Proteomics as a tool for biomarker detection using SELDI-TOF-MS and 2-D DIGE


The field of proteomics has developed rapidly in recent years. Until the mid-1990s scientists studied individual genes and proteins or a handful biologically related genes and proteins. Basis of proteomics is to characterize the behavior of the system rather than the behavior of any single component. The proteome is dynamic and in constant flux due to a combination of factors. These factors include posttranslational modifications and functional regulation of gene expression (1). Moreover, in proteomics protein identification is not necessarily performed by complete sequence analysis, but can also be performed by partial sequence analysis with the aid of database matching tools.

Proteomics

Clinical proteomics focuses at the identification of protein disease biomarkers in biological fluids or tissue samples. Such biomarkers can function as indicators of abnormalities before clinical symptoms arise, enabling initiation of therapy at early stages. By unraveling the pathogenic mechanisms, these biomarkers might also facilitate therapeutic interventions by providing clues for targeted therapy and by reflecting the biochemical effect of new developed drugs. Analyzing the human proteome requires specialized quantitative proteomic technologies and expertise. Between different proteomic methods, variations exist concerning the range of molecular weight of proteins that are optimally detected. As a consequence, a combination of proteomic approaches are needed to identify and quantify the whole human proteome (2). There are different techniques for expression analysis of proteins, like two-dimensional electrophoresis (2-DE) and two-dimensional difference gel electrophoresis (2-D DIGE) combined with Matrix Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF-MS), Surface Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (SELDI-TOF-MS), LC-MS/MS and the isotope-coded affinity tags (ICAT) and isotope tags retention on chromatographic chip surfaces and their subsequent analysis by a simple laser desorption/ionization mass spectrometer (5).

For the SELDI-TOF-MS technique different ProteinChip arrays (Ciphergen Biosystems Inc.) are used. The chromatographic surfaces that make up the various ProteinChip arrays are uniquely designed to retain proteins from a complex sample mixture according to specific properties such as hydrophobicity or charge (5). By choosing different ProteinChip arrays with array-specific surface components, different proteins will be analyzed depending on the chip characteristics. In fact the interaction of the analyte and the chip introduces a purification step. Each combination of ProteinChip arrays together with the binding and washing buffers of choice results in a unique binding capacity for a special subset of peptides and proteins. After introducing the ProteinChip array into the SELDI-TOF-MS, a laser beam is directed onto the sample on the spot. Upon laser activation, the sample becomes irradiated and the desorption and ionization proceeds to liberate gaseous ions from the ProteinChip arrays. These gaseous ions enter the TOF-MS region of the instrument, which measures the mass-to-charge ratio (m/z) of molecular ions of each protein, based on its velocity through a vacuum tube (5).

Proteomic Platforms

SELDI-TOF-MS

There have been many reports on the application of SELDI-TOF-MS technology since its first introduction in 1993 by Hutchens and Yip (3). SELDI-TOF-MS is an approach that tries to overcome the requirements for purification and separation of proteins prior to mass spectrometry analysis (4). It is a novel approach to biomarker discovery that combines two powerful techniques: chromatography and mass spectrometry (MALDI-TOF). One of the key features of SELDI-TOF-MS is its ability to provide a rapid protein expression profile from a variety of biological and clinical samples (5). It consists of selective protein extraction and retention on chromatographic chip surfaces and their subsequent analysis by a simple laser desorption/ionization mass spectrometer (6).

For the SELDI-TOF-MS technique different ProteinChip arrays (Ciphergen Biosystems Inc.) are used. The chromatographic surfaces that make up the various ProteinChip arrays are uniquely designed to retain proteins from a complex sample mixture according to specific properties such as hydrophobicity or charge (5). By choosing different ProteinChip arrays with array-specific surface components, different proteins will be analyzed depending on the chip characteristics. In fact the interaction of the analyte and the chip introduces a purification step. Each combination of ProteinChip arrays together with the binding and washing buffers of choice results in a unique binding capacity for a special subset of peptides and proteins. After introducing the ProteinChip array into the SELDI-TOF-MS, a laser beam is directed onto the sample on the spot. Upon laser activation, the sample becomes irradiated and the desorption and ionization proceeds to liberate gaseous ions from the ProteinChip arrays. These gaseous ions enter the TOF-MS region of the instrument, which measures the mass-to-charge ratio (m/z) of molecular ions of each protein, based on its velocity through a vacuum tube (5).

