Haematological Malignancies

Proteomics in Myeloproliferative Disorders

a report by

Pascal Mossuz

Haematology Laboratory, University Hospital Grenoble

DOI: 10.17925/EOH.2008.02.1.78

Myeloproliferative disorders (MPDs) are haematopoietic stem cell malignancies characterised by overproduction of one or more mature myeloid lineage. The term ‘myeloproliferative syndrome’ was first introduced in 1951 by Dameshek, who defined chronic myeloid leukaemia (CML), polycythaemia vera (PV), essential thrombocythaemia (ET) and idiopathic myelofibrosis (IMF) as a group of closely inter-related neoplasms caused by proliferative activity of the bone marrow cells.

Chronic myeloid leukaemia was the first neoplasia to be associated with a specific chromosomal abnormality, and became the paradigm for targeted therapy. The central role of the tyrosine-kinase activity and fusion of the oncprotein BCR-ABL in reciprocal translocation between chromosomes 9 and 22 – t(9;22)(q34;q11) – made it a suitable drug target. It is now the front-line CML therapy.

The molecular mechanisms of PV, ET and IMF remained elusive. The clonal nature of these disorders was objectified by their capacity to form endogenous colonies without haematopoietic factor. In 2005, 50 years after Dameshek’s definition, several groups identified the molecular basis of these MPDs as a JAK2-V617F ‘gain-of-function’ mutation. The role of the JAK2-V617F mutation has been widely documented, along with its critical role in the pathogenesis of MPD-Phi.

The term ‘proteome’ was proposed by Wilkins et al. in 1995 and referred to the whole set of proteins present in a cell or a biological fluid at a given time. The proteome is a fundamentally dynamic entity that reflects the best functional status of a biological system. There is increasing evidence that the large range of gene products is not necessarily driven by an alteration of levels of transcript. Due to messenger RNA (mRNA) alternative splicing, allelic variations and post-translational modifications, a genome can produce over a million protein species. The proteome represents a dramatic expansion and alteration of the genome message, and as such is a very promising field of investigation as a complement to mRNA-based measurement and may achieve significant insights into complex biological systems. Proteomics refers to the comprehensive study of the proteome, including detection, identification, dosage and characterisation of protein modifications. Proteomic methods have been available in clinical laboratories for several decades, but have been limited to the evaluation of only a few proteins at a time.

In the past few years, mass spectrometry (MS) has emerged as an efficient alternative to 2D electrophoresis. This has considerably increased the relevance and accuracy of proteomics. Numerous technical advances have led to MS-based methods as powerful tools for the identification of proteins in complex mixtures. Tandem-MS (nano-LC MS/MS) has been applied to characterise the proteome of several organisms, organelles and multiprotein complexes. Progress in genome sequencing and better, faster bio-informatics tools have increased the potential of proteomics as a powerful analytical tool. Comparison of protein profiles, obtained from biological fluids or tissue sections, reveals an informative diagnostic approach for tumour discrimination and for the discovery of the diagnostic markers of solid tumours. In the field of haematology, proteomics has allowed the investigation of proteome modifications associated with different haematological neoplasms, identified therapy-related proteomes and predicted clinical behaviour in adult lymphocytic leukaemia (ALL). Evidence that some solid tumors and AML are associated with quantitative changes of serum proteins has been published.

This review illustrates the ability of proteomic approaches to gain insight into MPD pathophysiology in three ways:

- research for biomarkers;
- investigation of molecular mechanisms of drug resistance; and
- phenotypic characterisation of patients.

Proteomics Insights into Myeloproliferative Disorders

Serum Biomarker Identification

The fact that the allelic ratio (JAK2-V617F/JAK2-wild-type) can differ between patients, in the same patient between cells and during disease follow-up suggests that the JAK2-V617F burden may have a clinical impact. It was initially suggested that the haematological and clinical parameters in PV and ET may be, in part, dependent on JAK2-V617F.
V617-F status. Quantitative polymerase chain reaction (PCR) determination of the JAK2-V617F load provided contradictory results about the relationship between JAK2-V617F status and both haematological parameters and vascular complications. In this context the identification and correlation of serum markers, proportionally to mutated alleles, may provide a useful tool for the characterisation of MPD JAK2-V617F+ patients at diagnosis.

