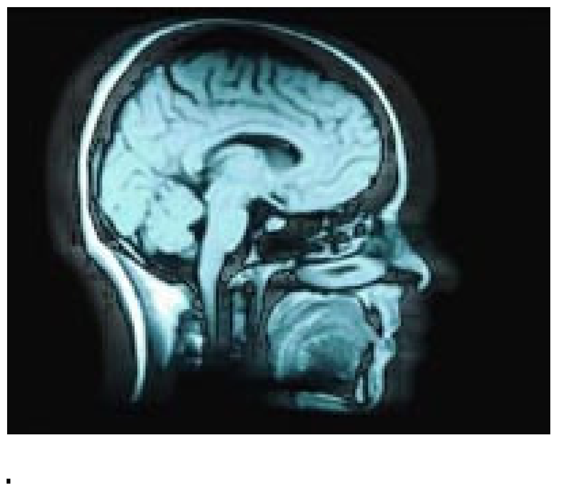
PROTEOMIC PROFILING OF CEREBROSPINAL FLUID AND SERUM IN SCHIZOPHRENIA USING A LABEL FREE MASS SPECTROMETRIC APPROACH

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Michael R Morris[1], Thérèse McKenna[1], Hilary Major[1], Emanuel Schwarz[2], Markus Leweke[2], 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

OVERVIEW

- We employed a **'label-free' mass spectrometry based approach** to determine proteomic profiles from CSF and sera samples from first-onset, drug-naïve paranoid schizophrenia patients and healthy controls.
- Seventy seven proteins were detected with high confidence, in at least three of the samples. Of these, seven proteins were newly identified in CSF.
- Partial least squares discriminant analysis showed a significant separation of first-onset, drug-naive schizophrenia patients away from healthy controls in this study.
- From this study a number of putative biomarker proteins, some associated with other neurological disorders are undergoing further validation.



INTRODUCTION

Schizophrenia is characterized by hallucinations, delusions, inappropriate affects and bizarre or inappropriate behaviours and 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 the present study, we have applied a label-free proteomics methodology to protein biomarker discovery. We examined the reproducibility of protein expression results as well as the sensitivity in detecting CSF proteins in clinical samples.

