Proteomic screening of cerebrospinal fluid
Rosenling, Ann Therese Isabell

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2010

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 18-10-2019
Chapter 4
Profiling and identification of Cerebrospinal Fluid (CSF) proteins in a Rat EAE Model of Multiple Sclerosis

Therese Rosenling\textsuperscript{1}, Marcel Stoop\textsuperscript{2}, Amos Attali\textsuperscript{3}, Hans van Aken\textsuperscript{3}, Ernst Suidgeest\textsuperscript{3}, Christin Christin\textsuperscript{1}, Christoph Stingl\textsuperscript{2}, Frank Suits\textsuperscript{4}, Peter Horvatovich\textsuperscript{1}, Rogier Hintzen\textsuperscript{2}, Tinka Tuinstra\textsuperscript{3}, Rainer Bischoff\textsuperscript{1}, Theo Luider\textsuperscript{2}

\textsuperscript{1}Department of Analytical Biochemistry, Centre for Pharmacy, University of Groningen, Groningen, The Netherlands
\textsuperscript{2}Department of Neurology, Erasmus University Medical Center, Rotterdam, The Netherlands
\textsuperscript{3}Abbott Healthcare Products B.V., Weesp, the Netherlands
\textsuperscript{4}IBM TJ Watson Research Centre, Yorktown Heights, NY, USA

\textit{Manuscript in preparation}
Abstract

The experimental autoimmune encephalomyelitis (EAE) model resembles certain aspects of multiple sclerosis (MScl), with common features such as motor dysfunction, axonal degradation and infiltration of T-cells during the affected period. In MScl the main cause of the disease yet to be discovered and there is still no specific molecular biomarker in use for early diagnosis and patient classification. We here chose to study the EAE rat model to discover novel MScl related proteomic biomarker candidates using a screening approach with the goal to translate them into the human situation.

EAE was induced in male Lewis rats by injection of myelin basic protein (MBP) together with complete Freund’s adjuvant (CFA). An inflammatory control group was injected with CFA alone and a non-treated group served as healthy control. CSF was collected from the cisterna magna at day 10 (disease onset) and 14 (peak of symptoms) after treatment and analyzed by bottom-up proteomics. The analysis was based on Orbitrap LC-MS and QTOF LC-MS platforms in two independent laboratories. The obtained data were processed in a crossover design by a combination of two different approaches including our in-house developed data processing pipeline, the Progenesis pipeline and the nearest shrunken centroid classifier (NSC) as well as principal component analysis (PCA). By combining results, we discovered 41 proteins significantly increasing in EAE animals compared to the control groups, 28 of these have not been discovered as discriminatory in the EAE model before. Among the discriminatory proteins are Vitamin D binding protein, lysozyme C1, fetuin B, T-kininogen, serum paraoxonase/arylesterase 1, glutathione peroxidase 3 and afamin. To our knowledge, this is the first study of its kind focusing on a screening approach of rat CSF from the EAE model.
1. Introduction

Multiple Sclerosis (MScl) is an autoimmune, demyelinating, neurodegenerative disorder of the central nervous system (CNS) (1). Presently there is no rapid and specific diagnostic test that excludes other diseases or that can distinguish the different sub-classes of MScl from each other. Diagnosis is currently based on clinical evaluation of the patient together with evidence of lesions by magnetic resonance imaging (MRI) as well as the presence of non-specific oligoclonal IgG bands in CSF (2, 3).

The discovery of reliable molecular biomarkers to diagnose and classify MScl is of great importance to initiate in an early phase the most appropriate therapy in order to at least slow disease progression and to improve the quality of life of MScl patients. An additional value of MScl-specific molecular biomarkers is the possibility of gaining further insight into the disease mechanism and to monitor the effect of new drugs that are developed for MScl.

