A New Paradigm in Proteomics

Peptides and proteins have been associated with many disease states such as cancers, diabetes, neurological and cardiovascular diseases [1-4]. Despite the limited success of a handful of biomarkers, most diseases lack sensitive and specific biomarkers. One of the most successful biomarkers, Prostate Specific Antigen (PSA) has a fairly high false-positive rate and very low clinical sensitivity (~25%).

Many proteomics researchers are involved in developing integrated suites of algorithms, statistical methods, and computer applications to perform proteomics profiling directly from mass spectra [5-13]. These strategies have demonstrated diagnostic sensitivity and specificity superior to conventional single biomarkers. Mass profiling of diseases is a major paradigm shift in proteomics; the entire pattern contains important diagnostic information. The use of mass spectrometry as a clinical tool provides many key advantages over traditional immunoassays. The speed of mass spectroscopy may enable diagnosis ‘on the fly’. Since the mass spectrometer can accurately measure proteins and peptides without the need for antibodies, the development time is drastically reduced.

Spectral profiling is also likely to play a key role in toxicoproteomics, where specific profiles may be used to signal toxic events. Detection of such signatures in animal or early human clinical trials could help pharmaceutical companies detect adverse drug effects, improving the efficiency of clinical trials and reducing toxic effects.

A Complete Biomarker Discovery Platform

PerkinElmer has joined forces with industry innovators Vivascience and Predictive Diagnostics to offer a complete solution for biomarker discovery in drug discovery and diagnostics. The BioXPRESSION™ Biomarker Platform combines three powerful technologies: high throughput, automated, carrier-protein biomarker enrichment with novel membrane chromatography, exceptional high resolution detection using prOTOF™
MALDI O-TOF mass spectrometry and powerful, dedicated, biomarker analysis capability with BAMF™ Biomarker Amplification Filter algorithms. The introduction of this new platform gives researchers the ability to screen thousands of samples and obtain the highest confidence results for population segmentation, pre-clinical trials and disease profiling.

**Reproducible Sample Complexity Reduction**

Reproducible sample complexity reduction is an essential first step in biomarker discovery experiments. The large dynamic range of serum/plasma necessitates the effective removal of abundant serum proteins to allow analysis of the lower concentration analytes.

ProXPRESSION™ Protein Enrichment Kits enable efficient, reproducible, microscale fractionation for complex protein mixtures using patented membrane absorber (MA) chromatography technology (Vivascience AG) combined with PerkinElmer’s proprietary buffers and elution chemistries.

Vivascience’s MA technology, in contrast to traditional resins and beads, features a rigid format that makes it highly reproducible and robust. The technology also has significant ease-of-use advantages when incorporated into the PerkinElmer biomarker screening solution.

**ProXPRESSION Acidic and Basic Protein Enrichment Kits**

The ProXPRESSION Basic Protein Enrichment Kit and Acidic Protein Enrichment Kit are designed to fractionate a protein sample so that the recovered fraction is enriched in proteins having either a basic or an acidic pH. Reagents, spin columns, and protocols are provided to prepare the sample for further analysis through clarification (to remove debris from the sample), charge fractionation, and salt removal.

In one example, human brain (neocortex) tissue samples were fractionated based on charge and separated on 2-D gels. An example of the enrichment is shown in Figure 1. Many protein spots not visible in the total protein gel are identifiable in a gel containing the basic protein fraction.

Fractionation of acidic and basic proteins resulted in significant complexity reduction of protein patterns and better resolution of individual protein spots as compared to the protein patterns for non-fractionated proteins [14].

**ProXPRESSION Biomarker Enrichment Kits**

Analyzing low molecular weight peptide biomarkers in serum and plasma has been extremely difficult due to the vast number of contaminating species (salts, lipids, etc.) present. These contaminants pose problems for separation methods as well as detection systems. High concentrations of lipids and salts can suppress the ionization of peptide biomarkers in MALDI-TOF MS analysis. 2-D gels cannot effectively separate proteins below 10 kilodaltons, the range where MS has the greatest sensitivity and specificity.