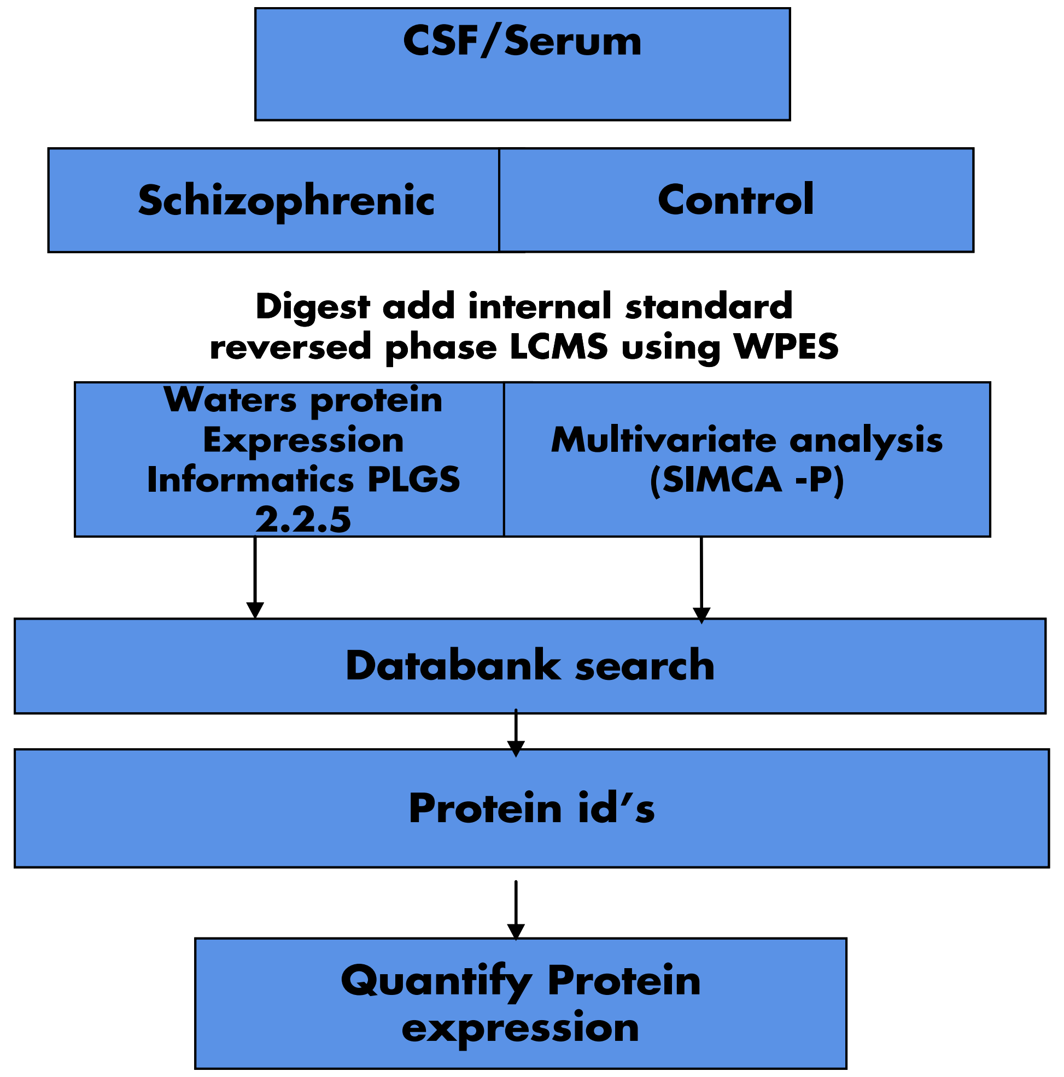
METHODS

CLINICAL SAMPLES

The Ethical committee of the Medical Faculty of the University of Cologne reviewed and approved the protocol of this study and the procedures for sample collection (lumbar puncture) and analysis. All study participants gave their written informed consent. All clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki.

CSF samples were collected from 10 drug-naïve patients diagnosed with first episode paranoid schizophrenia due to duration of illness (DSM-IV 295.30) and from demographically matched healthy volunteers (n=10).

As outlined in Figure 1 below, 75µL aliquots of CSF 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.MA). The resulting mixture was diluted and spiked with internal standard (yeast Enolase) prior to analysis.



CSF/Serum

Schizophrenic | Control

Digest add internal standard reversed phase LCMS using WPES

Waters protein Expression Informatics PLGS 2.2.5 | Multivariate analysis (SIMCA -P)

Databank search

Protein id's

Quantify Protein expression

***Figure 1.** Experimental work flow*

LC CONDITIONS

- 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_20/MeCN/formic acid at 300nL/min

MS CONDITIONS

- 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 samples 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

Statistical analysis

To efficiently evaluate the proteomic variability within and between samples derived from patients and controls, data from each EMRT cluster were exported into SIMCA-P (version 10.5, Umetrics AB, Umeå, Sweden), in which a range of multivariate statistical analyses were conducted. Any EMRT with more than 20 missing value in the total of 60 runs was omitted from the dataset. Initially principal components analysis was applied to the data in order to discern the presence of inherent similarities in spectral profiles. Only one spectrum was excluded from the analysis on the basis of the Hotelling's t-test, which provides a 95% confidence value for a model based on the sample composition (see Results). In order to identify the biomarkers differentiating the patients with schizophrenia from matched controls, partial least square discriminant analysis (PLS-DA) was employed. To avoid overfitting in PLS-DA, cross-validation was applied using a built-in method in SIMCA.

RESULTS AND DISCUSSION

A total of 77 proteins were identified and quantified from more than 3 CSF samples across the sample set. A representative 2D map (exact mass *vs* retention time) from a control sample showing more than 4500 features is presented below in Figure 2.

