ABSTRACT

Conventional cancer therapy relies heavily on genomics for the screening of patients, with a number of cancers being identified and classified through the presence of specific mutations. Tests for well-described, key genes along various biochemical pathways are used to assess the risk of malignancy, diagnosis and prognosis. However, such analysis is insufficient for many patients and needs to be supplemented with information related to the downstream protein phenotype, and which is more patient-specific. A comprehensive understanding of the changes in the proteome over the progression of cancer, from the pre-malignant to the metastatic phase, is required. Proteomic technologies are slowly becoming more accessible and are being employed at various stages of cancer therapy. Such techniques are capable of providing patient-specific data which improves decision-making. Various proteins have been proposed as biomarkers for specific malignancies either for the screening of tumour biopsies or serum proteome as a form of non-invasive testing. Moreover certain changes in protein properties are linked to tumour-specific aberrant post-translational modifications e.g. abnormal protein activity due to
phosphorylation. In the very near future cancer therapy will rely on a combination of mutational data from the current, standard genomic technologies and data from other levels including revolutionary proteomic technologies integrated into a patient-centred system for the diagnosis, sub-classification, and administration of individualized therapies.

**Keywords:** Proteomics; Biomarkers; Mass spectrometry; Protein microarrays; Post-translational modifications

**Abbreviations:** LCM: Laser-Capture Microdissection; MS: Mass Spectrometry; PTM: Post-Translational Modification; 2D-PAGE: Two-Dimensional Polyacrylamide Gel Electrophoresis; DIGE: Differential In-Gel Electrophoresis; MudPIT: Multi-dimensional Protein Identification Technology; SELDI-TOF: Surface-Enhanced Laser Desorption Ionization Time-of-Flight; MALDI-TOF: Matrix Assisted Laser Desorption Ionization Time-of-Flight; FPPA: Forward Phase Protein Arrays; RPPA: Reverse Phase Protein Arrays; FRET: Fluorescence Resonance Energy Transfer; MRI: Magnetic Resonance Imaging; REIMS: Rapid Evaporative Ionization Mass Spectrometry

**INTRODUCTION**

At present, genomics holds the top spot in the clinical setting due to the ease and accessibility of genetic techniques but unfortunately this provides an incomplete analysis and limited success. Gene arrays are routinely used to classify patients by using either cancer stage or survival outcome [1-3] but are limited in that these offer no information as to the functional properties of the encoded proteins.

Cancer therapy in the past decade has advanced at a fast pace, expanding from single gene analysis, through high-throughput genomics and transcriptomics, up to proteomics among various other biomolecular analysis such as metabolomics and lipidomics.

Proteomics, in the broadest sense, involves the comprehensive analysis of changes in protein expression patterns due to post-translational modifications (phosphorylation, ubiquitination, acetylation and methylation among many others), sub-cellular localization, protein–protein interactions, changes instructural conformations and protein function.

This expansion of analysis to incorporate proteomics is based on increasing insight that although key aberrations in cancers may be genetically encoded, the final downstream effect and biochemical phenotype occurs through changes in protein function. This stems from the fact that gene expression assays give no indication of protein localization, half-life or turnover, post-translational modifications, protein–protein interactions or functional activity, which is crucial for therapeutics.

Thus while genetic mutations may encode growth advantages, metastatic properties and resistance to therapy, it is aberrant protein activity, both within the tumour and in its surrounding microenvironment, that ultimately drives the process of carcinogenesis [4]. Simply identifying the genetic changes involved in carcinogenesis are not enough for clinical application and successful
treatment of patients and a concomitant understanding of the functional consequences of the disease at a proteomic level is also required.

In an effort to improve cancer therapy, a comprehensive understanding of the changes in the proteome over the progression of cancer, from the pre-malignant to the metastatic phase, is required. A number of key proteomic technologies are being applied to cancer screening and the development of new drugs and therapies. This has led to numerous potential biomarkers having been proposed for diagnosis, prognosis and drug selection. Based on these advances, a view beyond the simple total protein levels and into protein activity combined with other properties controlled by post-translational modifications is being applied to provide patients with more suitable treatment regimens.