The low molecular weight region of the human proteome is comprised of peptides and fragments of proteins. This ‘fragmentome’ has attracted much interest because it is a relatively ‘unmined’ region of the proteome [15]. These lower molecular weight species have the ability to bind to carrier proteins (such as albumin), extending their half-lives in serum. The ProXPRESSION Biomarker Enrichment Kits utilize cibacron blue to capture these carrier proteins. Proprietary elution buffers are then used to generate high fidelity peptide biomarker signatures.

A key advantage of using a mass spectroscopy-based approach in mining this low molecular weight proteome is that the mass spectrometer can reproducibly record protein fragment peaks. If a biomarker for a given disease state

![Figure 1. Application of ProXPRESSION Basic Protein Enrichment Kit to human neocortical brain tissue. Many protein spots were enriched, as evident in comparison of the 2-D gels prior to (A) and after fractionation (B). Approximate range of proteins: 6-200 kDa, pl 6.0 – 10.0.]()
is a fragment of a larger protein, it may be extremely difficult or impossible to produce effective antibodies for the conventional ELISA diagnostic tests. Such antibodies would most likely demonstrate cross-reactivity to the larger parent protein as well.

Application of these kits to human serum/plasma results in high-content, reproducible signatures of the low molecular weight proteome (Figure 2). This method results in reproducible and robust protein signatures. Surface-based biomarker enrichment strategies have been criticized for their limited binding capacities and tendency to bind the more abundant serum proteins [16, 17]. The ProXPRESSION biomarker enrichment kits offer a unique opportunity to discover biomarkers within the low molecular weight proteome.

MALDI Orthogonal-TOF Mass Spectrometry: Enhanced Biomarker Feature Detection

MALDI mass spectrometry has become an important tool for proteomics researchers. MALDI Orthogonal-TOF mass spectrometry provides superior biomarker discovery performance characteristics over traditional axial MALDI systems [18-22]. Pulsing ions orthogonally into the time-of-flight detection region separates ionization conditions in the sampling chamber from detection. The influence of ionization conditions on mass accuracy is effectively eliminated.

The excellent mass accuracy and sensitivity and high resolution achievable over a wider mass range assures high information content in the MALDI spectra. This translates into ease-of-use and higher confidence in the resulting data.

Wider Mass Range

MALDI O-TOF MS is capable of collecting data for a broader mass range (100 Da – 300,000 Da) in a single acquisition. This reduces the number of runs and eliminates the need for ‘spectral-stitching’ to combine multiple acquisitions. Figure 3 illustrates the ability to acquire peptides and proteins in a single acquisition.

Figure 2. MALDI O-TOF MS spectra of three healthy human serum replicates processed using ProXPRESSION Biomarker Enrichment Kits (1,000–10,000 Da displayed, insert shows zoom of reproducibility in the 5,000–8,000 Da range).

Figure 3. Wider Mass Range Analyses: Collect from 100 Da to 300 kDa in a single acquisition (700–25,000 Da shown). 1) Angiotensin I, 2) ACTH 18–39, 3) Cytochrome C (equine), 4) Myoglobin (equine) and 5) Trypsinogen.
Generating Disease Fingerprints with BAMF Technology

BAMF Technology generates a set of rules that distinguishes disease from non-disease (or responders from non-responders, etc.). The first step in generating a disease fingerprint involves recording quality protein signatures for well-characterized disease and non-disease samples. The high resolution, sensitivity and mass accuracy of MALDI O-TOF MS provides very reproducible, information-rich spectra. These protein signatures are uploaded to the BAMF Technology LINUX cluster processing center directly from the prOTOF acquisition computer via a 128-bit encrypted web interface.

BAMF Technology is a proprietary pattern recognition engine and a discovery platform whose primary purpose is the illumination of uncommon features present in ensembles of spectra. These features are used to develop classification models and screen unknowns. The BAMF Technology suite consists of several analytical components, including some common computational algorithms, and other, highly specialized custom algorithms. The platform uses known techniques and algorithms as well as proprietary techniques and combinations of techniques and selected algorithms to achieve superior and more reproducible results as compared to known and existing methods and systems that are currently available.

