, D. Pappin, T. Naven, M. Gough, P. Harnden, P.J. Selby, Electrophoresis 20 (1999) 689.
[21] R.A. Craven, R.E. Banks, Proteomics 1 (2001) 1200.
[22] J.D. Wulfkuhle, K.C. McLean, C.P. Paweletz, D.C. Sgroi, B.J. Trock, P.S. Steeg, E.F. Petricoin 3rd, Proteomics 1 (2001) 1205.
[23] B. Bjellqvist, K. Ek, P.G. Righetti, E. Gianazza, A. Gorg, R. Westermeier, W. Postel, J. Biochem. Biophys. Methods 6 (1982) 317.
[24] B. Bjellqvist, J.C. Sanchez, C. Pasquali, F. Ravier, N. Paquet, S. Frutiger, G.J. Hughes, D. Hochstrasser, Electrophoresis 14 (1993) 1375.
[25] S.M. Hanash, J.R. Strahler, J.V. Neel, N. Hailat, R. Melhem, D. Keim, X.X. Zhu, D. Wagner, D.A. Gage, J.T. Watson, Proc. Natl. Acad. Sci. USA 88 (1991) 5709.
[26] A. Gorg, W. Postel, S. Gunther, Electrophoresis 9 (1988) 531.
[27] A. Gorg, C. Obermaier, G. Boguth, W. Weiss, Electrophoresis 20 (1999) 712.
[28] M.P. Molloy, B.R. Herbert, B.J. Walsh, M.I. Tyler, M. Traini, J.C. Sanchez, D.F. Hochstrasser, K.L. Williams, A.A. Gooley, Electrophoresis 19 (1998) 837.
[29] T. Rabilloud, Electrophoresis 17 (1996) 813.
[30] T. Rabilloud, C. Adessi, A. Giraudel, J. Lunardi, Electrophoresis 18 (1997) 307.
[31] T. Rabilloud, Electrophoresis 19 (1998) 758.
[32] P. Chaurand, F. Luetzenkirchen, B. Spengler, J. Am. Soc. Mass. Spectrom. 10 (1999) 91.
[33] O.N. Jensen, M. Wilm, A. Shevchenko, M. Mann, Methods Mol. Biol. 112 (1999) 571.
[34] A. Shevchenko, M. Wilm, O. Vorm, O.N. Jensen, A.V. Podtelejnikov, G. Neubauer, P. Mortensen, M. Mann, Biochem. Soc. Trans. 24 (1996) 893.
[35] N.L. Anderson, A.D. Matheson, S. Steiner, Curr. Opin. Biotechnol. 11 (2000) 408.
[36] R. Tonge, J. Shaw, B. Middleton, R. Rowlinson, S. Rayner, J. Young, F. Pognan, E. Hawkins, I. Currie, M. Davison, Proteomics 1 (2001) 377.
[37] M. Unlu, M.E. Morgan, J.S. Minden, Electrophoresis 18 (1997) 2071.
[38] R.D. Appel, D.F. Hochstrasser, M. Funk, J.R. Vargas, C. Pellegrini, A.F. Muller, J.R. Scherrer, Electrophoresis 12 (1991) 722.
[39] R.D. Appel, J.R. Vargas, P.M. Palagi, D. Walther, D.F. Hochstrasser, Electrophoresis 18 (1997) 2735.
[40] J.I. Garrels, J. Biol. Chem. 264 (1989) 5269.
[41] P.F. Lemkin, Electrophoresis 18 (1997) 461.
[42] P.R. Jungblut, U. Zimny-Arndt, E. Zeindl-Eberhart, J. Stulik, K. Koupilova, K.P. Pleissner, A. Otto, E.C. Muller, W. Sokolowska-Kohler, G. Grabher, G. Stoffler, Electrophoresis 20 (1999) 2100.
[43] I. Ali, Y. Chan, R. Kuick, D. Teichroew, S.M. Hanash, Electrophoresis 12 (1991) 747.
[44] J.E. Celis, M. Kruhoffer, I. Gromova, C. Frederiksen, M. Ostergaard, T. Thykjaer, P. Gromov, J. Yu, H. Palsdottir, N. Magnusson, T.F. Orntoft, FEBS Lett. 480 (2000) 2.
[45] J.E. Celis, H. Wolf, M. Ostergaard, Electrophoresis 21 (2000) 2115.
[46] M. Ostergaard, H. Wolf, T.F. Orntoft, J.E. Celis, Electrophoresis 20 (1999) 349.
