

# Protein Biomarker Research Pipeline for Developing Protein Biomarkers for Diabetic Kidney Disease

Using AB SCIEX TOF/TOF™ and QTRAP® Systems

Scott Bringans<sup>1</sup>, Thomas Stoll<sup>1,2</sup>, Kaye Winfield<sup>1,2</sup>, Tammy Casey<sup>1,2</sup>, Tim Davis<sup>3</sup>, Jenny Albanese<sup>4</sup> and Richard Lipscombe<sup>1,2</sup>

<sup>1</sup>Proteomics International, <sup>2</sup>Centre for Food and Genomic Medicine, Australia, <sup>3</sup>University of Western Australia, <sup>4</sup>AB SCIEX, USA

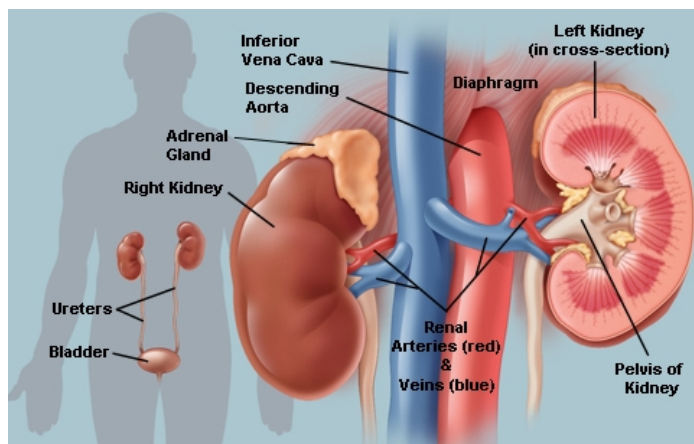
Diabetes mellitus is a chronic disease currently affecting 250 million people worldwide with the number projected to rise to 438 million by 2030<sup>1</sup>. Type 2 diabetes is by far the most common, affecting 85-90% of all people with diabetes. The complications of diabetes reflect tissue damage, especially to the small blood vessels in eyes, nerves and kidneys, and involvement of the large blood vessels to cause heart attack, stroke, impotence and foot problems. Early detection is critical so the disease can be controlled and complications reduced at an early stage<sup>2</sup>.

Diabetes is the largest cause of kidney disease (nephropathy) in developed countries and 10% to 20% of people with diabetes will die of kidney (renal) failure. While the mechanisms are complex, the cumulative result is that the renal filtering units or glomeruli cannot retain protein (albumin) which is excreted initially in the urine in small but abnormal amounts (microalbuminuria), with the prospect of heavier urinary protein losses (macroalbuminuria) as the complication progresses<sup>3</sup>.

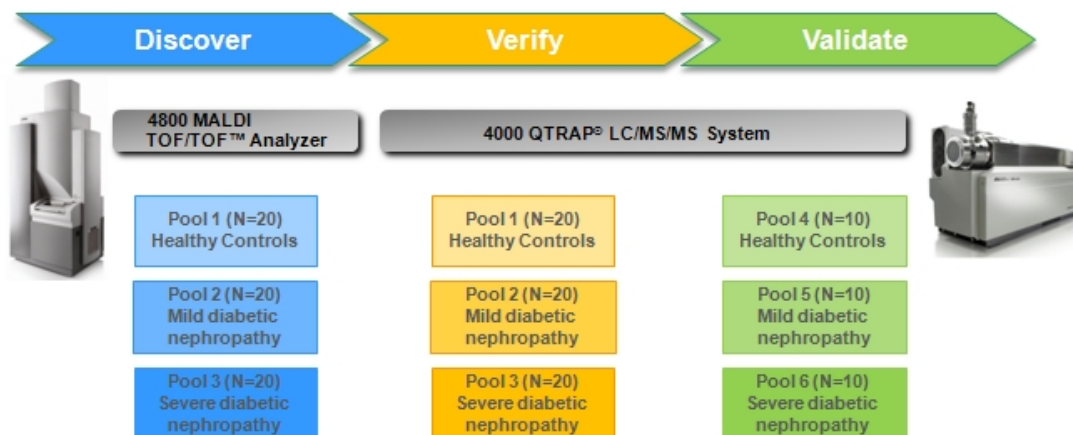


Diabetic nephropathy is detected primarily by measuring the ratio between the albumin and the creatinine excreted in the urine, to assess the degree of albuminuria. However, there are many difficulties with obtaining reliable results from this urine based test<sup>4</sup>.