The BAMF Technology platform comprises three process steps in the learning mode to train the system and provide the framework for analyzing data and providing results based on specific inputs. These steps include: Outlier Rejection (also known as “Pre-Processing”), Feature Discovery, and Model Building. In classification mode, the BAMF Technology platform comprises the process steps of Outlier Rejection and Categorization.

Pre-Processing

Data is received from a mass spectrometer as vector data (m/z along with an intensity level). The Pre-Processing comprises quality control steps used to identify and remove spectra represented by systemic error and to ensure an acceptable level of uniformity of the data. It does not affect or change any of the data that is passed on to the next step. Pre-Processing establishes the uniform set to proceed with in subsequent steps, such as Feature Selection and Model Building.
Feature Selection
In the Feature Selection Step, algorithms are run to look for features in the spectral data that are identified by distinguishing variations from group to group or spectra to spectra. Feature Selection includes three steps: Filtering, Human Intervention and Consensus. Feature selection can be bypassed if the operator wants to focus on specific features, for example, those that lie beneath the noise floor.

Model Building
The Feature Selection process will identify many biomarkers, only some of which may be useful as diagnostic predictors. A model is a particular combination of biomarkers. The Model Building process is used to build and test all the possible models of the identified biomarkers to determine which biomarkers in which combinations have the best diagnostic utility. The challenge is that the greater the number of features, the greater the number of models and the more time and computer processing intensive the search for useful biomarkers becomes. On the other hand, if only a sampling of models is tested in order to save time and computing power, the chances of finding the model with optimal utility decreases.

These various processes in the Model Building step may result in markers that are not exactly intuitive, as they do not represent known elements or significant peaks in the spectra. However they may act as helper markers to provide a more accurate picture of the indication for which detection is desired. Irrespective of the approach used, all classification algorithms that represent the Model Building step undergo a consensus function for determination of the final selected class of categorizers. The consensus function operates by polling the best and worst categories to identify overlaps, using a stringent and majority rules approach.

The results of the analysis provided by the BAMF Technology platform on an input submitted for testing and quantification after the system is fully trained includes the output of a confidence level that is based on the nature of the results and the closeness of the results to the criteria set during the system training. Again, consensus analysis of the output of more than one categorizer may be used during Model Building.

Models can be altered at will giving the user control over what features significantly impact screening. BAMF Technology then goes through several additional proprietary steps to generate the optimum disease predictor features. Once these biomarkers are calculated, a model (or disease ‘fingerprint’) is created for the disease state.

Classification Using BAMF Generated Disease Fingerprints
The procedure for classification of unknown samples is shown in Figure 4. The samples are processed and analyzed using MALDI O-TOF mass spectrometry. The resulting protein profiles are then run against the model to see how well they classify.

Figure 4. BAMF Technology is a proprietary, integrated pattern recognition engine and discovery platform for signal processing and image analysis to discover, differentiate, and validate biomarkers.
Benefits of BAMF Technology

The combination of MALDI Orthogonal-TOF mass spectrometry and the high throughput BAMF Technology provides the most comprehensive analysis of the disease state. Some of the benefits include:

- Data may be uploaded via a secure, 128-bit encrypted web interface. This obviates the need for continuous software updates and provides a seamless transition from data acquisition to informatics analysis.
- Delivers up to 100 percent specificity and sensitivity compared to single-biomarker tests.
- Includes many QC/QA preprocessing steps to eliminate poor spectra and ensure quality results.
- Terabytes of data can be managed in a single location.
- Provides visualization tools (including contour plots).

- Offers the potential for stand-alone diagnostics.
- Results are highly reproducible.