Although the cause of MScl remains elusive, studies based on the experimental autoimmune encephalomyelitis (EAE) rat model have revealed many pathogenic processes. The EAE model mimics several aspects of the disease such as, axonal and neuronal degradation, motor dysfunction in the lower limbs (4-6), demyelination, infiltration of CD4+/CD8+ T-cells, activation of microglia and macrophages, blood-brain-barrier (BBB) rupture and release of inflammatory cytokines (e.g. TNF-α and IL1-β) (4-8). However, the EAE model still remains a model and it does not cover all clinical aspects of MScl. Most proteomic studies based on the EAE model have been hypothesis-driven (8-16), while only a few large-scale whole proteome screening studies have been performed to date on affected tissue such as spinal cord an lesions (17-20). Until now, no studies are described that look into the proteome of rodent EAE models to identify and quantify EAE related proteins in CSF.

MScl lesions are usually found in the periventricular white matter and at the surface of the spinal cord (21) which are both in contact with cerebrospinal fluid (CSF). This makes CSF an ideal compartment to discover biomarkers. In the clinic CSF is collected from patients with relative ease albeit not as readily as blood. CSF has been used for MScl research in humans for many years (22). The EAE model, on the contrary, has not been studied extensively (8) from CSF. This is probably due to the relatively small volume of CSF that can be withdrawn from rodents (for rats < 60 µL (unpublished results) and mice < 15 µL (23)) without blood contamination of hemoglobin and apolipoprotein B100. It is obvious, that analysis of such small volumes requires sensitive, miniaturized techniques and highly skilled operators.
Here we report a comprehensive whole proteome analysis of CSF from MBP-EAE Lewis rats using two LC-MS/MS platforms in two independent laboratories. Samples were taken at the onset of disease symptoms (day 10) and at the apex of symptoms (day 14) to detect early-indicator, disease-specific proteins and link them to hallmarks of MScl disease and pathogenesis.

2. Material and Methods

2.1. Induction of acute EAE in the Lewis rat

At the start of the study (day 0) EAE was induced in 30 male Lewis rats (Harlan Laboratories B.V.) with a starting weight of 175 – 200 g, by injection of 100 μL saline-based emulsion containing 20 μg guinea pig myelin basic protein (MBP), 500 μg *Mycobacterium tuberculosis* type 37HRa (Difco) and 50 μL Complete Freunds’ Adjuvant (CFA) (EAE groups; E and F). Also at day 0, 30 additional male Lewis rats received the same saline-based emulsion as above but without added MBP (inflammatory control groups; C and D). The emulsion was injected subcutaneously in the left hind paw under isoflurane anesthesia. Another 30 rats were kept as untreated controls undergoing anesthesia only (healthy control groups; A and B). The animal groups and samples included in the experimental design of the present study are summarized in Table 1.
Table 1. Samples included in the study. Sample data that were analyzed by Orbitrap only are marked O, sample data analyzed by QTOF only are marked Q. Group E was divided in two groups according to the TIC area $E_{\text{high}}$ (high TIC) and $E_{\text{low}}$ (low TIC). Samples marked were not included because of blood contamination (samples 62, 78, 81). Sample 80 was excluded because the rat did not show visual signs of disease (considered as outlier).

<table>
<thead>
<tr>
<th>Group</th>
<th>Samples analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Healthy control day 10</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>21 O</td>
</tr>
<tr>
<td>7</td>
<td>22</td>
</tr>
<tr>
<td>8</td>
<td>23 O</td>
</tr>
<tr>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>11</td>
<td>26</td>
</tr>
<tr>
<td>12 O</td>
<td>27</td>
</tr>
<tr>
<td>13</td>
<td>28</td>
</tr>
<tr>
<td>14</td>
<td>29</td>
</tr>
<tr>
<td>15</td>
<td>30</td>
</tr>
</tbody>
</table>
The animals were randomized for treatment and individually marked. Three animals were kept in each cage (type III cages) where food and water were available ad libitum. Disease symptoms and weights of all animals were recorded daily. The following scores for motor dysfunction were used; 0: no pathological signs, 1: paralysis of the tip of the tail, 2: no curling reflex/no strength at tail basis, 3: slightly impaired walking, 4: unstable walk, 5: half hind limb paralysis, 6: complete hind limb paralysis, 7: midriff paralysis, 8: half fore limb paralysis, 9: moribund, 10: death due to EAE.