[47] J.E. Celis, P. Gromov, M. Ostergaard, P. Madsen, B. Honore, K. Dejgaard, E. Olsen, H. Vorum, D.B. Kristensen, I. Gromova, A. Haunso, J. Van Damme, M. Puype, J. Vandekerckhove, H.H. Rasmussen, FEBS Lett. 398 (1996) 129.
[48] D.L. Page, W.D. Dupont, L.W. Rogers, M.S. Rados, Cancer 55 (1985) 2698.
[49] B. Franzen, S. Linder, A.A. Alaiya, E. Eriksson, K. Uruy, T. Hirano, K. Okuzawa, G. Auer, Br. J. Cancer 74 (1996) 1632.
[50] B. Franzen, S. Linder, K. Uryu, A.A. Alaiya, T. Hirano, H. Kato, G. Auer, Br. J. Cancer 73 (1996) 909.
[51] B. Bhattacharya, G.L. Prasad, E.M. Valverius, D.S. Salomon, H.L. Cooper, Cancer Res. 50 (1990) 2105.
[52] D.K. Trask, V. Band, D.A. Zajchowski, P. Yaswen, T. Suh, R. Sager, Proc. Natl. Acad. Sci. USA 87 (1990) 2319.
[53] A. Alaiya, U. Roblick, L. Egevad, A. Carlsson, B. Franzen, D. Volz, S. Huwendiek, S. Linder, G. Auer, Anal. Cell Pathol. 21 (2000) 1.
[54] H. Hondermarck, A.S. Vercoutter-Edouart, F. Revillion, J. Lemoine, I. el-Yazidi-Belkoura, V. Nurcombe, J.P. Peyrat, Proteomics 1 (2001) 1216.
[55] C.S. Giometti, K. Williams, S.L. Tollaksen, Electrophoresis 18 (1997) 573.
[56] J. Stulik, K. Koupilova, J. Osterreicher, J. Knizek, A. Macela, J. Bures, P. Jandik, F. Langr, K. Dedic, P.R. Jungblut, Electrophoresis 20 (1999) 3638.
[57] J. Stulik, J. Osterreicher, K. Koupilova, Knizek, A. Macela, J. Bures, P. Jandik, F. Langr, K. Dedic, P.R. Jungblut, Electrophoresis 20 (1999) 1047.
[58] A.G. Roseth, J. Kristinsson, M.K. Fagerhol, H. Schjonsby, E. Aadland, K. Nygaard, B. Roald, Scand. J. Gastroenterol. 28 (1993) 1073.

[59] J. Stulik, L. Hernychova, S. Porkertova, J. Knizek, A. Macela, J. Bures, P. Jandik, J.I. Langridge, P.R. Jungblut, Electrophoresis 22 (2001) 3019.
[60] H.R. Schmid, D. Schmitter, P. Blum, M. Miller, D. Vonderschmitt, Electrophoresis 16 (1995) 1961.
[61] T. Hirano, K. Fujioka, B. Franzen, K. Okuzawa, K. Uryu, H. Shibanuma, K. Numata, C. Konaka, Y. Ebihara, M. Takahashi, H. Kato, G. Auer, Br. J. Cancer 75 (1997) 978.
[62] A.C. Bergman, T. Benjamin, A. Alaiya, M. Waltham, K. Sakaguchi, B. Franzen, S. Linder, T. Bergman, G. Auer, E. Appella, P.J. Wirth, H. Jornvall, Electrophoresis 21 (2000) 679.
[63] J.M. Oh, F. Brichory, E. Puravs, R. Kuick, C. Wood, J.M. Rouillard, J. Tra, S. Kardia, D. Beer, S. Hanash, Proteomics 1 (2001) 1303.
[64] A.A. Alaiya, B. Franzen, K. Fujioka, B. Moberger, K. Schedvins, C. Silfversvard, S. Linder, G. Auer, Int. J. Cancer 73 (1997) 678.
[65] A.A. Alaiya, B. Franzen, A. Hagman, C. Silfversward, B. Moberger, S. Linder, G. Auer, Int. J. Cancer 86 (2000) 731.
[66] C. Sarto, C. Deon, G. Doro, D.F. Hochstrasser, P. Mocarelli, J.C. Sanchez, Proteomics 1 (2001) 1288.
[67] S.P. Gygi, R. Aebersold, Methods Mol. Biol. 112 (1999) 417.
[68] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994.
[69] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J.R. Yates 3rd, Nat. Biotechnol. 17 (1999) 676.
[70] J. Madoz-Gurpide, H. Wang, D.E. Misek, F. Brichory, S.M. Hanash, Proteomics 1 (2001) 1279.
[71] H. Zhou, S. Roy, H. Schulman, M. Natan, Trends Biotechnol. 19 (Suppl. 10) (2001) S34.