Jean-Pierre Wery
INCAPS, 351 West Tenth Street, Suite 350, Indianapolis, Indiana 46202

Application of Proteomics Technologies to Biomarker Discovery and Development – Challenges and Solutions

In recent years, proteomic technologies have led to enormous advances in basic research and medicine. The impact these technologies can have on a more complete understanding of the processes of life is only now beginning to be realized. On the clinical side, proteomics is viewed as a promising new approach that will speed the discovery and validation of protein biomarkers that correlate with disease and allow for assessment of therapeutic regimens. However, despite many significant technological developments, serious challenges remain. For instance, greater sensitivity in detecting proteins from complex mixtures and easy access to fast and cost-effective assays to validate promising markers are improvements that will allow the field of proteomics to have a much greater impact. In this article, we will discuss how INCAPS is tackling these two challenges to offer unique proteomic capabilities to its customers.

Proteomics

Proteomics can be defined as the identification, characterization and quantification of all proteins involved in a particular pathway, organelle, cell, tissue, organ or organism that can be studied in concert to provide accurate and comprehensive data about that system. The analysis of proteomes is significantly more challenging and complicated than sequencing genomes for three main reasons. First, in higher eukaryotes, a single gene often produces many different forms of the protein, primarily due to alternative splicing and various post-translational modifications, which can result in important functional differences. Secondly, genomes are largely static throughout the lifetime of a cell or organism, but proteomes vary dramatically. The classic example is the caterpillar transforming itself into a butterfly, each with strikingly different proteomes dictating remarkably disparate phenotypes. Third, proteomics is currently more challenging than genomics because the technologies required for proteomics are more complex than, and thus not as robust as, nucleic acid-based methods.

Biomarkers

The National Academy of Sciences defines "biomarker" as an indicator that signals events in biological samples or systems. Molecular biomarkers hold the promise of transforming almost every field in biology and medicine. With the recent emergence of new technologies such as genomics-based and proteomics-based approaches, the field of biomarker discovery, development and application has been the subject of intense interest and activity. Ultimately, biomarkers offer the promise of more efficient discovery and development of novel therapies as well as improved and more individualized disease prevention and treatment. And one of the most valuable classes of biomarker over the past century has been protein biomarkers.

Applying Proteomics to Biomarker Discovery

The development of modern separations techniques coupled with advanced mass spectrometry technologies have ushered in a new paradigm that expands the scope of protein identification, quantitation and characterization. These proteomic-scale capabilities now enable thousands of proteins to be identified from complex mixtures. Despite the success of existing platforms, the current approaches are unable to tackle some important research problems of great interest to many scientists and clinicians. For instance, it is highly desirable for protein markers that will be used in medical practice to be detected and measured in body fluids such as plasma. These fluids are highly complex mixtures of proteins and exhibit a very broad dynamic range of protein relative abundances (up to 12 orders of magnitude). Therefore, it is believed that many important proteins, and potentially important markers, escape detection by the current proteomics technologies. Furthermore, although a number of protein biomarkers have found their way into clinical practice, very few of these have been discovered by proteomics approaches. Reasons for this situation include the significant time and cost required to validate, or invalidate, the markers suggested by proteomics experiments. The availability of fast, cost-effective, mass spectrometry-based validation approaches would change that situation. We can condense these problems in protein biomarkers discovery and development into two distinct proteomics challenges:

1. Detection, identification and quantification of low-abundance proteins from complex biological samples, and
2. Development of a high-throughput, low-cost mass spectrometry solution to validate promising protein markers identified in proteomics experiments.

To provide solutions to these challenging problems, INCAPS is developing and implementing novel protein separation approaches combined with cutting-edge mass spectrometry techniques. It is also turning these novel technologies into reliable, cost-effective services available to its academic and commercial customers. Following is a short description of the technologies implemented at INCAPS.

Development and Implementation of Multi-Dimensional Chromatographic Separation Technologies

One measure of success of a proteomics experiment is how much proteome coverage can be achieved. It is also clear that with complex biological samples such as human plasma, many peptides introduced into the mass spectrometer will not be detected due to their relatively low abundance, ultimately reducing the number of proteins that can be confidently identified. One way to address this problem is to reduce the complexity of the sample introduced into the mass spectrometer by using multi-dimensional separation techniques.

Over the years, one- or two-dimensional sodium dodecyl sulfate polyacryamide gel electrophoresis (1D or 2D SDS-PAGE) has proven to be a reliable and efficient method for separation of several hundred to a few thousand proteins on the basis of differences in their isoelectric point (pI) and molecular weight. Although gel electrophoresis has many advantages, there are also several disadvantages. The procedure can be time-consuming, create a heavy work load, and lack sufficient reproducibility. The insolubility of hydrophobic proteins and the difficulty in detection and separation of low-abundant proteins are additional weaknesses. Because of these shortcomings several labs have started using Reverse Phase (RP) High Performance Liquid Chromatography (HPLC) for separations of proteins and peptides. The effluent can be directly infused into the mass spectrometer resulting in an automated high throughput analysis. However there are also disadvantages associated with use of the HPLC method. There is a tendency in such experiments to "re-discover" over and over again the same set of proteins (i.e. the most abundant ones). Most low-abundance proteins, which often are very important functionally, are missed by this approach as well. To address these problems, several groups have proposed to pre-fractionate the samples prior to RP separations. There are currently two major approaches to pre-fractionate complex biological samples: (i) the chromatographic approach; and (ii) the electrophoretic approach. INCAPS uses both approaches and combines these methods to further fractionate the samples using two- and three-dimensional processes. INCAPS also optimizes the fractionation process for different types of samples (human plasma, human CSF and cultured human cells) and then combines this with its MS techniques to maximize the number of proteins that can be confidently identified from these complex biological samples.