PROTEOMIC TECHNOLOGIES

Great effort is being put into the development of proteomic technologies that improve the systematic analysis of thousands of proteins in parallel. This needs to be performed in a fast, inexpensive and low-sample-volume format, in a standardized manner in order to achieve reliable results that are comparable between different instruments and operators [5]. Such technologies are an asset for cancer therapy from basic biological research, for the development of a better understanding of the disease process, and the identification of novel therapeutic targets.

Laser-Capture Microdissection

When a biopsy sample is prepared for proteomic analysis, it is important to isolate the tumour cells from the non-cancerous tissue material, as the tumour cells of interest generally comprise only a small proportion of the biopsy material [6,7]. Furthermore, the tumour cells may be comprised of sub-populations, which need to be divided. The inclusion of non-cancerous cells (presenting a different biochemical phenotype) would reduce the significance of the results by shifting the outcome towards the normal biochemical situation.

In order to achieve meaningful results it is necessary to isolate the desired tumour cells from the rest of the biopsy material. One of the most useful techniques for such isolation is Laser-Capture Microdissection (LCM) [6,8]. The general set up involves a UV laser coupled with a microscope. A biopsy section mounted on a slide is loaded onto the microscope stage and then using the microscope software interface, the laser is set to cut out the desired cancer cell selection from the biopsy section. Once the cells are cut out, they are separated from the adjacent biopsy material and extracted for downstream processing [6].

This dearth of sample material means that micro-proteomic technologies need to be developed that can employ such limited amounts of cellular material. The technology that has been most useful in the field of proteomics is mass spectrometry, which has been applied in a variety of modes.
Mass Spectrometry

Mass Spectrometry (MS) covers a set of techniques that provide extremely rapid and high-throughput methods for proteomic applications. These can be divided crudely into two major categories: high-resolution for protein identification and low-resolution for proteome scanning (Figure 1).

![Diagram of mass spectrometry systems]

**Figure 1:** The two types of mass spectrometry (MS) systems: low-resolution for high-throughput proteome scanning and phenotyping (left) and high-resolution for protein identification and single protein quantification (right).

High-resolution MS platforms provide high-speed identification of purified microgram quantities of biopsy or serum proteome components. The sensitivity is such that it allows for the identification of highly similar protein molecules such as isoforms and more recently even the same proteins harbouring different PTMs that alter proteins by as little as a methyl group, causing minute changes in mass and depending on the chemical group, charge. Such MS platforms therefore assign an accurate mass-to-charge (m/z) ratio fingerprint to a specific proteomic component in a patient specimen, with the resolution and mass accuracy of the instrument used being critical to the complexity of that fingerprint.
In order to visualize a broader sub-set of the proteome, proteomic pattern analysis is applied. Despite the usefulness of such analysis, the combination of multiple fingerprints from high-resolution MS platforms is technologically demanding and extremely laborious. Traditional proteomic pattern analysis or profiling is based on comparative analysis using Two-Dimensional Polyacrylamide Gel Electrophoresis (2D-PAGE) or the more recent variation, Differential In-Gel Electrophoresis (DIGE) [9], which overcomes limitations due to inter-gel variation by running dye-labeled samples on the same gel, directly comparing normal and disease samples. A non-gel-based technique, Multi-Dimensional Protein Identification Technology (MudPIT) [10], uses 2D liquid chromatography to separate complex protein mixtures and identify individual components.

Proteomic profiling using MS platforms such as Matrix Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) [11] or its variant Surface-Enhanced Laser Desorption Ionization Time-of-Flight (SELDI-TOF) [12], can generate unique protein barcodes (a complex low-resolution fingerprint of ion peaks) from small amounts of human serum or biopsy material [13,14], such that the pattern itself, independent of the protein identity, is used to discriminate between normal and diseased samples. The elimination of the need to use identified proteins improves the speed of application for changes in diseased samples and may thus be clinically useful at different stages of cancer therapy [15].

The drawback with MS techniques, and most other proteomic technologies that can work at such a scale, is that protein samples cannot be analysed in their native form. The preparation processes denature the proteins, breaking protein complexes apart, destroying the three-dimensional conformation, and modifying protein residues, such that these methods make it impossible to quantify protein activity [16]. For this reason protein arrays are preferred to analyse enzymes or signalling molecules which may present different activity due to PTMs, without any change in the total protein levels.