Centraal Diagnostisch Laboratorium, Maastricht Universitair Medisch Centrum, Maastricht

Email: will.wodzig@mumc.nl

2-D DIGE

For many years, 2-DE has been the workhorse of proteomic research. 2-DE is a powerful technique, which allows simultaneous resolution of thousands of proteins. The principle of 2-DE is gel-based separation of a protein mixture in two dimensions. The first dimension of 2-DE is iso-electric focusing (IEF), during which proteins are separated based on their charge. Subsequently, the proteins are separated in the second dimension by SDS-PAGE according to their molecular weight. A drawback of classic 2-DE is the high degree of gel-to-gel variation that makes it difficult to distinguish any true biological variation from experimental variation. 2-D DIGE overcomes this problem by multiplexing and by using an internal standard. Multiplexing, the ability to separate multiple samples on the same 2D gel, is enabled by labeling each sample using spectrally resolvable, size and charge-matched fluorescent dyes known as CyDye’s, with a linear dynamic range over circa four orders of magnitude. The internal standard comprises a pool of equal amounts of each of the experimental samples being compared. So, it represents the average of all samples being analyzed and ensures all proteins present in the samples are represented. Each protein spot on a gel can be measured relative, as a ratio, to its corresponding spot in the internal standard present on the same gel. Protein spots of interest can be excised from the gels and subjected to sensitive and accurate mass spectrometry (MS), these mass spectral data are then used to interrogate databases to establish identity of the excised protein spots. An extensive overview of this 2-D DIGE technology has been described by Ünlü et al. (7) and Marouga et al. (8). Complementary to 2-D DIGE, SELDI-TOF-MS can resolve proteins below 20 kDa accurately according to mass in a high throughput manner.

Proteomic applications in chronic inflammatory diseases

The fast development and improvement of proteomics techniques has led to many studies looking for potential biomarkers in the field of oncology such as ovarian (9, 10), prostate (11) and lung (12) cancer as well as for chronic inflammatory diseases in human tissue and body fluids. The proteins we have detected and identified in different studies will be compared and discussed.

Biomarker discovery using SELDI-TOF-MS

The true scientific goal of serum proteomic pattern analysis is in fact biomarker discovery. However, since the study by Petricoin et al. (9) on proteomic patterns to detect ovarian cancer, the use of SELDI-TOF-MS protein profiling as a diagnostic tool, has become an important subject of investigation too (10). In our department, the SELDI-TOF-MS technique was used to detect potential biomarkers for two chronic inflammatory diseases; sarcoidosis, ankylosing spondylitis (AS), and to compare protein profiles of two types of lacunar stroke (PhD thesis J.A.P. Bons, March 28, 2008).

Sarcoidosis

Sarcoidosis is a multi-systemic inflammatory disorder, which affects the lungs in 90 percent of the cases. The main pathologic feature is chronic inflammation resulting in non-caseating granuloma formation. Until now there is no satisfying biomarker for diagnosis or prognosis of sarcoidosis. For detection of potential biomarkers, protein profiles of anion exchange fractionated serum of 35 sarcoidosis patients and 35 healthy controls are compared using SELDI-TOF-MS. An optimal classification is achieved with metal affinity binding ProteinChip arrays coupled with copper (IMAC-Cu⁺³ ProteinChip array). A single marker with a mass-to-charge (m/z) value of 11,955 results in a sensitivity and specificity of 86% and 63%, respectively. A multimarker approach of two peaks, m/z values of 11,734 and 17,377, results in a sensitivity and specificity of 74% and 71%, respectively. These sensitivities and specificities are higher compared to measurements of ACE and sIL-2R. Identification of the peak at m/z 17,377 results in the alpha-2 chain of haptoglobin (13).

