IMSC 2006

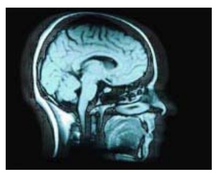
A NOVEL LC-MS APPROACH TO METABOLIC AND PROTEOMIC PROFILING OF CEREBROSPINAL FLUID AND SERUM IN SCHIZOPHRENIA

Thérèse McKenna[1], Hilary Major[1], Christopher Hughes[1], Johannes P.C. Vissers[1], Jeffrey T.J. Huang[2], Sabine Bahn[2], James Langridge[1]

[1]Waters Corporation, Manchester, UK [2] Institute of Biotechnology, University of Cambridge, Cambridge, UK

OVERVIEW

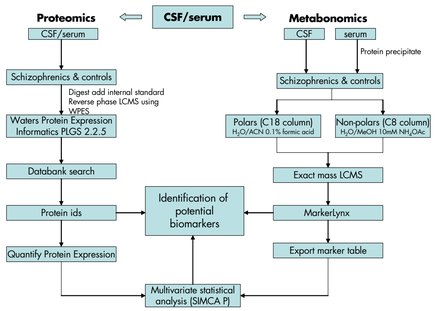
- We employed a **'label-free' mass spectrometry-based approach** to analyze metabolic and proteomic profiles from CSF and sera samples from first-onset, drug-naïve paranoid schizophrenia patients and healthy controls.
- Partial least squares discriminant analysis showed a significant separation of first-onset,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



INTRODUCTION

Schizophrenia, characterized by hallucinations, delusions, inappropriate affects and bizarre or inappropriate behaviors, is a common, chronic and disabling neuropsychiatric disorder which will affect approximately 1% of the population during their lifetime. Current diagnosis of schizophrenia relies on a complicated clinical examination/interview of the patient's family history, personal history, current symptoms (mental state examination) and the presence/absence of other disorders. An objective diagnostic test/tool for schizophrenia would help improve current diagnosis and aid the monitoring of individuals over the course of illness (treatment response, compliance etc.) and may also be useful in determining prognosis. Patients treated at early onset have the best chance of recovery. 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 quantitative 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. As collection of CSF is a highly invasive procedure matched serum samples, which can be collected more routinely, were also analysed for the presence of potential biomarkers.



METHODS

For the **metabonomic** study: the CSF samples were analyzed neat and the sera samples were protein precipitated with 3 volumes of acetonitrile to 1 volume sample prior to analysis. For the **proteomic** study: 75µL aliquots of each CSF sample and 20µL of sera (depleted of the 6 most abundant proteins) were reduced alkylated and digested with sequencing grade Trypsin (Promega) in the presence of 0.1% Rapigest SF (Waters Corp.). The resulting mixture was diluted and spiked with internal standard (yeast Enolase) prior to analysis.

LC CONDITIONS

Metabonomics :Waters ACQUITY™ UPLC System
Column: ACQUITY BEH C18 or C8 column (2.1 x 100mm, 1.7 µm particle size)
Gradient: H_2O/MeCN/formic acid or H2O/MeOH/ ammonium acetate at 0.6mL/min

Proteomics: Waters nanoACQUITY™ UPLC System
Column: Trap cartridge; Symmetry® C18 (180µm x 20mm, 5µm particle size) Analytical column; Atlantis™ C18 (75µm x 100mm, 3 µm) or BEH (75µm x 100mm, 1.7 µm)
Gradient: H_2O/MeCN/formic acid at 300nL/min

MS CONDITIONS

Metabonomics: Waters LCT Premier™ (oa-ToF) MS
Mode: ESI +ve or –ve, 10,000 resolution (FWHM)
Acquisition range: 50-1500Da over 0.15sec and 0.05sec ISD
Capillary voltage: 3.0kV positive, 2.6kV negative
Cone voltage: 35V
Lock reference: leucine enkephalin
Calibration: sodium formate