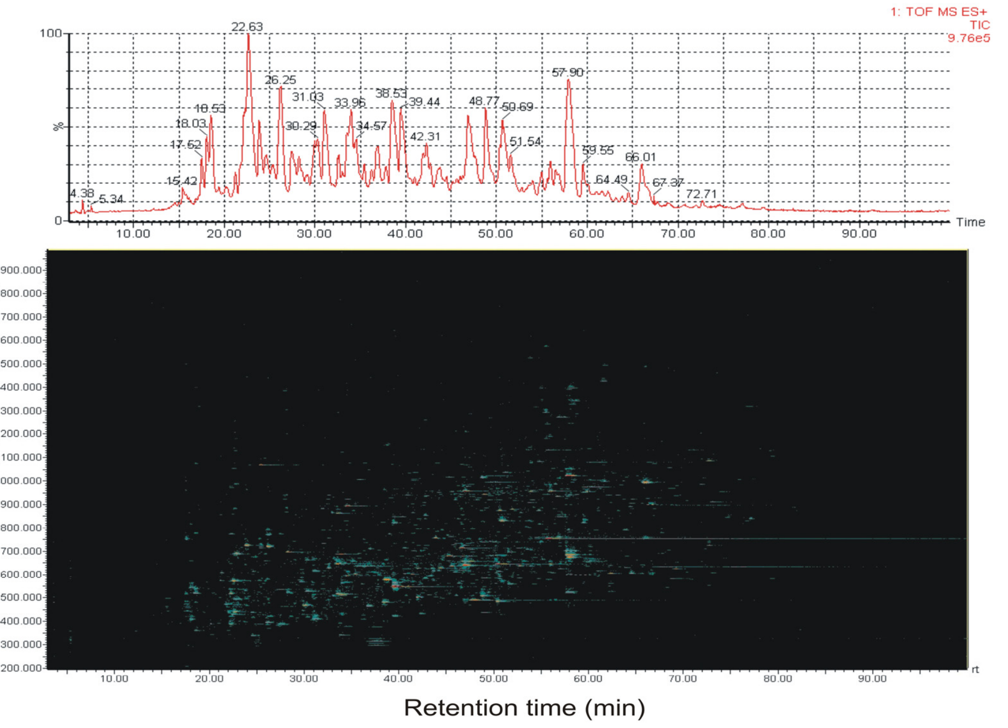


***Figure 2.** 2 Dimensional representation of the data from a control CSF sample. Upper panel shows the associated TIC.*

Figure 3 shows the PLGS results browser obtained following a databank search showing the identification of serum albumin.

***Figure 3.** Results browser showing serum albumin precursor identification from 37 peptides with good mass accuracy.*

The proteins detected in at least three of the CSF samples are shown opposite in Table 1. Among these 77 proteins, 25 were consistently detected in all 60 runs. The majority of these proteins have been reported previously in other studies in which much higher protein levels were used and/or abundant protein species (such as albumin) had been depleted in a prior step. Furthermore, we report 7 newly identified CSF proteins (blue font) which fulfill the acceptance criteria used in this study. These include beta-Ala-His dipeptidase precursor, calsyntenin-1 precursor protein, contactin precursor, MAC-2 binding protein precursor, plasma protease C1 inhibitor precursor, hevin-like protein and protein GS3786. Our results suggest that even without removing the most abundant protein fraction, the method is sensitive enough to detect at least 77 proteins with high confidence in CSF. Further analysis ((SIMPCA-P), produced a PLS-DA scores plot, Figure 4, of PC 1 v PC 2 where the schizophrenic samples (redtriangles) 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.