Ankylosing spondylitis

AS is a chronic systemic inflammatory rheumatic disorder that primarily affects the axial skeleton, with spondylitis as its hallmark. Spinal structural damage can be assessed on conventional radiographs as destructive and proliferative lesions ultimately leading to syndesmophyte formation. Sera of 38 AS patients and 38 healthy controls are used to detect potential biomarkers. Analyses of all AS and healthy control samples on CM10 arrays results in a sensitivity of 66% and a specificity of 74% using a multimarker approach of two peaks. M/z 4,172 is used as first splitter in the decision tree and is up-regulated in the AS group and m/z 28,144 is used as second splitter. Analyses of all AS and healthy control samples on IMAC-Cu⁺³ arrays results in a sensitivity and specificity of 70% using a multimarker approach of two peaks. M/z 6,644 is used as first splitter in the decision tree and is down-regulated in the AS group and m/z 13,875 is used as second splitter. The peaks at m/z 28,144 and 13,875 are both successfully identified as apolipoprotein A-I (ApoA1) (14).

Lacunar stroke

Lacunar infarcts are small, deeply in the brain located infarcts, mostly caused by occlusion of a small perforating artery. Lacunar stroke patients in whom cerebral imaging shows only a single symptomatic lesion differ clinically from those patients with multiple additional “silent” lacunar lesions. Lacunar stroke patients with multiple lesions have more extensive cerebral white matter lesions on neuro-imaging, have more often hypertension, and have worse prognosis on functional outcome, a higher stroke recurrence rate, higher short- and long-term mortality and higher rate of asymptomatic lesion progression. Two groups are defined according to pre-defined criteria. Group 1 consists of 8 patients in whom brain magnetic resonance imaging (MRI) show only one single symptomatic lacunar lesion (type I). Group 2 consists of 8 patients in whom brain MRI additionally show multiple (4 or more) asymptomatic lacunar lesions as well as extensive white matter lesions (type II). An anion exchange procedure is used, which allows high-throughput fractionation of all 16 serum samples. The best distinctive pattern is
found on the IMAC-Cu$^{2+}$ ProteinChip array with denatured serum. One clearly potential marker at m/z 16,122 is up-regulated in type I vs type II with mean intensities of 12.5 and 5.0, respectively. Protein identification is performed by 1-DE and 2-DE followed by MALDI-TOF-MS. The peak at m/z 16,122 is identified as the alpha-chain of haptoglobin. The alpha-chain of haptoglobin exists in 2 variants, alpha-1 (8.9 kDa) and alpha-2 (16 kDa), the latter being compatible with the marker at m/z 16,122. The haptoglobin concentration and phenotype distribution are determined. As the total haptoglobin concentration does not differ between the two lacunar groups, the up-regulation of the alpha-2-chain in type I compared to type II represents a higher haptoglobin-2 allele frequency in the former. Yet, in comparison to the reference population, in both lacunar stroke groups haptoglobin-1 outweighs haptoglobin-2 allele frequency. The even higher haptoglobin-1 allele frequency in type II implies a promoting role for haptoglobin-1 in developing multiple silent lacunar lesions and cerebral white matter lesions (WML). The association between haptoglobin-1 and lacunar stroke brings in a new candidate gene in the study of genetic factors in cerebral small vessel disease etiology. The trend for a difference in haptoglobin-1 association between two lacunar stroke types could be a reflection of a difference in underlying vascular pathology (15).

Biomarker discovery using 2-D DIGE

Multiple sclerosis
Multiple sclerosis (MS) is a chronic inflammatory demyelinating disease of the central nervous system (CNS). Multiple sclerosis is a multicausal disease influenced by genetic predispositions, environmental factors and immunological mechanisms which damage the CNS (16-19). Multiple sclerosis presents itself for the first time as clinically isolated syndrome (CIS) in the majority of patients (≥ 85%). Patients with CIS will often (30-70%) develop clinically definite MS (20, 21). Diagnosing MS is based on the McDonald criteria; a combination of established clinical, radiological and laboratory analyses. The analyses comprise MRI detection of lesions in brain and spinal cord, abnormal visual evoked potentials, demonstration of intrathecal oligoclonal IgG and/or an elevated IgG index (22). However, their diagnostic value is limited in terms of predicting long-term clinical outcome and excluding other potentially treatable diseases (17, 23). Therefore, biomarkers with improved diagnostic and prognostic accuracy are needed as well as biomarkers for clinical sub-typing and monitoring disease progression. The cerebrospinal fluid (CSF) is considered a promising source of potential biomarkers for neurological diseases such as MS, as it is in direct contact with the brain interstitial fluid.

