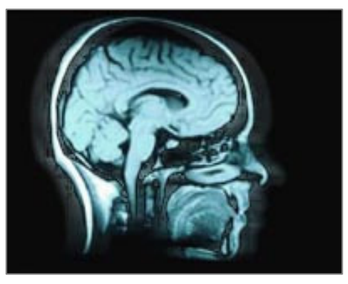
A COMBINED PROTEOMIC AND METABONOMIC APPROACH TO BIOMARKER DISCOVERY FOR SCHIZOPHRENIA USING ACCURATE MASS LC-MS

Thérèse McKenna[1], Hilary Major[1], Jan Claereboudt[3], Jennifer Burgess[1], Christopher Hughes[1], Jeffrey T.-J. Huang[2], Sabine Bahn[2], and James Langridge[1]
[1]Waters Corporation, Manchester, UK; [2]Institute of Biotechnology, University of Cambridge, Cambridge, UK; [3]Waters Central Europe, HRMS Division, Zellik, Belgium

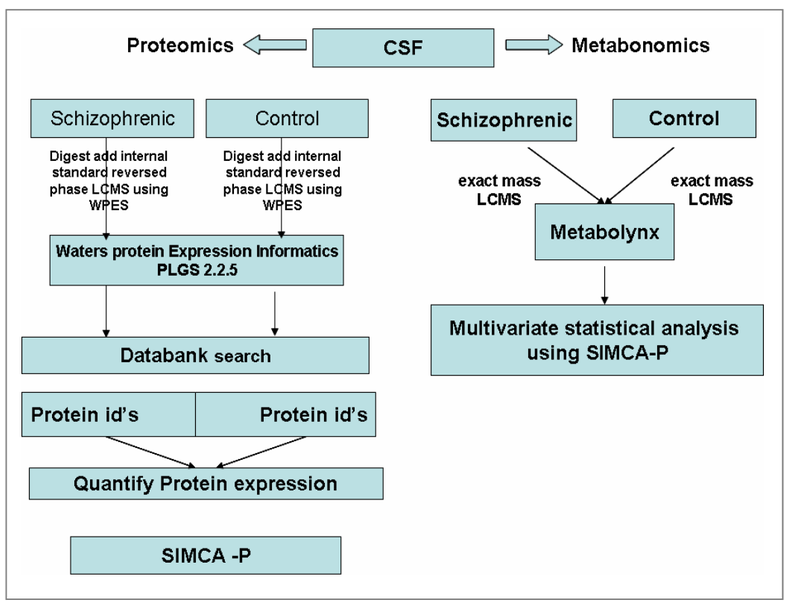
OVERVIEW

- We employed a **'label-free' mass spectrometry-based approach** to analyze metabolic and proteomic profiles from CSF samples from first-onset, drug-naïve paranoid schizophrenia patients and healthy controls.
- Partial least squares discriminant analysis showed a highly significant separation of first-onset, drug-naive schizophrenia patients away from healthy controls in both metabolic and proteomic studies.
- In the **metabonomics study**, potential metabolite biomarkers contributing predominantly to the separation were identified.
- In the **proteomics study**, putative peptide biomarkers were identified and quantified.
- The results indicate a potential association between CSF metabolites and CSF proteins.



OVERVIEW

Schizophrenia is a common, chronic disabling neuropsychiatric disorder which is characterized by hallucinations, delusions, inappropriate affects and bizarre or inappropriate behaviors. Approximately 1% of the population will be affected during their lifetime and as diagnosis of schizophrenia relies on a lengthy procedure of clinical examination of family history, personal history, symptoms and the presence/absence of other disorders the possibility of early intervention and therefore better outcome, is often delayed. An objective diagnostic test would help improve current diagnosis and allow monitoring of individuals over the course of illness and may also prove useful in determining prognosis. The discovery of Biomarkers for schizophrenia is a fundamental step towards a molecular diagnosis of the disease and the delivery of a diagnostic test. In this study we employed a label free mass spectrometry-based approach to analyze the metabolic and protein profiles in cerebrospinal fluid (CSF) samples from first-onset, drug naïve schizophrenia patients and demographically-matched healthy controls.(20 samples in triplicates for proteomic study and 96 samples for the metabonomics study).