**Protein Arrays**

Protein arrays are chips composed of a supporting surface onto which proteins are immobilized either indirectly (via antibodies) in Forward Phase Protein Arrays (FPPA) or directly in Reverse Phase Protein Arrays (RPPAs) and probed to quantify protein activity (Figure 2).
Figure 2: Workflow for the analysis of proteins within biological samples using protein arrays, which consist of chips onto which proteins are immobilized either indirectly (via antibodies) in Forward Phase Protein Arrays (FPPA) (left) or directly in Reverse Phase Protein Arrays (RPPAs) (right), ideal for the quantification of protein activity.

Protein activity screens using FPPAs allow the simultaneously screening of enzymes on a variety of short amino acid sequences as substrates, with immobilization making the substrates much more sensitive to PTM addition [17,18]. The interactions translated into protein activity can be quantified by incorporating bioluminescent reporter molecules into the assay, thus enabling combination with techniques such as Fluorescence Resonance Energy Transfer (FRET) [19].

RPPAs [20-23] tend to be the preferred modality because the lysates of small numbers of microdissected cells from biopsies by LCM can be denatured and immobilized on a chip and specific proteins or PTMs can be detected simultaneously in a large panel of protein lysates by probing with validated PTM-specific antibodies (one antibody per array slide) allowing for the
detection of PTM-specific signalling-related aberrations [24,25]. The information provided by such protein arrays is then used in the selection of appropriate pathway inhibitors.

**PROTEIN BIOMARKERS**

The use of clinical proteomic profiling by both MS platforms and protein arrays is an integral part of patient-centred molecular medicine. These technologies have enabled the generation of a long list of potential candidate biomarkers to aid decision-making at a variety of stages throughout cancer therapy including diagnosis, prognosis and treatment of cancer [26]. Among other areas of application are drug discovery and development, the analysis the desired drug effects on target pathways and the unwanted toxic effects on normal cells, the monitoring of patients progress through therapy (mainly through serum proteomic pattern analysis) and finally the determination of the therapeutic outcomes for changing the choice therapy [27] (Figure 3).

**Figure 3:** Applications of biomarkers throughout the various stages of cancer therapy. Biomarkers aid in decision-making at the critical stages of diagnosis, prognosis and treatment of cancer, with input from other areas on which they are heavily dependent such as drug discovery and protein activity testing.

A biomarker can be any biological molecule that allows the identification of a particular physiological or pathological state. In this review the term will be narrowed down to include only proteins. The ideal biomarker should be sensitive and specific for the protein of interest in an assay that should be cost-effective, fast and robust against both inter-operator and inter-institutional variability. Fora biomarker to be regarded as reliable, it needs be validated in controlled clinical studies of a wide selection of patients using rigorous standards for all steps in the process from sample collection, up to result analysis to make them reproducible between different laboratories [5].
By now it has become clear that a single biomarker for a cancer type does not exist as there is considerable heterogeneity in the proteome of patients in combination with a variety of underlying disease processes. However, so far, most biomarker discovery efforts employ laborious searches for one or a small number of up or down regulated proteins in cancer samples. As a result, panels of biomarkers make better clinical tests to cover a range of proteins related to aberrant protein presence, function or interactions.

A significant benefit for clinical application is if a protein biomarker can be quantified in an easily accessible body fluid such as serum, urine, or saliva. Serum is a unique source of protein information, as it contains traces of various biological processes that the blood is exposed to as it flows around the body. Cancer-related changes in the serum proteome include the increased or decreased release of intact, cleaved or modified proteins, making it well-suited for the discovery of biomarkers [7].

Protein biomarkers include cell surface receptors, tumour antigens, PTMs (mainly phosphorylations), glycoproteins (cross-reactive carbohydrate determinants), and peptides released by tumours into serum or body fluids. Subsequently, patterns of biomarkers, particularly in serum, might prove more selective and potentially more useful than individual biomarkers [5].