For the proteomic analysis, we have selected CSF samples from five relapsing-remitting MS (RR-MS) patients, five individuals with CIS and five control individuals who were diagnosed not to have a neurological disease. In total, hundred spots showed significant differential expression using 2-D DIGE between the three study groups and twenty-two of these hundred spots were identified using mass spectrometry. Proteomic analysis revealed significantly higher abundances of immunoglobulin peptides in CSF of MS patients which were confirmed using clinical routine diagnostics. IgG concentrations were higher in CSF of MS patients compared to both CIS patients and controls. Although transferrin spot abundances were significantly lower in MS patients compared to both CIS patients and controls, no significant differences in total CSF transferrin concentration were found between MS and CIS patients and controls. Furthermore, we investigated the glycosylation pattern of CSF transferrin in the three study groups. Both MS and CIS patients showed more asialotransferrin in CSF compared to controls (22.5 ± 4.3% and 24.3 ± 4.0% vs 16.6 ± 3.3%). We showed for the first time that asialotransferrin was increased in CSF of patients with MS, suggesting that the N-glycosylation of proteins in the CNS is altered (24).

Abdominal aortic aneurysm
Abdominal aortic aneurysm (AAA) is a degenerative disease of the abdominal aorta leading to progressive dilatation, intraluminal thrombus (ILT) formation and rupture. Two key features of AAA pathophysiology are inflammation and extracellular matrix (ECM) destruction. In 2009 our group published 2 reviews on existing and potential biomarkers of abdominal aortic aneurysm progression regarding inflammation and extracellular matrix degeneration (25, 26). Survival associated with AAA rupture is extremely low, with rates ranging from 10% to 15%, while elective surgery has an overall survival rate of 95%. It is widely recognized that the risk of rupture increases with increasing diameter, with the threshold for elective repair at ≥ 55 mm diameter (Dmax). However, a significant number of AAA ≥ 55 mm never experience rupture (27), while small AAA can rupture unexpectedly (28). Early detection and elective AAA repair is, for those at risk of rupture, a critical step to limit mortality associated with aneurysm rupture. Better approaches are needed for follow-up of patients with small asymptomatic AAA. Most patients with small AAA are prospectively followed by serial imaging until elective repair is indicated, although growth rate differs substantially between individuals. Besides identifying patients at risk for AAA progression and rupture new biomarkers may also help to elucidate the molecular mechanisms behind this disease and eventually generate treatments to reduce progression rate.

We studied a population consisting of 218 patients (189 men and 29 women) aged 71.7 ± 7.6 y and 69 age-matched controls (diameter aorta: < 30 mm). For biomarker analysis patients were divided into three groups based on their aneurysm diameter: small (30–44 mm; n=59), medium-sized (45–54 mm; n=64) or large (≥ 55 mm; n=95) AAA. The concentration of various biochemical markers, such as α1-antitrypsin (α1-AT), creatinine, cystatin C, haptoglobin, and IgG...
tend to be higher in AAA patients compared with controls. Both serum HDL and IgG concentration showed an inverse relationship with aneurysm size. Furthermore, higher serum creatinine and hsCRP values were observed in patients with a large aneurysm compared with patients with either a small or medium-sized aneurysm. Plasma concentration of MMP-9 was higher in patients with a large AAA compared with patients with small and medium-sized aneurysms. Linear regression modeling showed that age and hsCRP concentration were positively associated with aneurysm diameter, whereas HDL and IgG concentrations were negatively associated with aneurysm diameter. This model also indicates that in general men have larger aneurysms compared with women. These biomarkers may be useful in predicting AAA progression (29).