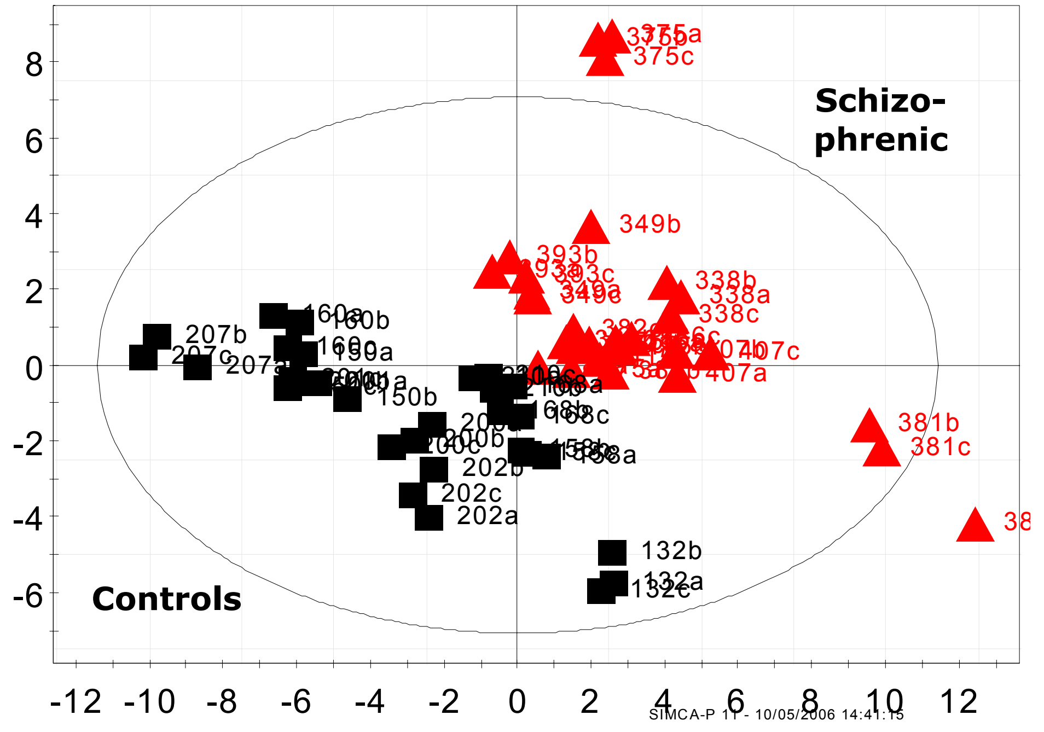
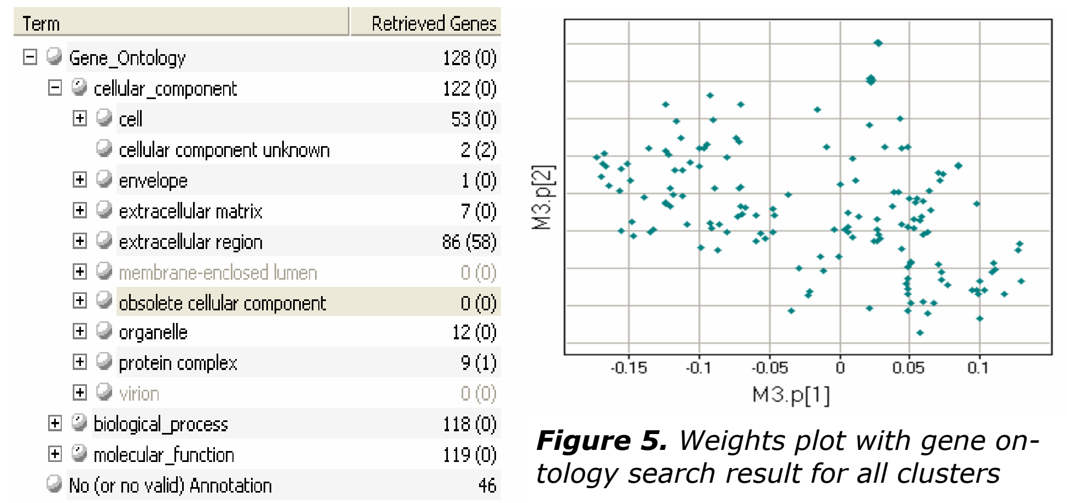