Optimization of Mass Spectrometry Technologies to Maximize Protein Identification from Complex Biological Samples

The proteolytic digest of a biological sample such as plasma or CSF can contain several hundred thousand peptides. HPLC and ESI (electrospray ionization) coupled with data-dependent MS/MS acquisition has been shown to be an exceptionally useful tool for the analysis of complex proteomics samples. Similarly, MALDI (matrix assisted laser desorption ionization) in conjunction with MS/MS analysis has also been used to acquire product ion spectra from proteomics samples, generating data sets both analogous and complementary with those produced by ESI LC/MS/MS.

Despite advances in software and instrumentation that have enabled rapid acquisition of very large data sets, the extreme complexity of proteomic samples typically leads to peptides being introduced into the mass spectrometer at a rate far exceeding that of MS/MS data acquisition. This results in an under sampling of the proteome – at least at the peptide level and most often also at the protein level. At present, two ways have been proposed to increase proteome coverage in proteomics experiments. The first is use of multiple dimensions of chromatography to further reduce the complexity of the samples introduced into the mass spectrometer (discussed above). The second proposed approach takes advantage of the complementary nature of ESI and MALDI. The ESI and MALDI ionization processes are complementary in that they often suppress ion currents for different peptides. The combination of both techniques on the same sample takes advantage of the benefits of both. In particular, the combination of high-throughput real time ESI MS/MS spectra acquisition using an ion trap with high scanning speed followed by intelligent, non-redundant off-line MALDI TOF/TOF analysis will significantly increase the total number of detected peptides leading to the identification of more proteins with improved sequence coverage.

INCAPS is optimizing the performance that can be achieved by combining ESI and MALDI MS technologies applied to complex biological samples such as plasma, CSF and cultured human cells. As a result of our work, we are consistently able to detect and identify more proteins from proteomics studies, thereby offering an improved service to our customers and enhancing the biological information obtained.

Development and Optimization of Fast and Cost-Effective Quantitative Mass Spectrometric Assay for Biomarker Validation

Global proteomics experiments can often suggest many potential protein markers. Orthogonal protein quantification techniques are needed to validate changes in levels of these candidate biomarkers. Typically antibody-based assays such as enzyme-linked immuno-sorbent assays (ELISA), radioimmunoassay (RIA), or quantitative Western blot analysis are used to measure proteins in complex biological samples. Development of these assays is a time-consuming and resource-intensive effort and cannot practically be undertaken for many proteins. Furthermore, results from these assays often provide quantitative observations for the total composition of the candidate proteins and do not discriminate among closely related forms of specific proteins such as differentially modified proteins. Rapid development of simpler, more specific and more quantitative methods will greatly aid in validation of protein biomarker candidates. INCAPS is demonstrating that MS-based assays can successfully compete with antibody-based methods in cost, speed, and specificity. Such MS-based assays also represent the only viable way to characterize proteins for which no suitable antibody has yet been developed. In addition, MS-based assays are capable of simultaneously characterizing

several proteins from the same sample in the same experiment. This is a significant advantage over conventional monoplex antibody-based assays, especially given the fact that many scientists believe that successful diagnostics will involve the characterization of multiple proteins. Multiplexed antibody assays do exist, but they take even longer to develop and must be optimized to minimize cross-reactivity while managing the significant challenge represented by the huge dynamic range of body fluids such as plasma. MS-based methods are inherently multiplexed and do not need to be optimized in this manner.

Since it is difficult to predict whether the protein of interest will be better suited for ESI or MALDI MS, INCAPS is pursuing both technologies for implementation of specific assays. INCAPS is currently developing and validating assays for several biologically important proteins in neuroscience and cancer. These assays, which are run directly on plasma or CSF samples, are fast (a few minutes per sample), very accurate and very specific. In addition, they are very comparable in cost with antibody-based assays.

Conclusion

Proteomic technologies have the potential to revolutionize the field of protein biomarker discovery and development. However, before the full impact of these technologies can be realized, some technical hurdles must be resolved. INCAPS is developing and implementing technologies to address these limitations. It is combining novel protein separation technologies and mass spectrometry approaches to improve the detection and identification of low-abundance proteins. This will have an impact on our ability to discover novel and important markers. INCAPS is also developing a fast and cost-effective mass spectrometry-based assay, greatly improving our ability to validate (or invalidate) promising new markers. These new approaches are available to our customers as reliable and cost-effective services.

Dr. Jean-Pierre Wery is Chief Scientific Officer for INCAPS, a contract research organization in Indianapolis, Indiana focused on protein analysis, proteomics and biomarker discovery and validation. Prior to joining INCAPS, Dr. Wery spent three years at Vitae Pharmaceuticals, Inc. where he was Vice President of Computational Drug Discovery. Prior to Vitae Pharmaceuticals, Dr. Wery spent 12 years at Eli Lilly and Company in various scientific and management positions. He first joined Eli Lilly as a protein crystallographer and participated in several structural biology projects targeting enzymes such as HIV-Protease and Secretory Phospholipase A2. Later, Dr. Wery accepted management responsibilities, leading the company's effort in structure-based and computational drug discovery and more recently managing the company's Statistics and Information Sciences Division.

Dr. Wery received his B.S. and Ph.D. in Physics from the University of Liège, Belgium. During his Ph.D. studies, Dr. Wery worked on the three-dimensional structure of bacterial enzymes involved in antibiotic resistance and also helped develop novel methodologies, such as MAD phasing, for protein structure determination by x-ray crystallography. Following his Ph.D., Dr. Wery did postdoctoral studies at Purdue University with Professor Jack Johnson in the area of virus crystallography. Dr. Wery has authored more than 50 abstracts and publications.