Overview of the experimental strategy

METHODS

LC CONDITIONS

For the metabonomic study: 5 µL of neat CSF sample was analyzed.For the proteomic study: 75µL aliquots of each CSF sample were reduced alkylated and digested with in sequencing grade Trypsin (Promega) in the presence of 1% Rapigest SF (Waters Corp.). The resulting mixture was diluted (x5) and spiked with internal standard (yeast Enolase) prior to analysis. Sample sets were analyzed on either a Waters ACQUITY™ Ultra Performance Liquid Chromatography system using a (2.1 x 100) mm ACQUITY BEH C18 or C8 column (1.7 µm particle size) for the metabonomic study or by direct flow nanoLC on a Waters nanoACQUITY™ Ultra Performance Liquid Chromatography system using a Trap cartridge, 180µm ID x 20mm (5µm Symmetry® C18) and a 75µm x 100mm Atlantis™ C18 analytical column (3 µm particle size), at a flow rate of 300nL/min. All samples were separated and eluted using a water/acetonitrile/formic acid gradient.

MS CONDITIONS

Metabonomic mass spectrometric data was collected over a 13 minute period on a Waters LCT Premier™ orthogonal acceleration time of flight (oa-ToF) mass spectrometer operating either in positive ion or negative ion mode at 10,000 resolution (FWHM). The acquisition range was from 50-1500 m/z using an acquisition time of 0.15 seconds and an inter-scan delay of 0.05 seconds. A sample cone voltage of 35V was used throughout to minimise fragmentation. A capillary voltage of 3.0 kV positive ion and 2.6kV in negative ion mode was used for the analysis. The lock mass reference compound in both positive and negative ion analysis was leucine enkephalin introduced through the LockSpray™ reference probe. The instrument was calibrated in both positive and negative ion mode from an infusion of sodium formate solution.

The data was de-convoluted and aligned using the MarkerLynx™ application manager within MassLynx™. The table of m/z and retention time pairs with associated intensities generated was either analysed by principal components analysis (PCA) within MarkerLynx or was exported to SIMCA-P multivariate statistical analysis software for further analysis.

The CSF digests were analysed in triplicate on a Waters Q Tof™ Premier orthogonal acceleration time of flight (oa-ToF) mass spectrometer operating in positive ion V-Optics™ mode at 10,000 resolution (FWHM) using the 'Expression' mode of acquisition described previously (1,2) where LC-MS data is acquired in two functions at alternating collision energy values, 4eV in function 1 and elevated collision energy of 15eV to 40eV in function 2. The instrument was equipped with a NanoLock-Spray™ source from which the lock mass reference compound (Glu Fibrinopeptide B) used for mass alignment was sampled every 30 seconds. Instrument calibration was achieved using a mixture of NaI and CsI.

The resulting data-set was processed using Waters Protein Expression System Informatics (ProteinLynx Global SERVER (PLGS) 2.2.5) and searched against a Human database which contained 21,975 entries and 72,471 random protein entries. Relative expression levels of proteins between the sample groups (disease and control) were obtained using the quantitative capabilities of PLGS 2.2.5. In addition the resulting exact mass and retention time (EMRT) pairs with associated intensities were exported to SIMCA-P multivariate statistical analysis software for further analysis.

RESULTS

METABONOMICS

The PLS-DA scores plots (Figure 1and Figure 2) of both positive and negative ion data (PC 1 against PC 2) show all the diseased samples (red squares) cluster together in the top right of the plots, while the controls (blue circles) are clustered in the bottom left.

The m/z and retention time pairs responsible for any clustering or separation within the data can be readily determined from the associated weightings plot. The compounds responsible for the clustering were principally amino acids and lipids. In addition caffeine RT 3.87 min, observed mass 195.0876, calculated mass = 195.0882 ($C_8H_{11}N_4O_2$) , -0.6mDa, -3.1ppmwas found to be higher in the control population.