The protein-chip surface, enhanced, laser desorption/ionization, time-of-flight (SELDI-TOF), MS-based method of analysis had proved itself useful for the research of serum proteins in patients carrying the JAK2-V617F mutation. This technology, developed by Ciphergen Biosystems Inc., uses chip-based protein sample assays with different chromatographic surfaces designed to capture and retain subsets of proteins based on specific protein characteristics such as affinity, charge, hydrophobicity and metal-binding capabilities. It provides a surface-based enrichment prior to MS analysis. In this study, two subgroups of PV patients could be distinguished on the basis of their protein profiles. The distinct protein profiles demonstrated a significant variation in the mean percentage of the JAK2-V617F load (p=0.02). PV patients harbouring different proportions of JAK2-V617F alleles (>50%, 51–74%, ≥75%) displayed significant variations of serum protein expression in parallel with the modification of PRV-1 transcript levels, platelet numbers and haematocrits.

The data suggest that the JAK2-V617F load had a significant impact on the serum level of some proteins. Indeed, the authors isolated a 28kDa biomarker, identified by peptide fingerprint as apolipoprotein-A1 (Apo-A1), which highly correlated with the percentage of JAK2-28kDa biomarker identified by peptide fingerprint as apolipoprotein. Indeed, the authors isolated a transcript levels, platelet numbers and haematocrits.

The protein-chip surface, enhanced, laser desorption/ionization, time-of-flight (SELDI-TOF), MS-based method of analysis had proved itself useful for the research of serum proteins in patients carrying the JAK2-V617F mutation. This technology, developed by Ciphergen Biosystems Inc., uses chip-based protein sample assays with different chromatographic surfaces designed to capture and retain subsets of proteins based on specific protein characteristics such as affinity, charge, hydrophobicity and metal-binding capabilities. It provides a surface-based enrichment prior to MS analysis. In this study, two subgroups of PV patients could be distinguished on the basis of their protein profiles. The distinct protein profiles demonstrated a significant variation in the mean percentage of the JAK2-V617F load (p=0.02). PV patients harbouring different proportions of JAK2-V617F alleles (>50%, 51–74%, ≥75%) displayed significant variations of serum protein expression in parallel with the modification of PRV-1 protein expression, parallel with the modification of PRV-1 transcript levels, platelet numbers and haematocrits.

The data suggest that the JAK2-V617F load had a significant impact on the serum level of some proteins. Indeed, the authors isolated a 28kDa biomarker, identified by peptide fingerprint as apolipoprotein-A1 (Apo-A1), which highly correlated with the percentage of JAK2-V617F alleles. The highest levels of Apo-A1 specifically identified PV patients with ≥75% mutated alleles, suggesting that the Apo-A1 assay might provide helpful information for the characterisation of PV patients at diagnosis. The JAK2-wild-type was shown to selectively modulate the apolipoprotein interaction with the ATP-binding cassette transporter-A1 (ABC-A1), which mediates the transport of excess cholesterol from cells to apolipoprotein. Apo-A1 may also stimulate JAK2 auto-phosphorylation, which in turn activates apolipoprotein interaction with ABC-A1. Modifications of Apo-A1 levels might reflect alterations of cholesterol metabolism in MPD patients that may participate in vascular complications.

**Molecular Mechanisms of Drug Resistance**

Proteomics was revealed as particularly effective in elucidating the mechanism of imatinib resistance in CML. Although most patients show excellent responses to imatinib treatment, clinical resistance may develop. Resistance to glivec is now the main concern in the successful treatment of CML patients. Resistance is predominantly caused by point mutations in the ABL-kinase domain, but may also be mediated by BCR-ABL-independent mechanisms, including drug sequestration, or overexpression of the chemoresistance protein. Investigation of imatinib resistance is a critical issue. Proteomic approaches, based on comparative proteome profiling of cell lines, demonstrate an effective method of identifying candidate proteins implicated in the regulation of therapy sensitivity.

Comparative proteome analysis of CML cell lines (resistant versus sensitive) combined with LC-MS/MS peptide analysis of differentially expressed proteins (from 2-D gel electrophoresis) implicate different functional classes of protein, including RNA and DNA repair enzymes, cell structures and metabolised and heat-shock proteins.

The role of the overexpression of Hsp-70 in drug resistance was highlighted in two papers using two different cell line models. They further highlighted a role for Hsc-70 in the regulation of resistance through the regulation of Hsp-70 levels. Downregulation of heterogeneous nuclear ribonucleoprotein (hnRNP) contributed to the stability of mRNA. The Bcr-Abl fusion gene was also associated with a significantly increased resistance. It is likely that downregulation of the transcriptional repressor activity of hnRNP allows the expression of gene-implicated sensitivity to imatinib.