Alzheimer’s Disease

An estimated 15 million people worldwide suffer from Alzheimer’s disease (AD). It is a progressive illness that kills nerve cells and destroys nerve connections in the brain. The disease is not reversible and currently there is no cure. The lifespan of an Alzheimer’s disease victim is generally reduced, although a person may live anywhere from 3 to 20 years after diagnosis. The disease is marked by mental changes resulting from damage in the brain tissue. Because these changes cannot be visualized until an autopsy, diagnosis for the disease is based on symptoms that patients have. Symptoms include gradual loss of awareness, memory, and judgment as well as mood and behavioral disturbances. Current proteomics research efforts are focused on identifying disease-specific protein biomarkers to aid in diagnoses.

Analysis of Alzheimer’s Disease Samples

Serum and brain (cortex) tissue samples were analyzed using the biomarker discovery platform described herein. ProXPRESSION Acidic and Basic Protein Enrichment Kits were used to fractionate the tissue samples prior to peptide mass fingerprinting (PMF) experiments. The ProXPRESSION Biomarker Enrichment Kits were used to generate disease fingerprints using BAMF Technology. The experimental design is shown in Figure 5.

Peptide Mass Fingerprinting of Alzheimer’s Disease Samples

Serum samples were obtained from the Religious Orders Study, a study funded by the National Institute on Aging in 1993. The study includes

Figure 5. Alzheimer’s disease samples were fractionated and analyzed using two complementary strategies: peptide mass fingerprinting (green) and BAMF disease fingerprinting (red).

Figure 6. Differentially expressed proteins obtained from basic fractionation of AD brain lysates.
more than 1,000 older nuns, priests, and brothers without known dementia from about 40 groups across the U.S. All study participants participate in annual cognitive and motor testing and all have agreed to brain donation at the time of death. The study has greater than 95% follow-up of survivors with up to 10 data points.

The sample set consists of two groups, approximately fifteen samples each of Alzheimer’s human brain tissue (neocortex) and controls. 335 ‘blinded’ sera were also obtained. Brain lysate and sera samples were diluted 10X in binding buffers. The diluted samples were fractionated using ProXPRESSion Acidic and Basic protein enrichment kits. 2-D gels were run as described in reference 24. Gel images were acquired using a PerkinElmer ProXPRESS™ Proteomic Imaging System. Sample preparation and trypsin digestion were carried out using the Montage® In Gel Trypsin Digest Kit (Millipore). Spectra were acquired on the prOTOF 2000 MALDI O-TOF and analyzed with TOFworks™ software (PerkinElmer).

Over 100 statistically significant differentially expressed protein spots were identified by differential gel electrophoresis (Figure 6). Spots were ranked based on the following criteria: the number of gels present, the P-value, the differential expression value and spot resolution. The top-ranked gel spots were excised, digested with the enzyme trypsin and identified using MALDI O-TOF mass spectrometry.

When the 2-D gels images from AD and non-AD basic fractions were analyzed, 17 protein spots (13 different proteins) were identified as being differentially expressed. Table 1 lists the identities of these proteins and their expression in both groups. Seven-protein spots were decreased 2- to 7-fold in AD including protein phosphatase 1, phospholipase C, pyruvate kinase, phosphoinositide-3-kinase and fructose biphosphate aldolase.

Generating an Alzheimer’s Disease Fingerprint Using BAMF Technology

The 335 randomized serum samples were processed in triplicate using ProXPRESSION Biomarker Enrichment Kits. Reference serum and no-serum controls were also