From each of the 3 groups (EAE, inflammatory control, non-treated control) 15 animals were sacrificed at day 10 (disease onset) (groups A, C and E) and another 15 rats were sacrificed at day 14 (disease peak) (groups B, D and F). CSF was also collected and pooled from 10 healthy animals to be used as quality control (QC) sample during LC-MS analysis, for protein concentration measurements, and for the set-up of protocols and analysis methods. All animal experiments described in this study were approved by the local Ethical Committee for Animal Experiments.

2.2. CSF Sampling
At day 10 and 14, the animals were euthanized using CO₂/O₂, the head of the rats were fixed using a holder to reveal the Arachnoid membrane and a skin incision was made followed by a horizontal incision in the musculus trapezius pars descendens. A maximum volume of 60 µL CSF was obtained by direct insertion of an insulin syringe needle (Myjector, 29 G × 1/2" - 0.33 × 12 mm, 0.3 mL = 30 units) via the arachnoid membrane into the cisterna magna. Within 20 minutes after collection, each sample was centrifuged at a force of 2000 g for 10 minutes at 4 °C. After centrifugation, the supernatant was aliquoted in five tubes of ~10 µL each and stored (up to 6 months) at -80 °C until further analysis. Samples that according to visual control as well as to the detection of hemoglobin-derived peptides (analysis by Orbitrap LC-MS/MS) were considered to be contaminated with blood and were subsequently excluded from the study. The samples analyzed by the two platforms are listed in Table 1.

2.3. Sample preparation
Protein digestion was performed as follows: 10 µL CSF and 10 µL 0.1% RapiGest™ (Waters) dissolved in 50 mM ammonium bicarbonate were added to a sample tube. The sample was reduced by adding 0.5 µL 1,4-dithiotreitol (DTT) (0.5 M) followed by incubation at 60 °C for 30 min. After cooling to room temperature the sample was alkylated with 1 µL iodoacetamide (IAM) (0.3 M) in the dark for 30 min at room temperature. To the sample 1 µL sequencing grade modified
trypsin (Promega, Madison, WI, USA, part # V5111) (1 µg/µL) (enzyme to protein ratio; 1:10-50) was added and the sample was incubated for ~16 h at 37 °C under slight agitation (450 rpm). Thereafter, 3 µL hydrochloric acid (0.5 M) was added followed by incubation for 30 min at 37 °C. The sample was centrifuged at 13250 g for 10 min at 4 °C to remove precipitated, hydrolyzed RapiGest™. The samples were transferred to sample vials and analyzed in random order. Each of the samples was exposed to three freeze-thaw cycles prior to LC-MS analysis.

2.4. ChipLC-QTOF MS proteomic analysis

The sample order was randomized with quality controls (cytochrome C spiked into the pooled CSF sample after digestion at 200 fmol/µL) and blanks injected after every tenth sample. The sample injection volumes were normalized according to the TIC area; a first injection of 0.5 µL of samples from all group revealed a TIC area that was 5 times higher in group E and F compared to group A-D. To normalize the TIC area a five times lower volume was injected of group E and F (0.2 µL) compared to group A-D (1 µL). Peptides were separated on a reverse phase LC-chip (Protein ID chip #3; G4240-63001 SPQ110: Agilent Technologies; separating column: 150 mm × 75 µm Zorbax 300SB-C18, 5 µm; trap column: 160 nL Zorbax 300SB-C18, 5 µm) coupled to a nano LC system (Agilent 1200) with a 40 µL injection loop. Ions were generated by electrospay ionization (ESI) and transmitted to a quadrupole-time-of-flight (QTOF) mass spectrometer (Agilent 6510). Instrumentation was operated under the MassHunter Data Acquisition software (version B.01.03; Build 1.3.157.0; Agilent Technologies, Santa Clara, USA).