Proteomics: Waters Q Tof™ Premier (oa-Tof) MS
Mode: ESI +ve at 10,000 resolution (FWHM) using "Expression" mode[1,2]
Collision Energy: Function 1- 4eV; Function 2 - 15 to 40eV
Lock reference: Glu fibrinopeptide B
Calibration: NaI + CsI mix

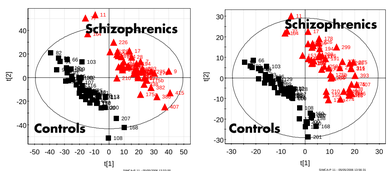
DATA PROCESSING

The **metabonomics** 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 (Umetrics) for further multivariate statistical analysis. The samples from the **proteomics** study were analysed in triplicate and the resulting data-set was processed using Waters Protein Expression System Informatics incorporated in ProteinLynx Global SERVER 2.2.5 (PLGS) 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 (schizophrenic and control) were obtained, together with an estimation of the molar amounts of each protein present, using the quantitative capabilities of PLGS 2.2.5 In addition the resulting exact mass and retention time (EMRT) pairs with associated intensities, protein identifications and molar ratios were exported to SIMCA-P multivariate statistical analysis software for further processing.

RESULTS AND DISCUSSION METABONOMICS

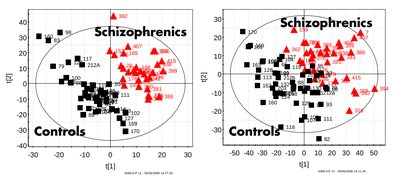
- The PLS-DA scores plots (Figure 1A and B) of the CSF polars data shows all the diseased samples (red triangles) cluster together in the top right of the plots, while the controls (black squares) are clustered in the bottom left.
- The retention times and masses of potential markers were obtained from the associated weights plot (not shown).
- The sera data did not show as good a separation by PLS-DA (see Figure 2A and B) as the CSF samples but this can be explained by greater variability within the sera samples, part of which was associated with the date the samples were collected. However, by using Orthogonal PLS within SIMCA-P it was possible to determine potential markers.
- All 4 sera data sets (polars and non-polars, positive and negative) were combined to build a PLS-DA model by extracting the 18 most important variables.
- PLS regression was repeatedly executed on random subsets of the samples to cross validate the model.
- The variables responsible for the separation were identified by their exact mass measurements and isotope ratios to give a postulated elemental composition.
- Among the postulated markers were peptides, lipids and fatty acid derivatives. Their structures are currently being confirmed by exact mass MS/MS.

Figure 1 Metabonomics PLS-DA analysis of CSF (polars)



A. Positive ion data B. Negative ion data

Figure 2 PLS-DA scores plot of human serum data (positive)

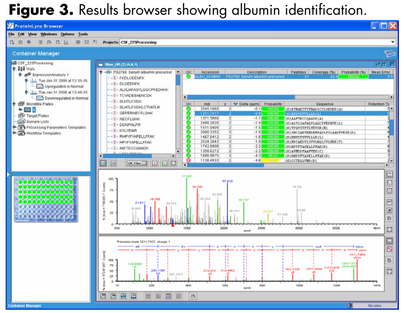


A. Polars data B. Non-polars data

PROTEOMICS

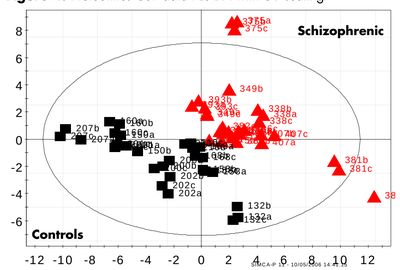
A total of 174 proteins were identified and quantified from more than 2 peptides and in more than 2 replicates across the sample set for the CSF samples. Figure 3 shows the search results from one sample replicate in the results browser.

Figure 3. Results browser showing albumin identification.