An early detection protein biomarker that could be analyzed from blood or plasma would be ideal for early initiation of management and a slowing or prevention of disease progression. In addition, such biomarkers would be valuable in screening the effectiveness of new therapies, particularly those that are focused on preventing the development of diabetes complications, and be useful in longitudinal monitoring and as an indicator of disease prognosis. Here, a biomarker research pipeline established at the Centre for Food and Genomic Medicine in Australia in collaboration with Proteomics International was applied to the study of diabetic nephropathy.



**Figure 1. Kidneys and Nephropathy.** The kidneys are a pair of organs located in the back of the abdomen. The kidney's function is to filter the blood. All the blood in our bodies passes through the kidneys several times a day. The kidneys remove wastes, control the body's fluid balance, and regulate the balance of electrolytes. As the kidneys filter blood, they create urine, which collects in the kidneys' pelvis – funnel-shaped structures that drain down tubes called ureters to the bladder.



**Figure 2. Discovery to Validation Protein Biomarker Pipeline.** In the discovery phase (blue), 60 samples from three pools (healthy controls, mild diabetic nephropathy, severe diabetic nephropathy) were labeled with iTRAQ® reagents and analyzed using the AB SCIEX 4800 TOF/TOF™ Analyzer. In the verification phase (orange), the same cohort of 60 samples was used to cross validate the found biomarkers from the discovery phase. The MRM assay was developed and verified using the MIDAS™ Workflow, then the protein expression differences were confirmed on the second set of 3 pools. In this validation phase (green), individual samples from each pool were analyzed using the final MRM Assay.

## Materials and Methods

**Study Design:** In the discovery phase, 20 samples from three sample groups were pooled (healthy controls, mild diabetic nephropathy, severe diabetic nephropathy), labeled with iTRAQ® reagents and analyzed by the 4800 TOF/TOF™ system. In the verification phase, the same samples were used to cross validate the putative biomarkers from the discovery phase. In the statistical validation phase, new samples were used, consisting of 10 individuals from each sample group to apply the final MRM assay for a subset of verified biomarkers (Figure 2).

**Sample Preparation:** All clinical plasma samples were provided by Fremantle Diabetes Studies (FDS I & II) and Busselton Diabetes Study, Australia (BDS). Plasma samples were first depleted of the top 14 high abundance proteins using the Agilent MARS14 column (Agilent, USA). Proteins in the depleted samples were then trypsin digested and labelled with iTRAQ® reagent according to the manufacturers protocol (AB SCIEX, USA).

**First Dimension Ion Exchange:** Labelled peptides were desalted on a Strata-X 33 µm polymeric reversed phase column (Phenomenex) before separation by strong cation exchange liquid chromatography (SCX) on an Agilent 1100 HPLC system (Agilent Technologies, USA) using a PolySulfoethyl column (4.6 x 100 mm, 5µm, 300 Å, Nest Group, USA). Peptides were eluted with a linear gradient of 0-400 mM KCl. A total of 8 fractions were collected, desalted on a Strata-X 33 µm polymeric reversed phase column and dried.

**Second Dimension Reverse Phase nanoLC:** Peptides were loaded onto a C18 pre-column and then separated on a C18 PepMap100, 3 µm column (Dionex, USA) using the Ultimate

3000 nano HPLC system (Dionex, USA). For the MALDI analysis, a gradient of 10-40% acetonitrile in 0.1% trifluoroacetic acid at a flow rate of 300 nL/min was used with the eluent mixed 1:3 with matrix solution (including Calibration Mixture) and spotted onto a 384 well Opti-TOF® plate (AB SCIEX, USA) using a Probot Micro Fraction Collector (Dionex, USA). For electrospray (ESI) analysis, the gradient was 2-30% acetonitrile in 0.1% formic acid with a flow rate of 400 nL/min and the eluant delivered directly to the NanoSpray® Source.

**MALDI Mass Spectrometry:** The eluant was analyzed using an AB SCIEX 4800 TOF/TOF™ system. Using a Nd:YAG laser at 355 nm and a frequency of 200 Hz, 400 shots per spot were used for MS data acquisition. Using reflector mode in MS/MS mode, 4000 shots per spectrum were acquired. A job-wide interpretation method selected the top 20 peaks per spot but only on the spot where they were most intense and excluding internal calibrants.

**Labeled Reference Plasma Preparation:** A reference plasma sample obtained from healthy volunteers was labeled with 18O water<sup>5</sup> and the same amount spiked into each cohort sample (1:1) prior to LC-MRM/MS analysis to correct for spray efficiency and ionization differences between runs.