For LC separation the following eluents were used; A: ultra-pure water (conductivity 18.2 MΩ, Elga Labwater, Ede, The Netherlands) with 0.1% formic acid (98-100%, pro analysis, Merck, Darmstadt, Germany); B: acetonitrile (HPLC-S gradient grade, Biosolve, Valkenswaard, The Netherlands) with 0.1% formic acid. The samples were injected on the trap column at a flow rate of 3 µL/min (3% B). After 10 min the sample was transferred to the separation column at a flow rate of 250 nL/min and the peptides were eluted using the following gradient: 100 min linear gradient from 3 to 50% B; 5 min linear gradient from 50 to 70% B; 4 min linear gradient from 70 to 3%.

The MS analysis was done in the 2 GHz extended dynamic range mode under the following conditions; mass range: 275-2000 m/z, acquisition rate: 1 spectrum/sec, data storage: profile and centroid mode, fragmentor: 175 V, skimmer: 65 V, OCT 1 RF Vpp: 750 V, spray voltage: ~1900 V, drying gas temp: 325 °C, drying gas flow (N₂): 6 L/min. Mass correction was done during analysis using internal standards; m/z: 371.31559 (originating from a ubiquitous background ion (Diocyl adipate, DOA, plasticizer) and m/z: 1221.990637 (HP-1221 calibration...
Proteome profiling in a rat EAE model

standard, continuously evaporating from a wetted wick inside the spray chamber).

Repeatability of the LC-MS analyses was assessed in terms of mass accuracy, retention time and peak area based on four selected peptide peaks originating from a cytochrome C digest spiked into the QC sample. The peaks were smoothed (Gaussian function width; 15 points, [15 sec]) and integrated to calculate the relative standard deviation (RSD) of the peak area (< +/- 30%) and the retention time (< +/- 0.3% or 5 sec). Mass accuracy calculated as the mean of five measurements from each selected cytochrome C peak, was within +/- 5 ppm of the expected value.

For protein identification, samples were analyzed by auto MS/MS in the 4 GHz mode using the following settings; fragmentor: 175 V, skimmer: 65 V, OCT 1 RF Vpp: 750 V, precursor ion selection: medium (4 m/z), mass range: 200-2000 m/z, acquisition rate for both MS and MS/MS: spectrum/sec, isolation width: ~ 4 amu, ramped collision energy 3.7 V/100 Da, offset: 2.5 V, precursor setting: maximum 3 precursors/cycle, absolute threshold: 1000, relative threshold: 0.01% of the most intense peak, active exclusion enabled after 2 selections, release of active exclusion after 0.5 min, precursors sorted by abundance only. The MS/MS files were stored in centroid mode. The LC separation was identical except for a slightly modified enrichment and gradient elution program. After injection, peptides were enriched for 15 min on the trap column before being transferred to the separation column. Elution was done using the following gradient: 97 min linear gradient from 3 to 60% B; 15 min linear gradient from 60 to 3% B.