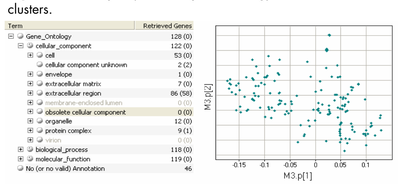
Analysis using SIMPCA-P, shows a PLS-DA scores plot, Figure 4 of PC 1 v PC 2 where the schizophrenic samples (red triangles) cluster together in the top right of the plot with the exception of one replicate from sample 381and the 375 group, which fall outside of the 95% confidence limit. In contrast the controls (black boxes) are clustered in the bottom left corner.

Figure 4. Proteomics CSF data PLS-DA with UV scaling



The proteins responsible for any separation within the data were determined from the associated weights plot and subjected to a gene ontology search. The results are presented in Figure 5 below.

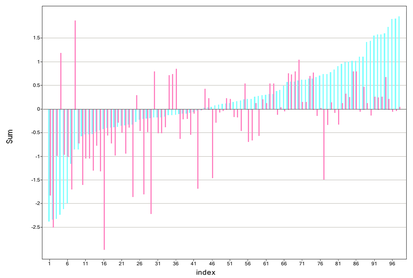
Figure 5. Weights plot with gene ontology search results for all clusters.



Term	Retrieved Genes
Gene_Ontology	128 (0)
cellular_component	122 (0)
cell	53 (0)
cellular component unknown	2 (2)
envelope	1 (0)
extracellular matrix	7 (0)
extracellular region	86 (98)
membrane-enclosed lumen	0 (0)
obsolete cellular component	0 (0)
organelle	12 (0)
protein complex	9 (1)
vesicle	0 (0)
biological_process	118 (0)
molecular_function	119 (0)
No (or no valid) Annotation	46

The difference in protein expression between the schizophrenic and control groups was determined in two ways. Quantitation was performed at the protein level, with all the identified peptides contributing to the fold change of a particular protein (3) and secondly the absolute molar amounts of each protein (after normalisation) was determined, using the method described previously (4). The ^{2}log ratio for 98 proteins which exhibit changes between the groups are plotted in Figure 6. The ratios for male (blue) and female (pink), are ordered from down regulation to up regulation for the schizophrenic male. Of these 98 proteins, 52 were regulated by more than 30%, 34 were up regulated in the schizophrenic male and 18 were down regulated. Differences between the male and female CSF samples can be seen with only 10 of the 34 up-regulated proteins in the male being up regulated in the female schizophrenic samples when compared to the female control. However, there is more similarity between the genders in the number and level of down regulated proteins.
From this study 23 proteins are currently being evaluated as potential biomarkers of schizophrenia. Several of these proteins have been associated with other neurological conditions and many of them were also detected in the serum samples analysed to date.

Figure 6 Relative protein concentration (^{2}log ratio) of the male (blue) and female (pink) schizophrenic CSF samples versus controls.



CONCLUSION

- A label free, exact mass, mass spectrometry-based approach was used to analyze metabolic and proteomic profiles from CSF and sera samples from first-onset, drug-naïve paranoid schizophrenia patients and healthy controls.
- Partial least squares discriminant analysis showed a significant separation of first-onset, drug-naive schizophrenia patients away from healthy controls in both metabolic and proteomic studies.
- The 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 and sera.
- A total of 174 different proteins were identified from the CSF samples. Of these proteins, 98 proteins exhibited a significant expression change, with 52 regulated by more than 30%. In the male samples 34 were up regulated and 18 were down regulated between the schizophrenic and control samples.
- Some gender differences were observed.
- From this study 23 proteins, some associated with other neurological disorders are undergoing further validation.
- 18 metabolites were found to show a significant difference between the schizophrenic and control samples. A further 200 sera samples are currently being analysed to independently validate the metabolite model.

REFERENCES

1. Bateman R *et al*, J Am Soc Mass Spectrom. 2002 Jul;13(7):792-803
2. McKenna T. *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 *et al*. Quantitative Proteomic Analysis by Accurate Mass Retention Time Pairs Anal. Chem. 2005, *77*, 2187-2200
4. Silva, J.C., *et al*. (2006) Absolute quantification of proteins by LCMS^E; a virtue of parallel MS acquisition, Mol. Cell. Proteomics 5, 144-156

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