For the proteomic analysis we used 2-D DIGE together with tandem mass spectrometry (MS/MS) to analyse the serum proteome from patients with small (Dmax 30-54 mm) AAA, either stable (increase Dmax < 5 mm/year; n=8)) or progressive (increase Dmax ≥ 5 mm/year; n=8) and large (Dmax ≥ 55 mm; n=8) AAA. We showed that several proteins were differentially expressed in serum of small stable, small progressive and large AAA. To confirm the 2-D DIGE differential protein expression, quantitative validation of the results was carried out on serum or citrate plasma using a larger sample set (n=80). This study enrolled 48 consecutive patients with small (Dmax 30-54 mm) asymptomatic AAA, either stable (increase Dmax < 5 mm/year; n=27) or progressive (increase Dmax ≥ 5 mm/year; n=21), and 32 consecutive AAA patients undergoing elective open or endovascular repair (Dmax ≥ 55 mm). Significantly higher amounts of IgG, α1-AT and Factor XII activity were found in stable and progressive AAA compared to large AAA. Notably, serum IgG concentrations were elevated above the upper reference limit in almost all small AAA and were positively correlated with serum α1-AT concentrations. In almost all AAA patients, plasma D-dimer concentrations were elevated above the reference limit and were positively correlated with both expansion and diameter. Aneurysm size was significantly negative correlated with serum IgG, serum α1-AT and Factor XII activity and positively correlated with plasma D-dimer. Size combined with IgG, α1-AT or Factor XII activity had minimal effect on the prognostic value in predicting aneurysm progression compared to size alone.

In conclusion: The present study indicates that IgG, Factor XII and α1-AT are found in increased amounts in the serum of patients with stable or progressive small AAA. However, combination of either Factor XII or α1-AT with aneurysm diameter had little effect on prediction of aneurysm progression versus diameter alone (30).

**Limitations SELDI-TOF-MS and 2-D DIGE**

A growing body of evidence reveals the potential of clinical proteomics in biomarker discovery, but overfitting is a known phenomena. Data over-interpretation, due to multivariate statistical models in the analysis of large datasets, can lead to irreproducible results (31, 32). However, through double cross-validation, the chance of overfitting can be markedly reduced. The effect of pre- and post-analytical variables on protein profiling needs further and more systematic investigation. Therefore, a stringent standardized protocol is needed, not only for pre- and post-analytical aspects, but also for calibration and quality control performance. In our study a standard protocol for calibration of the MALDI-TOF-MS part of the PBS IIC instrument was defined and acceptance criteria for the independent certified QC samples were established. This is also possible for other instrument types. By checking the calibration every week, the QC procedure acts prospectively, while in some studies the quality control acts retrospectively by including the QC samples in the profiling experiments and in some studies there is no quality control procedure described at all (33).

Since reliability and stability of the ProteinChip arrays have been questioned worldwide, improvement of their quality and stability are needed urgently. Systematic investigations of the effects of various experimental variables on the SELDI-TOF-MS proteomic profiles are needed. With reliable ProteinChip arrays, optimized standard operating procedures, appropriate study design and experimental precautions, the proteomic profiling results should become more reliable (34).

Because the limited sensitivity of the SELDI-TOF-MS causes difficulties in the identification of potential diagnostic proteins present at concentrations below the mg/l level, advancements in the MS instrumentation part of the SELDI-TOF-MS technology or similar technologies, like tandem mass spectrometry (MS/MS), are needed and will lead to the identification of potential biomarkers at µg/l or even lower level. 2-D DIGE is less efficient in detecting expression differences for hydrophobic proteins, low molecular proteins (<10 kDa) and proteins at extreme pH’s. Also, 2-D DIGE gels can give rise to overlapping spots, leading to problems with quantification and identification of protein patterns. Currently, the pipeline from translation of new biomarkers into diagnostic tests appears to have a bottleneck. A number of technical obstacles, like the limited sensitivity and relative low reproducibility, remain before routine proteomic analysis can be achieved in the clinic (35, 36).