2.5. ChipLC-QTOF MS data processing

Raw data were visually examined and exported to MzData.XML files in centroid mode using the Agilent MassHunter Qualitative Analysis software (Agilent version B.01.03; Build 1.3.157.0). Files were converted to ASCII format. In order to limit file size, thresholds were used for the conversion (conversion of data was limited to the time and mass range corresponding to eluting peptides and data were converted with intensity filtering to reduce background noise) according to following parameter settings; peak intensity: > 400 counts, mass range: 275-1600 m/z, retention time range: 22-95 min. The ASCII files were pre-processed using an in-house data processing workflow developed in C++ (24, 25) producing a common peak matrix of 5336 peaks (based on selection of 10 000 most intensive peaks from each data file, this number is selected based on the quality of the data and the noise threshold). The raw peak matrix was analyzed by principal component analysis (PCA) to visualize the overall variability prior to feature selection and analysis with double cross validated Nearest Shrunken Centroid (NSC) algorithm (26) (MatLab R2009a, Mathworks, Natick, MA, USA).
The NSC classifier was applied on the generated peak matrix in order to find the most discriminatory peptide peaks in the comparisons; A vs. E, C vs. E, B vs. F and D vs. F. Double cross-validation was implemented by counting classification error rate in function of the shrinkage in the inner loop and to measure the classification performance in the outer loop. When the lowest classification error rate was recorded for broad continuous range of shrinkage, the optimal shrinkage value was selected manually in order to generate the peak list. The NSC-selected features were subsequently visualized by PCA.

NSC-selected peaks were subsequently visualized using extracted ion chromatograms (EICs) with the MassHunter software. EICs for selected m/z values were smoothed and integrated, and the generated peak areas were exported in Excel format. Peak areas from both the peak matrix as well as the chromatograms were compared between groups using one way ANOVA with Bonferroni correction (SPSS Inc., Chicago, IL, USA). Peak area differences with a p-value below 0.01 were considered significant.

For processing of the same raw data with a different workflow, the data files were converted to MzXML format using the open source msconvert parser in the proteowizard pipeline (27). The generated files were pre-processed in terms of feature extraction, retention time alignment and matching the same peaks across multiple chromatograms using the Progenesis LC-MS software package (version 2.5, Nonlinear Dynamics, Newcastle-upon-Tyne, United Kingdom) generating a raw peak matrix. The resulting peak matrix was exported in Excel format (Microsoft Office Excel 2007, Microsoft Corporation, Redmond, WA, USA) and the 10 000 most intensive peaks were selected for further analysis. The NSC classifier was applied and the same procedure as described above was applied on the data set.

Discriminatory features were searched for matching m/z, RT and charge states (Δ RT: max ± 1 min, Δ m/z: max ± 10 ppm) in an Excel file containing all identified ions from the auto MS/MS analysis (see following section). After having assigned a discriminatory peptide peak to a given protein, all other identified peptides from that protein were examined to evaluate whether they followed the same pattern. Only proteins where at least 80 % of all identified peptides had a p-value below 0.01 when compared by ANOVA with Bonferroni correction (A vs. E, C vs. E, B vs. F and D vs. F) were considered discriminatory. The peak areas for all discriminatory peptides belonging to the same protein were summed within the sample groups and an average was calculated. A protein with a difference of at least a factor of two for proteins that increased in the diseased groups (E and F) was considered as significantly discriminatory. For decreasing proteins a factor of ten times was considered as significant (since group E and F was injected at five times lower volume compared to group A-D, a protein decreasing ten times are
decreasing only two times at native relative comparison [raw samples without any normalization]).

2.6. **Protein identification (QTOF data)**

For protein identification one sample from group F and one sample from group A were analyzed by data-dependent MS/MS, the generated data files were converted to the .mgf format and exported to the PhenyX data base search tool (version 2.6, Geneva Bioinformatics, Geneva, Switzerland). MS/MS spectra were searched against the uniprot_sprot and uniprot_sprot_rev databases (version: 20090608, Taxonomy: *Rattus norvegicus*). The search parameters were set as follows: scoring model: ESI-QTOF (QTOF), parent charge: +2, +3, +4, peptide/AC score: ≥ 5, peptide length: ≥ 4, p-value < 0.001, enzyme: trypsin, missed cleavage: 2, parent tolerance: 10 ppm, amino acid modifications: carbamidomethylation of cysteine (fixed, all), oxidation of methionine (variable, <4), cleavage mode: normal, false positive rate upon searching the reverse database < 2%. Only proteins that were identified with at least two separate peptides were considered as true positives.