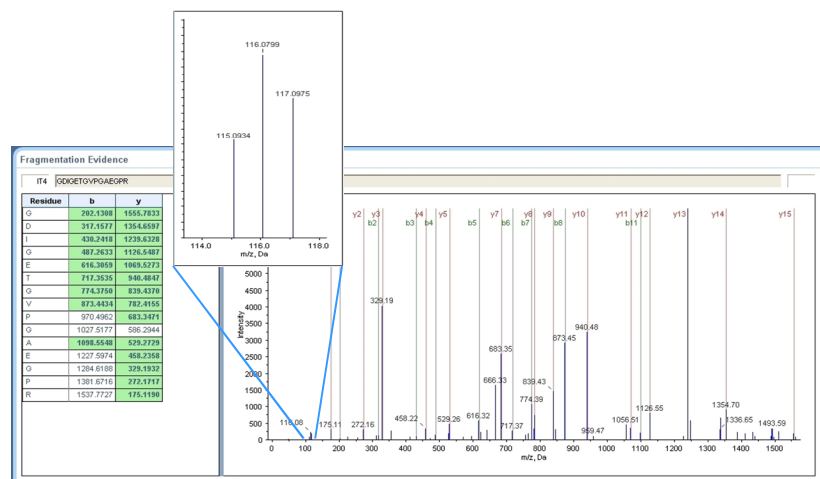
**ESI Mass Spectrometry:** MIDAS™ workflow and MRM assays were developed using an AB SCIEX 4000 QTRAP® system equipped with a NanoSpray® Source using Skyline Software (U Washington, USA). Multiple MRM transitions were developed per peptide for each putative protein biomarker for both the light version of the peptide and the 18O labeled heavy version of the peptide. The pooled verification samples were analyzed using MIDAS™ workflow acquisition method to verify the MRM assay

and do early verification on the sample. High throughput MRM assays were created using the *Scheduled MRM™* algorithm.

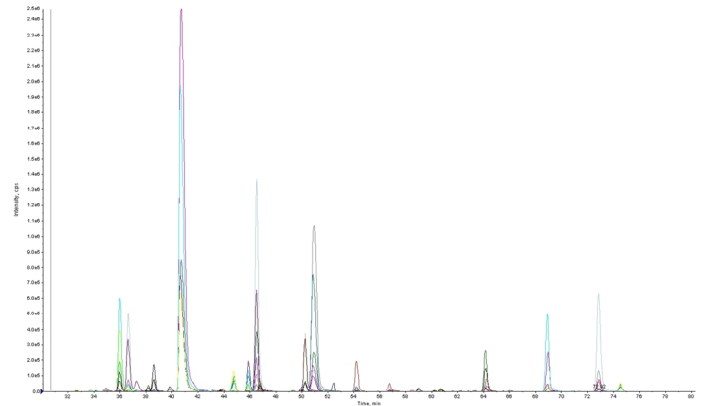
## Discovery using 4800 MALDI TOF/TOF™ Analyzer and iTRAQ® Reagents

iTRAQ® reagent labeled plasma samples from subgroups of each cohort (20 samples) were analyzed by 2D LC MALDI MS/MS. In each sample cohort, 150-200 proteins were identified from over 155,000 MS/MS spectra. A total of 275 proteins were identified and quantified using ProteinPilot™ Software. In the discovery phase of the study, >50 proteins showed statistically significant differences.

Figure 3 shows a MS/MS spectrum of adiponectin from a ProteinPilot software search result. This peptide was significantly differentially expressed as seen by the iTRAQ® reagent reporter ions (Figure 3, inset). The reporter ion peak at m/z 115 is indicative of the level of this peptide found in the healthy control samples. The m/z 116 reporter ion is indicative of the peptide level in the mild diabetic nephropathy sample and similarly the m/z 117 peak represents the severe diabetic nephropathy sample. This adiponectin protein shows an increase in expression in the mild disease case and then decreases as the disease becomes more severe.



**Figure 3: MS/MS Spectrum of a Differentially Expressed Peptide from Adiponectin.** This peptide shows good ion matching for a high confidence peptide identification. The iTRAQ® reagent reporter ions show the peptide to be up and down regulated across the different sample types.



**Figure 4. MRM Assay for 23 Putative Protein Biomarkers.** 32 peptides were analyzed using the MRM assay, which consisted of 4 MRM transitions for each peptide. 10 proteins were unchanging and therefore used as a reference point for normalization between samples.