**Conclusion**

SELDI-TOF-MS has been used frequently, discovering many new biomarkers. However, reproducibility, validation and identification still remain important pitfalls concerning implementation of SELDI-TOF-MS in biomarker research. 2-D DIGE counteracts some of these limitations, by using an internal standard, major improvements are made concerning reproducibility. Furthermore, identification and validation of differentially expressed proteins is more accessible using 2-D DIGE, since differential spots can be excised directly from the gel. In comparison to SELDI-TOF-MS, 2-D DIGE also provides information on the molecular
weight and isoelectric point of the proteins. Differences in proteins can result from proteolysis, as well as from post-translational modifications (PTMs) that affect the charge and isoelectric point of a protein. In this overview, we focused on two complementary proteomic technologies and their capability to detect novel biomarkers for chronic (inflammatory) diseases and we showed that biomarker discovery in human biofluids is clearly one of the areas with enormous potential. Currently, the pipeline from translation of new biomarkers into diagnostic tests appears to have a bottleneck. A number of technical obstacles, like the limited sensitivities and relative low reproducibility, need to be improved before routine proteomic analysis can be achieved in the clinic. Standardization of methodologies and dissemination of proteomic data into publicly available databases is starting to overcome part of these hurdles. At present the most promising application for proteomics is the detection of specific subsets of protein biomarkers for certain diseases, rather than large scale full protein profiling. Directions for future research comprise specifically, PTMs including phosphorylation, glycosylation, oxidation, etc. which have been particularly of interest in the field as it has been demonstrated relevant to disease pathology and useful targets for therapeutics. Clinical proteomics has much promise in medicine with new types of proteomic Technologies combined with advanced bioinformatics currently being used to identify molecular signatures of diseases based on protein pathways and signaling cascades.

References


Breast cancer research on chip

F. WOLBERS, H.R. FRANKE, J.M. KLAASE, M. BRINKHUIS, A. van den BERG and I. VERMES

Recently, significant advances in the prevention, diagnosis and management of breast cancer have been made. Nevertheless, worldwide, every year, 460,000 women die of breast cancer (1). The conventional approach to cancer therapy is to provide treatment according to the organ or tissue in which the cancer originates. Currently, the selection of which breast cancer therapy is based on a broad scale of factors, including a patient’s age and tumour characteristics, such as nodal stage, the presence of oestrogen receptors and the Her-2/neu status (2). However, the various protocols that exist for chemo- and hormone therapy have different and limited rates of success. Often, this approach to cancer treatment is referred to as ‘trial and error’ or ‘one-size-fits-all’ (3). This practice is inefficient and frequently results in inappropriate therapy and treatment-related toxicity. In contrast, personalised treatment has the potential to increase efficacy and decrease toxicity. Nowadays, it is known that cancer develops as a result of multiple genetic defects and that individuals with the same type of cancer often have dissimilar genetic defects in their tumours (4). This finding explains why patients who seem to have similar cancers respond in a heterogeneous manner to antitumour agents and show clearly the huge obstacle to providing effective treatments for cancer. The hypothesis that stem cells play an important role in tumour biology receives a lot of attention (5). These so-called cancer stem cells (CSC) have the ability for self-renewal and are pivotal in setting the heterogeneous character of a tumour. Besides influencing the origin and growth of tumours, these CSC play an important role in developing metastasis. For personalised medicine, individual treatment regimes have to be set to define the best treatment possible for every patient. Currently, personalised treatment is most advanced for breast cancer. To achieve personalised treatment for cancer, (bio)markers for determining prognosis, predicting response to therapy, and predicting severe toxicity related to treatment are needed (3). DNA/RNA-microarrays for breast cancer prognosis, but also prediction, are very promising and at present clinical validation is ongoing (6). Recently, the use of microtechnologies for cell biology applications, and specifically for cancer, has received rapidly growing attention (7). Lab-on-a-Chip technology is a promising platform for personalised oncology to predict response or resistance to therapy, so that the individual patient receives the right drug. Even though results from in vitro assays can’t be directly and uniformly translated to the in vivo situation, the in vitro approach to determining drug sensitivity and resistance continues to have great potential to spare patients the morbidity of ineffective treatment. Here, the development of a microfluidic chip (‘Apoptosis chip’) to screen the effect of well-known antitumour drugs on human breast cancer cells is described. Moreover, preliminary results on cancer stem cells are shown.

Apoptosis chip
The microfluidic ‘Apoptosis chip’ consists of a main channel which broadens into a chamber for cell culture (figure lab). The ‘Apoptosis chip’ is made of the polymer poly(dimethylsiloxane) (PDMS) and combines cell culture, drug screening and apoptosis detection in one single device.

BIOS®, Lab on a Chip Group, MESA® Institute for Nanotechnology, University of Twente, Enschede, The Netherlands: Department of Obstetrics and Gynaecology® and Department of Surgery®, Medisch Spectrum Twente, Hospital Group, Enschede; Laboratory for Pathology East Netherlands®, Enschede

E-mail: f.wolbers@ewi.utwente.nl