The emergence of imatinib resistance has led to the synthesis of a second generation of BCR-ABL inhibitors, approved for the treatment of patients with imatinib resistance. Clinical observation has highlighted different side effects and a patient-dependent pattern of sensitivity, which suggests differences in the molecular mechanisms of action. Biochemical studies have already revealed differences in kinase target engagement by different inhibitors. Global chemical proteomic approaches allow further research into the molecular action mechanisms of these drugs and the identification of a drug’s efficiency and side effects. For example, affinity purification of drug-related proteomes from CML leukocytes showed strong differences in proteome profiles between the three inhibitors and highlighted a potential impact of dasatinib on immune function. This study provided the first proof of implicating a non-kinase target (the oxydo-reductase NQO2) of these drugs.

**Phenotypic Characterisation of Myeloproliferative Disorder Patients**

The possibility to characterise a patient through his or her proteome profile has provided the opportunity to complement phenotype investigations of MPD patients. Clinical studies initially suggested that the presence of the JAK2-V617F mutation influenced the ET phenotype. Compared with V617F-negative ET, V617F-positive ET displays haematological characteristics that could assimilate it to pre-PV with a continuum between V617F-positive ET and PV.

A proteomic approach was used to evaluate the impact of the JAK2-V617F mutation on the ET phenotype. MS analysis was applied for the characterisation of ET serum protein profiles, which were then compared with PV. Significant differences in serum protein profiles between V617F-positive ET and PV were found. The presence of the JAK2-V617F mutation did not impact significantly on the MS serotype of ET patients and there were very few significant differences between the serum protein profiles of V617F-positive and -negative ET. Clustering of ET patients on the basis of their serotype did not correlate with their JAK2 status.

In ET, the MS protein profile distinguished two subgroups that did not display a significant difference in JAK2 burden, but did tend to have different polymorphonuclear cell counts. These results strengthen the argument that suggests that the JAK2 load weakly impacts the ET phenotype. Phenotypic heterogeneity exists in ET (consistent with previous clonality studies), and this heterogeneity is not directly linked to the JAK2-V617F mutation status. Phenotype characterisation of MPD through proteomic profiling may be a promising way to investigate the relationships between the MPD phenotype and genotype.
Haematological Malignancies

New Challenges
Proteomics has evolved to be an accurate strategy that provides informative data on disease pathogenicity. Three new specific challenges for proteomics in future clinical are:

- mining the proteome;
- improving MS-based quantitative proteomics; and
- interpreting the protein language using proteomics.

Mining the Proteome
One major problem with proteomic studies, regardless of the experimental approach, is the wide concentration range of proteins in biological samples. This is particularly true in body fluids, where the concentration range of proteins spans 10 or even 12 orders of magnitude. The most abundant protein, human serum albumin, constitutes over 50% of the total protein content in plasma. This is the reason why most proteomics studies publish non-specific biomarkers, which are more likely to be linked to the inflammatory response than to the pathogenic process. Given this, there is a clear need to improve the pre-analytical preparation of samples in order to detect low-abundance proteins and identify specific plasma or serum cancer biomarkers.

To overcome the limited range of MS detectors, several candidate technologies have been proposed. Depletion of the most abundant proteins, such as albumin, transferrin and immunoglobulins, by immobilised dyes or immuno-affinity may be the most intuitive strategies. While this may give rise to interesting results, it may also deplete the peptides (or low-molecular weight proteins) at the same time. These peptides are the best candidates for bio-marking in oncology.

Two strategies focus on the interest and effort of a proteomics platform. The first performs an extensive fractionation (anion exchange fractionation, reversed-phase chromatography) before analysis of individual fractions. The reduction in complexity of individual fractions allows for the discovery of low-abundance proteins further to LC-MS/MS. Selection of a subset of proteins by affinity chromatography using, for example, multiple lectins to capture glycoproteins, allows a decrease in the sample complexity. One pitfall of this strategy is that it relies on the persistence of major proteins in some fractions that do not modify the concentration range.

The second strategy consists on enriching samples in low-abundance proteins while decreasing the concentration of the most abundant. The goal is to obtain, through an equalisation-bead-technology and using a library of combinatorial ligands: ‘fixed-on’ beads. Highly specific affinity interactions between proteins and ligands ensure the binding of the greatest number and variety of proteins. The result is the concentration of rare species and the dilution of highly abundant species that narrow the concentration range of proteins present in the sample. Montasarrat and colleagues applied a highly purified fraction, obtained with equalisation bead technology (nanoLC-MSMS) to a LTQ-Orbitrap MS. They could confidently identify as many as 1,578 proteins in the cytoplasmic fraction of erythrocyte, which allowed a deep exploration of the classic red blood cell (RBC) pathway and the identification of unexpected minor proteins. Platelet proteome was also significantly explored.