## MRM Assay Development and Expression Verification using 4000 QTRAP® System

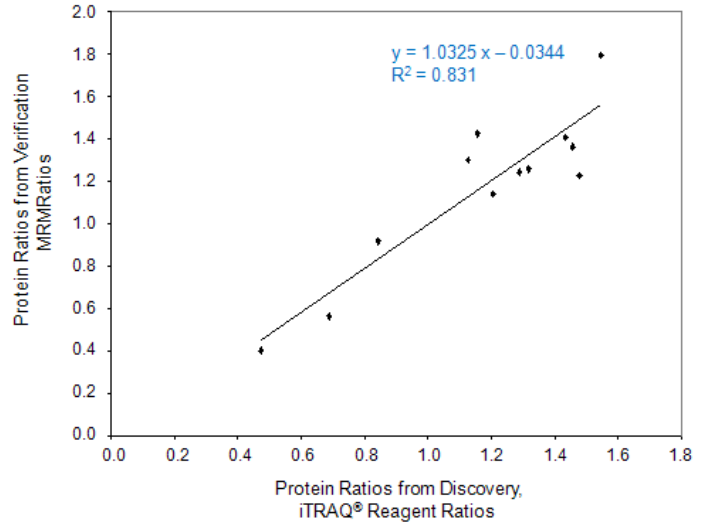
For the first stage of protein verification, an MRM assay was developed for the 50 putative biomarkers using the MIDAS™ workflow strategy. The unique nature of the QTRAP® System allows full scan MS/MS data to be obtained on every MRM peak, enabling rapid assay development and providing identification validation of the MRM assay. Proteins that showed no differential expression between sample types were included to enable normalization between samples. In addition, a reference plasma was introduced that had been labeled with 18O during the digestion step that could be used to determine a relative concentration for each peptide detected. All validated peptides were tested for their suitability to be used with the MRM 18O labeling method.

Using the developed MRM assay, the first cohort consisting of 20 samples pooled for each of the three sample types was analyzed again. Figure 4 shows example data on 32 peptides (190 MRM transitions) for 23 of the proteins in a single MRM run. For an MRM to be accepted in the assay, it had to show good signal intensity, provide a significant identification score during database search of the MS/MS spectra, work well with 18O labeling, and provide CV < 10% across replicate injections. These MRM transitions were passed through to the next stage: validation.

To ensure that the pipeline was effective, the results from the iTRAQ<sup>®</sup> reagent discovery data obtained on the first set of patient samples was correlated with the MRM quantification data obtained from the same set of patient samples. A good correlation between discovery and verification results was observed, with the MRM verification results confirming iTRAQ<sup>®</sup> reagent discovery results using a completely orthogonal workflow and MS strategy (Figure 5).

## Biomarker Validation

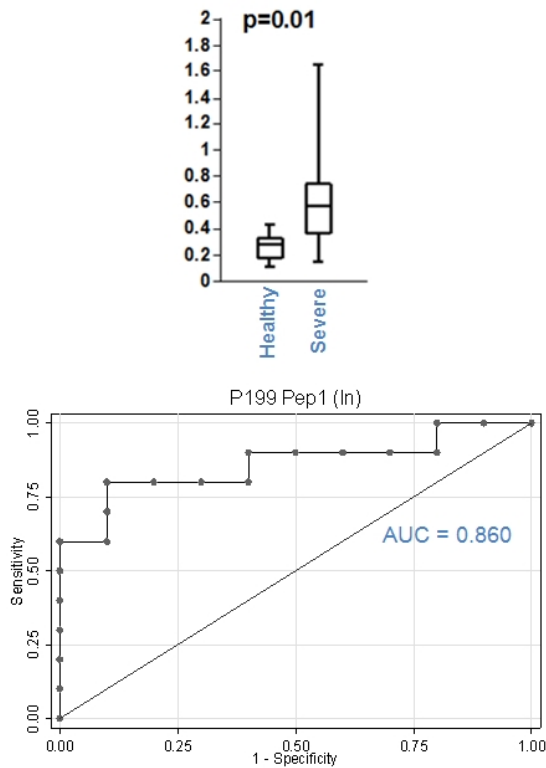
For the final step, the verified MRM assay was now applied to a second cohort of samples. These samples were analyzed as individual samples (not pools). Each sample set consisted of 10 individuals each for healthy controls, mild diabetic nephropathy and severe diabetic nephropathy.



**Figure 5. Strong Correlation Between the Discovery and Verification Protein Expression Results.** Cross validation of protein ratios from discovery iTRAQ<sup>®</sup> reagent data (pool) versus MRM data from the same pooled samples.

A number of statistical tools were used to interpret the data. Disease groups were first compared using box-whisker plots, with statistical significance calculated with the Mann-Whitney test. Statistical transformations were also used to improve power including natural logarithm (ln), inverse (inv) and square root ( $\sqrt{\quad}$ ), prior to plotting Relative Operating Characteristic (ROC) curves (Figures 6 and 7).