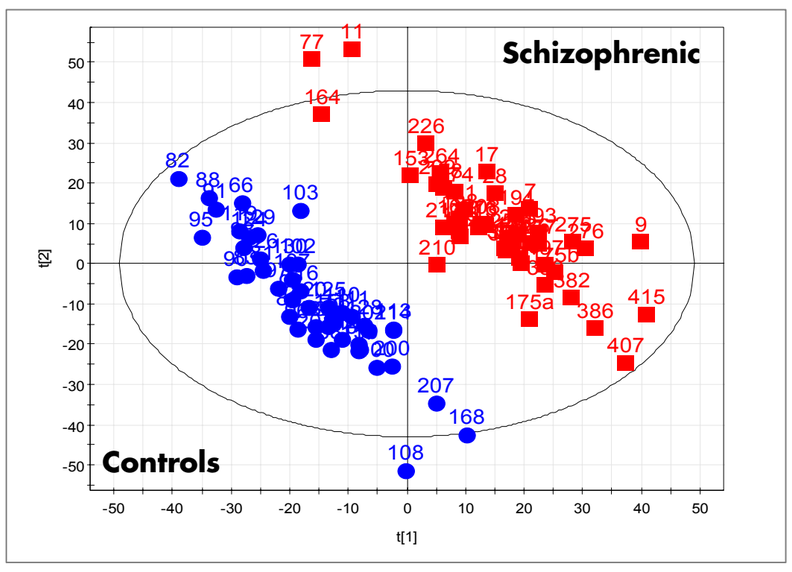


Figure 1 *Metabonomics positive ion data PLS-DA UV scaling*

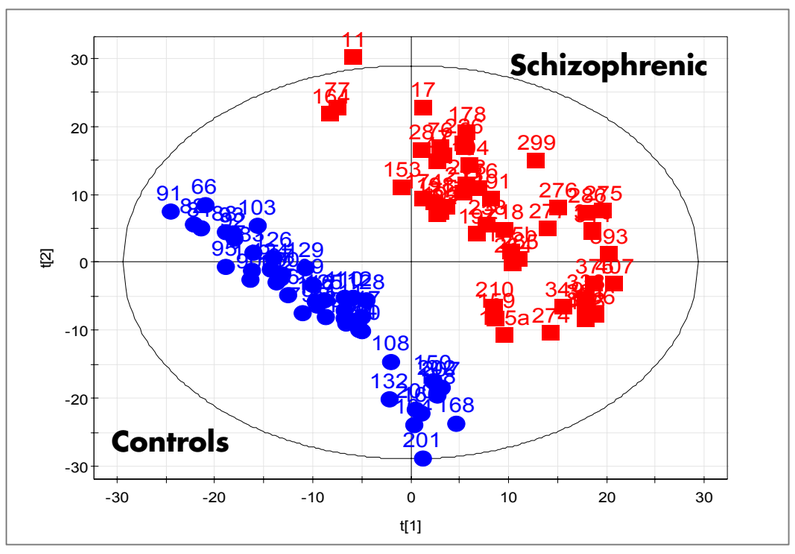


Figure 2. Metabonomics negative ion data PLS-DA UV scaling

PROTEOMICS

Qualitative Proteomic Analysis

Proteins were identified from a databank search using PLGS 2.2.5, for each sample replicate and filtered at the 100% confidence limit. A total of 118 different proteins were identified with an intersection of around 40 proteins. Figure 3 shows the search results from one sample in the results browser.

Figure 3. Results browser showing protein identification summary.

Qualitative Proteomic Analysis

The EMRT analysis in SIMPCA-P is presented below in Figure 4, where a PLS-DA scores plot of PC 1 against PC 2 shows that all the diseased samples (red squares) cluster together in the top right of the plot with the exception of the replicates from a single sample outlier (sample 381) which fall outside of the 95% confidence limit. In contrast the controls are clustered in the bottom left.