The main potential pitfall of this method relates to its impact on relative concentrations of proteins. Modification of concentrations of proteins by this technique could modify the protein ratio, and in doing so impede proteome comparison between different clinical or biological stages. Preliminary studies on cerebrospinal fluid have shown that quantitative information and the relative ratio between different samples were maintained after equalisation.

Improving Mass Spectrometry-based Quantitative Proteomics
MS is not inherently quantitative and allows for, at best, a semi-quantitative evaluation of protein based on the comparison of the peptide peak intensity between two different source samples. Quantification of proteins in parallel with their identification is necessary to produce proteomic clinically relevant data.

The ‘stable isotope labelling’ method remains the core technology for proteomic quantification. Interest in chemical labelling for quantitative MS was first shown in 1995, when the isotope-coded affinity tag (ICAT) technology, based on a cysteine-specific tagging of intact proteins, was introduced. Alternative chemical labelling methods, such as the isobaric tag (iTRAQ), were later developed and introduced the stable isotope of proteolytic peptides. A limitation of the isotope method, despite the relative complexity of the isotope, is the low number of samples that can be quantified at the same time (maximum of eight). An example of such a strategy is provided in a study that employed the isobaric tag peptide labelling (iTRAQ) coupled with nano-LC tandem MS to analyse the effect of imatinib on the CML cell proteome. Large populations of proteins and novel consequences of the action of the imatinib were identified. In particular, DEAD-box protein-3, Hsp105 and peroxiredoxin-3 were identified as potential biomarkers for a response to imatinib.

Due to continuing technical improvements, new generations of MS offer improved sensitivity and a higher sequencing speed, allowing for the label-free quantitative analysis of samples. Samples are separately analysed by nanoLC-MS/MS and the peak intensity of peptides is compared, providing information on quantitative differences. New generations of FT mass spectrometers offer improved sensitivity and significantly higher sequencing speeds, typically enabling the identification of several thousands of proteins. Linear ion trap (LTQ) combined with a new mass analyser, the orbitrap, can achieve very high resolutions and mass accuracy. It is very likely that by maximising the number of proteins identified during a nanoLC-MS/MS run and allowing the comparison of large numbers of patients, these tests will be frequently performed, despite the significant investment needed to achieve this.

Interpreting Protein Language Using Proteomics
One major element of the regulation of proteome function is the consequences of post-translational modification (PTM), which define the structural plasticity of proteins. PTMs such as phosphorylation, glycosylation and acylation increase molecular heterogeneity – more than 200 types of PTM have been characterised – and represent significant markers of the gene diversity playing a fundamental role in cell biology. Constitutive phosphorylation, due to mutations in tyrosine kinase, is present in a significant proportion of haemopathies and represents critical events in the pathogenesis of MPD-. PTM is recognised as an important target for molecular medicine.

MS allows for the site-specific assignment of PTM to resolve individual amino acids in proteins. The interest in an MS strategy as a functional...
Proteomics in Myeloproliferative Disorders

Screen for mutant kinase alleles in MPD was highlighted in a study by Mercher et al. This study used a phosphotyrosine, immunoprecipitation protocol to identify tyrosine phosphorylated proteins in human, acute, megakaryocytic, leukaemic cell lines with constitutive activation of STATS. The approach identified 346 phosphopeptides including several JAK family members. In particular, it identified a new JAK2-T875-N mutation in the CHRF-288 cell line, which was shown to be capable of conferring growth factor independence on the Ba/F3 cell line. Expression of JAK2-T875-N in a murine bone marrow transplant model caused a myeloproliferative disorder with splenomegaly, megakaryocytic hyperplasia, myelofibrosis and polycythemia and mild leukocytosis. These results demonstrate the interest in an MS-based PTM approach to gain new molecular insights into MPD pathology.

Conclusions

Tremendous technological advances in mass spectrometry coupled with the development of powerful bioinformatics tools and the assembly of protein databases containing putative protein information have led to a considerable increase in the potential of proteomics. The difficulties caused by the complexity of cellular or biological fluid proteomes are gradually being overcome. The development of new methods of protein quantification coupled with their identification represents a considerable improvement in the scope of quantitative profiling.

Better definition of pre-analytical steps (sampling and handling) has contributed to minimising non-specific disease-associated changes. Together with significantly increased levels of standardisation and reproducibility, it allows for the planning of efficient strategies in cancer research.

Proteomics will indubitably take a strategic place, together with other ‘omic’ technologies, in the investigation of MPD molecular biology and could highlight new clinically relevant hypotheses into MPD pathogenesis.