Diagnostic Biomarkers

Detecting cancer at an early stage, particularly in the pre-malignant state, has a profound effect on improving the success rate of treatment strategies. However, unfortunately, in a number of cases, early stage cancers lack specific symptoms and there are no accurate and reliable diagnostic biomarkers, in the form of non-invasive tests to identify a large proportion of malignancies. When diagnostic biomarkers are available, they can be sub-divided into three main categories based on their use for classification, grading and staging.

Classification

Classification of a cancer by tissue of origin is the first step towards the assignment of a prognosis and choosing the most suitable therapeutic regimen. Protein biomarkers are generally not required at this stage since the anatomical location of a tumour usually indicates its tissue of origin, while histological examination further confirms the diagnosis and identifies the tumour sub-type [5]. However, biomarkers aid in the differential diagnosis, to distinguish cancer from other diseases that present similar symptoms as well as in classifying tumours of unknown primary location, to determine the site of origin and map the evolution of a cancer as it progresses [28].

Grade

Grading systems are generally designed to classify tumours by histology, based on the degree of differentiation, where low-grade, well-differentiated tumours are usually less aggressive and present a more favourable prognosis than high-grade, undifferentiated tumours, which tend to
grow faster and metastasize earlier [5]. Inherent subjectivity, which is the major drawback of grade assignment, can be eliminated by relying on biomarker patterns to correctly score tumours in accordance with pathologist-assigned grades [29]. Through such a process, the assessment of histological grade using biomarkers increases the utility of grading for predicting the response to therapy and selecting appropriate treatment.

**Stage**

Cancer staging consists of four divisions and sub-divisions thereof, with the primary purpose (in solid tumours) being to discern if the cancer is localized or metastatic, which is crucial in predicting survival and the choice of first-line therapy. Conventional staging integrates information from sources such as physical examinations, imaging (X-rays, magnetic resonance imaging - MRI, etc.) and biopsies with the aim of determining the extent of disease through visual and histological information [5]. Proteomics can be employed for staging in the form of molecularly-targeted functional imaging, which involves the use of imaging modalities such as MRI to visualize cancer biomarkers [30].

**Prognostic Biomarkers**

Tumour classification, grade and stage are ultimately used to define the prognosis of a patient’s cancer and select the best therapy. Biomarker expression generally complements, but in some cases supersedes traditional clinical tumour classification, staging and grading when biologically-targeted therapeutics are being evaluated [5]. Both prognosis and predicted response to therapy are needed in the decision for the selection of neoadjuvant (given prior to surgery to shrink the cancer) or adjuvant (given after surgery to prevent recurrence) chemotherapy.

**Biomarkers in Drug Discovery and Design**

The screening of pathways for the selection of treatment strategies has been widely performed using pharmacogenomics, which relies on gene arrays to produce molecular profiles. However, proteomic pathway analysis of biopsy material should be used instead since most drugs target proteins and alter their characteristics, such as protein activity [16].

There is currently an intense focus by the biotechnology and pharmaceutical industries in applying high-throughput proteomic technologies for drug design and drug target discovery efforts as the proteome is too vast to screen for individual proteins [31]. Moreover, in order to successfully develop protein drugs (recombinant proteins or antibodies) and effective combinations of pharmacological or biological inhibitors that can specifically target key classes of molecules, a thorough understanding of the combination of proteins dysregulated along a regulatory network, is of major importance [16]. Similarly, the determination of selected interactions between proteins and other biomolecules, be it other proteins, nucleic acids, ligands or synthetic compounds, are crucial for the discovery and validation of novel therapeutic targets [32].
Therapeutic Biomarkers

Biomarkers that confer information about the response to therapy, for the selection of first-line therapy and change in therapy following the development of drug or chemoresistance, are another important category of biomarkers. As each tumour appears to have a unique selection of proteomic aberrations, it follows that a given class of therapy would only be effective in the subset of patients having susceptible proteomic aberrations. This justifies the use of a therapeutic strategy that selects from a range of treatment choices or combinations, to best counter the protein dysregulation within a tumour [33-37]. This requires the identification of critical proteins along a pathogenic pathway and the building of functional interaction maps for each tumour.