---

**Table 1.** Listing of identifications for the highest ranked proteins from basic fractionation.

<table>
<thead>
<tr>
<th>Spot ID</th>
<th>Protein</th>
<th>MW (kDa)</th>
<th>pI</th>
<th>Expression (Normal/AD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1780</td>
<td>Protein phosphatase 1</td>
<td>110.9</td>
<td>5.5</td>
<td>3.98</td>
</tr>
<tr>
<td>2153</td>
<td>Phospholipase C, delta</td>
<td>86.5</td>
<td>6.2</td>
<td>2.00</td>
</tr>
<tr>
<td>2324</td>
<td>Pyruvate kinase</td>
<td>58.5</td>
<td>7.2</td>
<td>6.15</td>
</tr>
<tr>
<td>2463</td>
<td>Phosphoinositide-3-kinase</td>
<td>54.6</td>
<td>5.8</td>
<td>2.19</td>
</tr>
<tr>
<td>3061</td>
<td>Fructose biphosphate aldolase</td>
<td>39.8</td>
<td>6.4</td>
<td>7.26</td>
</tr>
<tr>
<td>3049</td>
<td>Aldolase C</td>
<td>39.8</td>
<td>6.4</td>
<td>1.94</td>
</tr>
<tr>
<td>3370, 3436, 3383</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
<td>36.2</td>
<td>8.4</td>
<td>0.24, 0.10, 0.42</td>
</tr>
<tr>
<td>3334</td>
<td>GDNF receptor, alpha 3 precursor</td>
<td>46.1</td>
<td>9.2</td>
<td>0.71</td>
</tr>
<tr>
<td>2777</td>
<td>Enolase-1</td>
<td>47.4</td>
<td>7.0</td>
<td>0.32</td>
</tr>
<tr>
<td>2339</td>
<td>Copine III</td>
<td>61.0</td>
<td>5.7</td>
<td>0.68</td>
</tr>
<tr>
<td>387</td>
<td>KARP-1 Binding protein</td>
<td>164.9</td>
<td>6.8</td>
<td>0.43</td>
</tr>
<tr>
<td>2446</td>
<td>hCRMP-2</td>
<td>62.7</td>
<td>6.0</td>
<td>0.22</td>
</tr>
<tr>
<td>1260</td>
<td>Valosin-containing protein</td>
<td>90.0</td>
<td>5.1</td>
<td>0.58</td>
</tr>
<tr>
<td>2584</td>
<td>Protein kinase C</td>
<td>85.0</td>
<td>6.8</td>
<td>0.32</td>
</tr>
</tbody>
</table>

**Table 2.** Sensitivity and specificity of the BAMF Technology generated Alzheimer’s disease fingerprint.

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>94%</td>
<td>89%</td>
</tr>
</tbody>
</table>
processed in each experiment. Seventy percent of the 1,050 sample spectra dataset were grouped into three training sets:

- **Group 1** = Alzheimer’s patients and patients diagnosed with mild cognitive impairment (68 samples total).
- **Group 2** = only Alzheimer’s samples (43 samples).
- **Group 3** = control population (143 samples).

These samples were analyzed on a proTOF 2000 mass spectrometer. The protein signatures were uploaded to the BAMF processing center.

For the remaining spectra, 30% of the samples from each group (91 samples total) were pooled into a single “blinded testing group”. This dataset was run against the disease model generated from the training set. The Alzheimer disease model was able to discern control samples from Alzheimer’s samples with 94% sensitivity and 89% specificity. Four potential biomarkers with high sensitivity and specificity were discovered.

**Summary**

PerkinElmer’s powerful, integrated biomarker discovery platform includes fractionation/enrichment kits, automated MultiProbe® II liquid handling, MALDI O-TOF mass spectrometry and powerful BAMF informatics software. This platform was used to process more than 300 human serum samples from Alzheimer’s and normal patients.

Alzheimer’s disease and control samples were fractionated using ProXPRESSION Basic Protein Enrichment Kits. These kits reduce the complexity of proteomic samples and enhance the detection of low-abundance proteins. The combination of reproducible fractionation methods with precise and accurate separation, detection and identification results in successful, statistically significant protein identifications. Some of the proteins identified in this study have been reported in previous Alzheimer’s research studies. The cytosolic enzyme alpha-enolase has been identified as a target of protein oxidation and is involved in the glycolytic pathway in the pathological events of AD [25]. Other proteins identified as differentially expressed in this study have not been reported and their possible roles in neuro-degeneration are being explored.

Fractionation of Alzheimer’s disease and control samples using ProXPRESSION Biomarker Enrichment Kits generated highly reproducible spectra by MALDI O-TOF MS. These kits allow researchers to investigate the low

![Figure 7. BAMF generated contour plots for control and Alzheimer’s disease samples.](image-url)
molecular weight proteome (i.e. fragmentome), a virtually unexplored treasure trove of biomarkers. These spectra were used to generate a BAMF Alzheimer’s disease fingerprint. This fingerprint successfully classified Alzheimer’s disease samples from controls with high sensitivity and specificity.