An ROC curve is a graphical plot of the sensitivity, or true positive rate, vs. false positive rate (1-specificity or 1 – true negative rate), for a binary classifier system as its discrimination threshold is varied. The conventional interpretation of the clinical significance of the ROC curve area under the curve (AUC) is as follows: ROC curve with AUC < 0.75 indicates poor predictive ability; AUC of 0.75-0.849 indicates potentially interesting; and an AUC of >0.85 indicates a possible marker.



**Figure 6. Assessing the Utility of a Peptide as a Detection Biomarker.** (Top) Box and whisker plot was generated for the peptide P199pep1, showing the healthy controls (left hand box) vs severe disease samples (right hand box). (Bottom) The ROC curve for this peptide was generated to determine the utility of this peptide as a detection biomarker.

## Conclusions

Including discovery and validation, 270 patient samples across the disease states in 3 clinical studies have been analyzed quantitatively. In the study's validation phase, a putative biomarker panel consisting of 13 proteins has been established with MRM assays developed on the AB SCIEX 4000 QTRAP<sup>®</sup> system. Proteins in the panel include ones involved in inflammation, metabolism and oxidative stress. One of the proteins found here to be predictive for nephropathy was Adiponectin. This protein has been found to be a marker for diabetes in a separate large consortium study involving 4 multinational pharmaceutical companies under the direction of the NIH and FDA.

Next steps will involve two parallel tracks to further improve the biomarker panel:

- Developing an algorithm combining the 13 putative protein biomarkers.
- Expanding the number of subjects in the study.

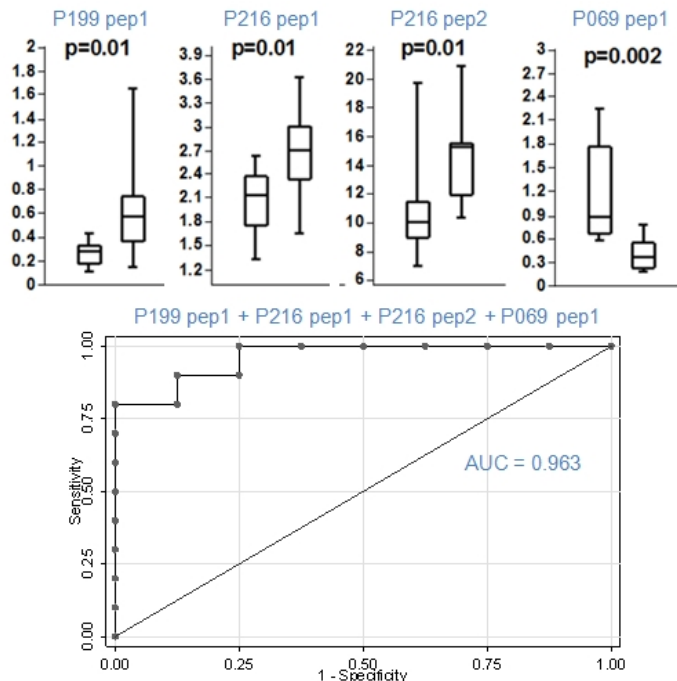
## Acknowledgements

The authors and the Centre for Food and Genomic Medicine would like to thank Fremantle Hospital and the Busselton Diabetes Study for providing all clinical plasma samples and Michael Philips of the Western Australian Institute for Medical Research for statistical analysis.

**For Research Use Only. Not for use in diagnostic procedures.**

© 2011 AB SCIEX. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

Publication number: 4250211-01



**Figure 7. Assessing the Utility of a Putative Biomarker Panel.** (Top) Box and whisker plots were generated for four peptides, showing the healthy controls (left hand box of each) vs. severe disease samples (right hand box of each). (Bottom) The ROC curve for the four peptides was generated to determine the utility of this putative protein biomarker panel.

## References

1. Wild S, Roglic G, Green A, Sicree R, King H. *Diabetes Care* 2004, **27**:1047-1053; and International Diabetes Federation ([www.idf.org](http://www.idf.org)).
2. Harris MI, Klein RE, Welborn TA, et al. *Diabetes Care* (1992) **15**: 815-819.
3. Shields J, Maxwell AP. *Clin Med* (2010) **10**: 500-504.
4. Ellam TJ. *Nephron Clin Pract* (2011) **118**: 324-330.
5. Niles R, Witkowska HE, Allen S, Hall SC, Fisher SJ, Hardt M. *Anal Chem* (2009) **81**(7): 2804–2809.