A collection of proteomic technologies have to be applied to gather information regarding which pathways are being dysregulated and used by the cancer cells and the surrounding stroma, and follow the changes as a consequence of the treatment given. The key point is the movement away from conventional pharmacological cancer therapy, which has been directed at a single upstream regulator as a drug target and towards a combinatorial therapy in which multiple steps along the pathogenic signalling pathway are targeted by a mix of drugs and protein inhibitors [38-41]. This is a critical shift in cancer therapy because in order to effectively shut down the entire pathway by blocking only a single upstream target, the high drug concentration required produces unwanted toxic side effects by affecting off-target molecules. However, using a combination of drugs and protein inhibitors targeting multiple steps along the pathogenic signalling pathway ensures increased efficacy through higher specificity, with a decrease in toxicity and side-effects due to lower doses [42-46].

The mapping of aberrant signalling pathways enables the determination of the core set of interconnected proteins that make the best drug targets. Protein kinases are one example of key molecules that make up signalling nodes, and their aberrant function is at the centre of numerous diseases, including cancer [47-49], making them popular drug targets [50-52].

The application of RPPAs to screen for protein activity can be used to monitor the disease state or response to therapy by comparing the level of activity of a protein at different stages of disease progression, or before and after treatment. This would improve patient-specific therapeutic strategies through a more effective selection of inhibitors. However this requires the development of high quality antibodies that are specific for the PTM son target proteins within key pathways [53].

POST-TRANSLATIONAL MODIFICATIONS

There is a wide variety of Post-Translational Modifications (PTMs) and among the more relevant, phosphorylation and acetylation are associated with protein activation, while methylation can alter most protein characteristics depending on the cellular context [54]. In cancers, aberrant signalling is generally not via a change in total protein levels but through a
change in protein activity encoded by PTMs on signalling proteins and transcription factors. This section will only focus on phosphorylation as this has been the most well-studied.

The human proteome is estimated to contain several hundred thousand potential phosphorylation sites, making up what is referred to as the phospho-proteome. This offers a huge selection of PTM combinations for protein regulation, producing a specific response to changes such as those induced by various cancer therapies [55,56]. The phosphorylation status of various protein classes is specifically regulated by selected kinase (that phosphorylate) and phosphatase (that dephosphorylate) enzymes acting at specific sites. However, kinase enzymes are known for only a third of all phosphorylation sites [57]. The high level of complexity offered by so many sites and enzymes makes the quantification of specific phosphorylation states along a pathway of interest for cancer therapy very difficult.

The high-throughput analysis of the phospho-proteome has been made possible by the development of techniques for the isolation of phosphorylated peptides from complex protein mixtures [58-60] and the combination with RPPAs [20,61]. This has facilitated the evaluation of the effects of specific cancer therapies on the phospho-proteome [62-65] by investigating biopsies from solid tumours using RPPAs [66]. The drawback of such analysis is that only previously identified phosphorylation sites can be analysed since validated antibodies need to be used.

Moreover, the phosphorylation status of a protein is not always representative of its activity since multiple phosphorylations can act together and additional PTMs (particularly methylation) can affect the net protein activity by additive effect or completion with phosphorylation [54].

CONCLUSION

Clinical proteomics thus has several important direct bedside applications for the detection, treatment and management of cancer. Early detection of cancer using serum proteomic would definitely increase the chances of a positive outcome by a non-invasive test. Diagnosing and classifying tumours correctly using reliable molecular biomarkers, at an early stage would increase the options for treatment. High-throughput array-based proteomics can screen for enzymatic activity and assist in the selection of patient-based selection of therapeutic regimens to target dysregulated nodes in disease-specific protein networks. Such technologies also prove useful in assessing therapeutic efficacy and toxicity in real-time such that therapy can be adjusted accordingly especially following drug or chemoresistance.

As proteomic technologies advance and the potential of clinical proteomics expands, the applications and benefits for patients will increase. The real-time analysis of proteomics data and integration of proteomics with other omics technologies will further improve the clinical application of advanced technologies and improve patient outcome as has already been demonstrated by the use of the Rapid Evaporative Ionization Mass Spectrometry (REIMS) [67] for the real-time analysis of phospholipids, directly during surgery.
References