The PerkinElmer biomarker platform enables complementary discovery strategies: identification using PMF, and rapid protein profiling (i.e. fingerprinting) using BAMF technology. Knowledge of a protein’s identity is important for learning more about a disease processes underlying pathology. The rapid screening capabilities of protein profiling offer incredible potential for improving clinical diagnostics. This comprehensive, integrated biomarker discovery platform, however, is not limited to biomarker discovery for clinical diagnostics. The performance characteristics described are well suited to meet the demands of toxicoproteomics and drug efficacy studies. The rapid classification capabilities of the integrated bioinformatics platform described herein can be used to screen for and detect adverse effects of drug trials. Ultimately these tools may result in better, safer drugs and more efficient patient stratification, fulfilling the unrealized promises of the genomics revolution.

For a complete BioXPRESSION Biomarker Platform product list, see inside back cover.

References

14. Lopez, M.F. et al. Microscale fractionation facilitates detection of differentially expressed


**BioXPRESSION Biomarker Platform Complete Product Listing**

A complete line of reagents and instrumentation from PerkinElmer and its partners supports your laboratory's proteomics needs at every step.

<table>
<thead>
<tr>
<th>Product</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BAMF™ (Biomarker Amplification Filter) Data Analysis</strong></td>
<td>A sophisticated analysis technology from Predictive Diagnostics, Inc. to identify patterns of proteins and peptides over a highly secure Internet connection from high-resolution mass spec data. Builds multiple models for high-confidence and develops the optimal biomarker set.</td>
</tr>
<tr>
<td><strong>MALDIchip™ Target Plates</strong></td>
<td>Our disposable MALDI targets simplify your sample throughput needs.</td>
</tr>
<tr>
<td><strong>MultiPROBE® II Liquid Handling System</strong></td>
<td>Reduce sample preparation errors with our flexible, robotic liquid handling solution.</td>
</tr>
<tr>
<td><strong>Nonlinear Dynamics Ltd. Software Products</strong></td>
<td>A complete suite of image analysis and bioinformatics software, including ProMST™ mass spectrometry bioinformatics package and Progenesis™ image analysis software.</td>
</tr>
<tr>
<td><strong>ProFINDER™ Image Analysis Software</strong></td>
<td>Find 50–100% more true spots than with standard software packages. Powerful, comprehensive and easy to use.</td>
</tr>
<tr>
<td><strong>prOTOF™ 2000 MALDI O-TOF Mass Spectrometer</strong></td>
<td>Our breakthrough MALDI orthogonal time of flight mass spectrometer with collisional cooling is the most stable and accurate single MALDI MS system available. It provides improved instrument stability, resolution, and mass accuracy across a wide mass range making it the best choice for biomarker discovery and peptide mass fingerprinting.</td>
</tr>
<tr>
<td><strong>ProXCISION™ Proteomics Gel Cutting Robot</strong></td>
<td>Our high precision ProXCISION excises spots from 2D gels and bands from 1D gels that were identified by the ProXPRESS, ensuring you get the spots you’re after every time. The ProXCISION’s highly accurate, precise protein picking interfaces with ProFINDER 2D Image Analysis Software for the utmost in accuracy and traceability.</td>
</tr>
<tr>
<td><strong>ProXPRESS™ 2D Proteomic Imaging System</strong></td>
<td>Our imager provides the most accurate and reproducible detection and quantification of low abundance proteins on large 2D and 1D gels of any CCD-based imaging system.</td>
</tr>
<tr>
<td><strong>SYPRO® Ruby Stain</strong></td>
<td>For gel imaging, Molecular Probes’ stain provides excellent sensitivity.</td>
</tr>
<tr>
<td><strong>TOFprep™ MALDI Spotter</strong></td>
<td>Our spotter automates the digestion, clean-up, concentration and spotting of samples for MALDI MS analysis.</td>
</tr>
</tbody>
</table>