2.7. **Nano-LC-ESI-Orbitrap MS/MS**

Digested rat CSF samples were analyzed by LC-MS/MS using an Ultimate 3000 nano LC system (Dionex, Germering, Germany) online coupled to a hybrid linear ion trap / Orbitrap mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific, Bremen, Germany). Five microliter digest were loaded onto a C18 trap column (C18 PepMap, 300µm ID x 5mm, 5µm particle size, 100 Å pore size; Dionex, The Netherlands) and desalted for 10 minutes using a flow rate of 20 µL/min. The trap column was switched online to the analytical column (PepMap C18, 75 µm ID x 150 mm, 3 µm particle and 100 Å pore size; Dionex, The Netherlands) and peptides were eluted with the following binary gradient: 0% - 25% solvent B for 120 min and 25% - 50% solvent B for further 60 minutes, where solvent A consisted of 2% acetonitrile and 0.1% formic acid in ultra pure water and solvent B consisted of 80% acetonitrile and 0.08% formic acid in water. The column flow rate was set to 300 nL/min. For MS detection a data dependent acquisition method was used: a high resolution survey scan from 400 – 1800 m/z was performed in the Orbitrap (automatic gain control (AGC) 10⁶, resolution 30,000 at 400 m/z; lock mass set to 445.120025 m/z [protonated (Si(CH₃)₂O)₆])(28). Based on this survey scan the 5 most intense ions were consecutively isolated (AGC target set to 10⁴ ions) and fragmented by collision-activated dissociation (CAD) applying 35% normalized collision energy in the linear ion trap. Once a precursor had been selected, it was excluded for 3 minutes.
2.8. **Orbitrap data processing**

The raw data was preprocessed using the Progenesis LC-MS software package (version 2.5) (no normalization was applied). Peptides were identified and assigned to proteins by exporting features, for which MS/MS spectra were recorded, using in the Bioworks software package (version 3.2; Thermo Fisher Scientific, Germany, peak picking by Extract_msn, default settings). The resulting .mgf file was submitted to Mascot (version 2, Matrix Science, London, United Kingdom) for identification to interrogate the UniProt-database (version 57.0, rattus taxonomy: 7114 sequences). Only ions with charge states between +2 and +7 were considered and only proteins with at least two unique peptides (with a Mascot score > 25) assigned to them were accepted as true identifications. Modifications: carbamidomethylation of cysteine (+57.021 m/z) was set as fixed and oxidation of methionine (+15.996 m/z) as variable modification allowing a maximum of 2 missed cleavages. Mass tolerance for precursor ions was set to 10 ppm and for fragment ions at 0.5 Da. The cut-off for mass differences between the measured and the theoretical mass of the identified peptides was set at 2 ppm. The Mascot search results were imported back into the Progenesis software to link the identified peptides to the detected abundances of these peptides. Subsequently the data were exported in Excel format.

The peak matrix containing only the identified features was analyzed by one-way ANOVA with Bonferroni post hoc test (A vs. E, C vs. E, B vs. F and D vs. F) (SPSS). All identified proteins with 80 % of the identified peptides having a *p*-value below 0.01 in the statistical test were kept for further analysis. The sum of all peptides originating from the same protein was calculated for each sample and the average of this sum was calculated for each group of animals and compared between diseased groups and controls (A vs. E, C vs. E, B vs. F and D vs. F). If the average sum in the diseased groups was increased at least 10 times or decreased more than 2 times, the protein was considered as significantly discriminatory (since the overall protein abundance was increased by five times in the diseased groups, we considered that proteins had to increase a factor of two more than this overall increase to be significantly discriminatory).