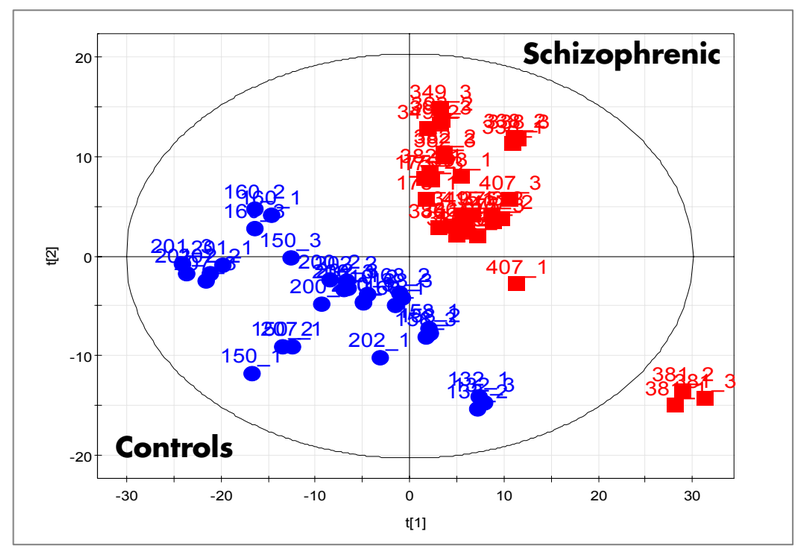


Figure 4. Proteomics data PLS-DA UV scaling

The EMRT components responsible for any clustering or separation within the data can be readily determined from the associated weightings plot and these peptides are currently undergoing further validation.

Quantitative Proteomic Analysis

Protein expression changes between the two groups (disease v control) were determined by comparison of the data at the protein level with all the identified peptides contributing to the fold change of a particular protein (3). Some 34 proteins exhibited changes which are statistically significant between the two groups. 18 were up regulated between 1.15 - 1.9 fold in the disease state and the remaining 16 were down regulated by up to 1.6 fold in the schizophrenic sample. These proteins are currently being subjected to further validation as potential biomarkers of schizophrenia.

The data set can be viewed at the peptide or EMRT level (Figure 5) where the log of the expression ratio of the intensities for peptides from the normal versus the schizophrenic samples are plotted against the average molecular weight for the cluster. The EMRT's coloured red are down regulated in Figure 5a and up regulated in 5b and can be selected and searched to identify the proteins. The proteins identified are a subset of those determined by the qualitative route.

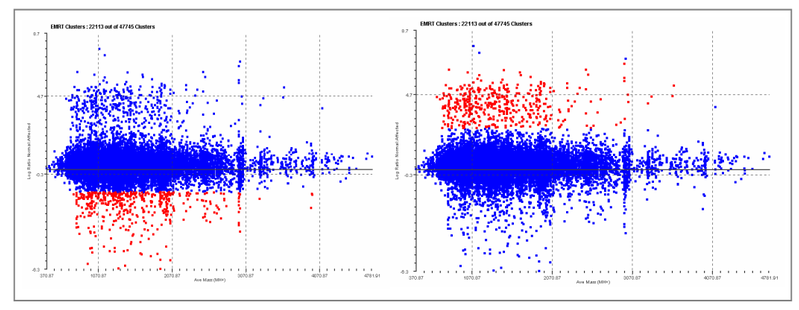


Figure 5a and 5b EMRT plots of log Ratio Normal:Affected with up– and down-regulated EMRT's in red.

CONCLUSIONS

- **A label free, exact mass, mass spectrometry-based approach was used to analyze metabolic and proteomic profiles from CSF samples from first-onset, drug-naïve paranoid schizophrenia patients and healthy controls.**
- **Partial least squares discriminant analysis showed a highly significant separation of first-onset, drug-naive schizophrenia patients away from healthy controls in both metabolic and proteomic studies.**
- **he PLS-DA results indicate an association between CSF metabolites and peptides, suggesting a potential association between metabolites and proteins.**
- **These proteomic and metabolite profiling methods can be employed to identify potential metabolite and protein markers for schizophrenia from CSF.**
- **A total of 118 different proteins were identified with an intersection of around 40 proteins. Of these proteins, 34 proteins exhibited a statistically significant expression changes, 18 were up regulated and 16 were down regulated in the schizophrenic samples.**

References

1. Bateman R et al, J Am Soc Mass Spectrom. 2002 Jul;13(7):792-803.
2. McKenna *et al.* A Novel Approach to Protein Identification: A Direct Comparison to Traditional Mass Spectrometric Techniques. 52nd meeting of the ASMS. Poster presentation (2004).
3. J.C. Silva, R. Denny, C.A. Dorschel, M.V. Gorenstein, I.J. Kass, G.-Z. Li, T. McKenna, M.J. Nold, K. Richardson, P. Young, and S.J. Geromanos Quantitative Proteomic Analysis by Accurate Mass Retention Time Pairs Anal. Chem. 2005, 77, 2187-2200.

TO DOWNLOAD A COPY OF THIS POSTER, VISIT WWW.WATERS.COM/POSTERS