***Figure 4.** Proteomics CSF data PLS-DA with UV scaling*

No.	Entry name	rep rate	Accession number	Description
1	A1AG_HUMAN	60	P02763	Alpha-1-acid glycoprotein 1 precursor (AGP 1) (Orosomucoid 1) (OMD 1)
2	AACT_HUMAN	60	P01011	Alpha-1-antichymotrypsin precursor (ACT)
3	A1AT_HUMAN	60	P01009	Alpha-1-antitrypsin precursor (Alpha-1 protease inhibitor) (Alpha-1-antiproteinase)
4	A2MG_HUMAN	60	P01023	Alpha-2-macroglobulin precursor (Alpha-2-M)
5	ANGT_HUMAN	60	P01019	Angiotensinogen precursor [Contains: Angiotensin I; Angiotensin II]
6	APA1_HUMAN	60	P02647	Apolipoprotein A-I precursor (Apo-AI)
7	APOD_HUMAN	60	P05090	Apolipoprotein D precursor (ApoD)
8	APE_HUMAN	60	P02649	Apolipoprotein E precursor (Apo-E)
9	CLUS_HUMAN	60	P10909	Clusterin precursor (Complement-associated protein SP-40
10	CO3_HUMAN	60	P01024	Complement C3 precursor [Contains: C3a anaphylatoxin]
	CO4_HUMAN	60	P01028	Complement C4 precursor [Contains: C4A anaphylatoxin]
11	CYTC_HUMAN	60	P01034	Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace) (Post-gamma-globulin)
12	DKK3_HUMAN	60	Q9UBP4	Dickkopf related protein-3 precursor (Dkk-3) (Dickkopf-3) (hDkk-3)
13	GELS_HUMAN	60	P06396	Gelsolin precursor
14	HEMO_HUMAN	60	P02790	Hemopexin precursor (Beta-1B-glycoprotein)
15	GC1_HUMAN, GC2_HUMAN,	60	P01857, P01859,	Ig gamma 1-4 chain C region
16	KAC_HUMAN	60	(P01834)	Ig kappa chain C region
	KV2D_HUMAN	10	(P01617)	Ig kappa chain V-II region TEW
	KV3G_HUMAN	40	(P04206)	Ig kappa chain V-III region GOL (Rheumatoid factor)
	KV3L_HUMAN	27	(P18135)	Ig kappa chain V-III region HAH precursor
	KV3M_HUMAN	27	(P18136)	Ig kappa chain V-III region HIC precursor
	KV3B_HUMAN	57	(P01620)	Ig kappa chain V-III region SIE
	KV3D_HUMAN	40	(P01622)	Ig kappa chain V-III region TI
	KV3E_HUMAN	57	(P01623)	Ig kappa chain V-III region WOL
	KV4C_HUMAN	10	(P06314)	Ig kappa chain V-IV region B17 precursor
	KV4B_HUMAN	10	(P06313)	Ig kappa chain V-IV region JI precursor
	KV4A_HUMAN	26	(P01625)	Ig kappa chain V-IV region LEN
	KV40_HUMAN	10	(P06312)	Ig kappa chain V-IV region precursor (Fragment)
17	LAC_HUMAN	60	(P01842)	Ig lambda chain C regions
	LV3B_HUMAN	7	(P80748)	Ig lambda chain V-III region LOI
18	PEDF_HUMAN	60	P36955	Pigment epithelium-derived factor precursor (PEDF) (EPC-1)
19	PGHD_HUMAN	60	P41222	Prostaglandin-H2 D-isomerase precursor (EC 5.3.99.2) (Prostaglandin-D synthase) (Glutathione-independent PGD synthetase) (Prostaglandin D2
20	O43532	60	O43532	RIG-like 7-1
21	TRFE_HUMAN	60	P02787	Serotransferrin precursor (Siderophilin) (Beta-1-metal binding globulin)
22	ALBU_HUMAN	60	P02768	Serum albumin precursor
23	TTHY_HUMAN	60	P02766	Transthyretin precursor (Prealbumin) (TBPA) (TTR) (ATTR)
24	VTDB_HUMAN	60	P02774	Vitamin D-binding protein precursor (DBP) (Group-specific component) (GC-globulin) (VDB)
25	ZA2G_HUMAN	60	P25311	Zinc-alpha-2-glycoprotein precursor (ZN-alpha-2-glycoprotein) (ZN-alpha-2-GP)
26	FBL3_HUMAN	59	Q12805	EGF-containing fibulin-like extracellular matrix protein 1 precursor (Fibulin-3) (FIBL-3) (Fibrillin-like protein) (Extracellular protein S1-5)
27	ALC1_HUMAN, ALC2_HUMAN	59	P01876, P01877	Ig alpha (1-2) chain C region
28	KLK6_HUMAN	59	Q92876	Kallikrein 6 precursor (EC 3.4.21.-) (Protease M) (Neurosin) (Zyme) (SP59)
29	APP1_HUMAN	58	P51693	Amyloid-like protein 1 precursor (APLP)
28	CERU_HUMAN	58	P00450	Ceruloplasmin precursor (EC 1.16.3.1) (Ferroxidase)
29	HV3E_HUMAN	58	(P01766)	Ig heavy chain V-III region BRO
	HV3D_HUMAN	13	(P01765)	Ig heavy chain V-III region TIL
	HV3C_HUMAN	41	(P01764)	Ig heavy chain V-III region VH26 precursor
30	A1BG_HUMAN	57	P04217	Alpha-1B-glycoprotein
31	A1AH_HUMAN	56	P19652	Alpha-1-acid glycoprotein 2 precursor (AGP 2) (Orosomucoid 2) (OMD 2)
32	O43505	56	O43505	N-acetyllactosaminide beta-1
33	SG1_HUMAN	54	P05060	Secretogranin I precursor (SGI) (Chromogranin B) (CgB) [Contains: GAWK peptide; CCB peptide]
34	FBL1_HUMAN	52	P23142	Fibulin-1 precursor
35	CMGA_HUMAN	50	P10645	Chromogranin A precursor (CGA) (Pituitary secretory protein I) (SP-I) [Contains: Vasostatin I; Vasostatin II; EA-92; ES-43; Pancreastatin; SS-18; WA-8; WE-14;
36	APOH_HUMAN	49	P02749	Beta-2-glycoprotein I precursor (Apolipoprotein H) (Apo-H) (B2GPI) (Beta(2)GPI) (Activated protein C-binding protein) (APC inhibitor)
37	B2MG_HUMAN	49	P01884	Beta-2-microglobulin precursor
38	A4_HUMAN	47	P05067	Alzheimer's disease amyloid A4 protein precursor (Protease nexin-II) (PN-II) (APPI) [Contains: Beta-amyloid protein (Beta-APP) (A-beta)]
39	HPT1_HUMAN	47	P00737	Haptoglobin-1 precursor
40	HPT2_HUMAN	46	P00738	Haptoglobin-2 precursor
41	HPTR_HUMAN	30	P00739	Haptoglobin-related protein precursor
42	ANT3_HUMAN	44	P01008	Antithrombin-III precursor (ATIII)
43	CNDP1_HUMAN	44	Q96KN2	Beta-Ala-His dipeptidase precursor
44	OSTP_HUMAN	42	P10451	Osteopontin precursor (Bone sialoprotein 1) (Urinary stone protein) (Secreted phosphoprotein 1) (SPP-1) (Nephropontin) (Uropontin)
45	VTNC_HUMAN	40	P04004	Vitronectin precursor (Serum spreading factor) (S-protein) (V75) [Contains: Vitronectin V65 subunit; Vitronectin V10 subunit; Somatomedin B]
46	A2HS_HUMAN	38	P02765	Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Alpha-2-Z-globulin) (Ba-alpha-2-glycoprotein)
47	CSTN1_HUMAN	35	O15179	NRCAM PROTEIN
48	RNP_HUMAN	35	P07998	Ribonuclease pancreatic precursor (EC 3.1.27.5) (RNase 1) (RNase A) (RNase UPI-1) (RIB-1)
49	O94985	32	O94985	Calsyntenin-1 precursor protein
50	KNG_HUMAN	31	P01042	Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains: Bradykinin]
51	Q9UQS6	28	Q9UQS6	Fibronectin (Fragment)
52	Q08380	27	Q08380	MAC-2 BINDING PROTEIN PRECURSOR (LECTIN
53	CFAH_HUMAN	25	P08603	Complement factor H precursor (H factor 1)
54	SODE_HUMAN	25	P08294	Extracellular superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1) (EC-SOD)
55	NPC2_HUMAN	23	Q15668	Epididymal secretory protein E1 precursor (Niemann-Pick disease type C2 protein) (EPI-1) (hE1) (Epididymal secretory protein 14.6) (ESP14.6)
56	OIF_HUMAN	23	P20774	Osteoinductive factor precursor (OIF) (Osteoglycin) (Mimecan)
57	PLTP_HUMAN	23	P55058	Phospholipid transfer protein precursor (Lipid transfer protein II)
58	CFAB_HUMAN	22	P00751	Complement factor B precursor (EC 3.4.21.47) (C3/C5 convertase) (Properdin factor B) (Glycine-rich beta glycoprotein) (GBG) (PBF2)
59	Q9UHG2_HUMAN	21	Q9UHG2	PROSAAS PRECURSOR (GRANIN-LIKE NEUROENDOCRINE PEPTIDE PRECURSOR)
60	PLTP_HUMAN	19	P55058	Phospholipid transfer protein precursor
61	G786_HUMAN	17	Q92520	Protein GS3786
62	CONT_HUMAN	16	Q12860	Contactin precursor (Glycoprotein gP135)
63	SPRL1_HUMAN	15	Q14515	Hevin-like protein
64	IBP7_HUMAN	15	Q16270	Insulin-like growth factor-binding protein 7 precursor (IGFBP-7),Prostacyclin-stimulating factor
65	THY1_HUMAN	15	P04216	Thy-1 membrane glycoprotein precursor (Thy-1 antigen) (CDw90) (CD90 antigen)
66	SAP3_HUMAN	14	P17900	Ganglioside GM2 activator precursor (GM2-AP) (Cerebroside sulfate activator protein) (Shingolipid activator protein 3) (SAP-3)
67	LAMP_HUMAN	14	Q13449	Limbic system-associated membrane protein precursor (LSAMP)
68	NCA1_HUMAN	14	P13591	Neural cell adhesion molecule
69	NCA2_HUMAN	13	P13592	Neural cell adhesion molecule
70	RETB_HUMAN	10	P02753	Plasma retinol-binding protein precursor (PRBP) (RBP)
71	NEL2_HUMAN	10	Q99435	Protein kinase C-binding protein NELL2 precursor (NEL-like protein 2) (Nel-related protein 2)
72	O15240	9	O15240	Neuro-endocrine specific protein VGF
73	IC1_HUMAN	8	P05155	Plasma protease C1 inhibitor precursor (C1 Inh) (C1Inh)
74	THRB_HUMAN	8	P00734	Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II)
75	HRG_HUMAN	6	P04196	Histidine-rich glycoprotein precursor (Histidine-proline rich glycoprotein) (HPRG)
76	PLMN_HUMAN	6	P00747	Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin]
77	TETN_HUMAN	6	P05452	Tetranectin precursor (TN) (Plasminogen-kringle 4 binding protein)

The proteins responsible for 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.



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 (58)
⊞ membrane-enclosed lumen	0 (0)
⊞ obsolete cellular component	0 (0)
⊞ organelle	12 (0)
⊞ protein complex	9 (1)
⊞ virion	0 (0)
⊞ biological_process	118 (0)
⊞ molecular_function	119 (0)
No (or no valid) Annotation	46

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

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). From this study a number of 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 can be detected in the serum samples analysed to date.

CONCLUSION

- A label free, exact mass, mass spectrometry-based approach was used to investigate the proteomic profiles from CSF.
- Without prior depletion of abundant protein species, using only 0.3µL of CSF per sample and with technical and biological variation controlled, seventy seven proteins were detected with high confidence, of which seven were newly identified.
- Twenty five of these proteins were detected consistently across all sixty replicates
- 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.
- Some gender differences were observed.
- From this study a number of proteins, some associated with other neurological disorders are undergoing further validation.

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