For further analysis of the Orbitrap data using a second data processing procedure a normalization (29) was applied in the Progenesis software (to compensate for the large TIC differences and detect differences between groups that were not caused by an overall increase in protein abundance). The resulting peak matrix was further processed in Microsoft Office Excel 2007. All non-identified peaks were filtered out and the peak matrix was imported into MatLab (version R2009a) for NSC and PCA analysis. The NSC analysis was done pair wise (A vs. E, C vs. E, B vs. F and D vs. F) for selection of discriminatory peaks between the diseased and control groups. PCA was applied to the NSC-selected features as well as to the complete peak matrix to visualize group separations. The NSC-
selected peaks were linked to the identified protein in the Excel peak matrix and listed for comparison of results between workflows and platforms.

3. Results

Trypsin-digested CSF samples from Lewis rats with induced EAE were compared to CSF from healthy as well as inflammatory control animals collected at day 10 or day 14 after injection of the disease-inducing agent (see Table 1 for details). The weight of the rats was monitored daily from the day of injection until euthanasia at day 10 or 14 (Figure 1). The average weight of animals belonging to the healthy (group B) or inflammatory control (group D) increased continuously from day 1 to day 14, while animals in the EAE group (group F) showed a marked decrease starting on day 11. The weight pattern of the rats from groups A, C and E was similar until day 10 (data not shown). The weight decrease in the EAE group (group F) was paralleled by an increase in neurological scores (Table 2). None of the other animals (group A-E) showed an increase in neurological score.

![Figure 1](image-url)

**Figure 1.** Average weight of rats in groups B, D and F (groups A, C and E showed no difference in the weight pattern between each other). A diverging weight pattern can be seen for group F from day 11 onwards. Rat groups; A: healthy control day 10, B: healthy control day 14, C: inflammatory control (CFA) day 10, D: inflammatory control (CFA) day 14, E: EAE day 10, F: EAE day 14.
Table 2. Neurological scores of the samples included in the analysis from animals in group F monitored at each day (Only group F showed increased neurological scores). 
Explanation; 0: no pathological signs, 1: paralysis of the tip of the tail, 2: no curling reflex/no strength at tail basis, 3: slightly impaired walking, 4: unstable walk, 5: half hind limb paralysis, 6: complete hind limb paralysis, 7: midriff paralysis, 8: half for-limb paralysis, 9: moribund, 10: death due to EAE.

```
<table>
<thead>
<tr>
<th>Sample nr.</th>
<th>Score day 9</th>
<th>Score day 10</th>
<th>Score day 11</th>
<th>Score day 12</th>
<th>Score day 13</th>
<th>Score day 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>76</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>77</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>79</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>82</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>84</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>85</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>86</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>87</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>88</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>89</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>90</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
```

Visual inspection of the LC-MS data revealed that the total protein concentration (measured as the total ion chromatogram or the UV-trace) was significantly higher (~5 times) for all animals in group F and for about half of the animals in group E. Differences were highly significant as shown by box and whisker plots of the area under curve of the total ion chromatograms (TIC) based on the chipLC-QTOF analyses (Figure 2A). In order to not overload the trap and separating columns on the microfluidics device and to make optimal use of the dynamic range of the LC-MS system, samples from groups E and F were injected a second time at a 5 times lower volume (0.2 µL) compared to group A-D (1 µL) giving comparable areas for the TIC (Figure 2B). Interestingly, variation was very high in group E (EAE animals at the general onset of disease, day 10). This is most probably due to variable inception of disease depending on the individual animal. Variability at day 14, when disease symptoms have reached their maximum, was again rather small (Figure 2A, 2B).

In the LC-Orbitrap system all samples were injected at the same volume (5 µL), with no problems of overloading, because a trap column with larger volume capacity was used for acquisition with Orbitrap MS. In order to, in the PCA and NSC analyses, rule out detections of differences caused by an overall increase of
Proteome profiling in a rat EAE model

protein abundance in the group E and F analyzed by the Orbitrap, the data was normalized at data level (29). For peak area comparisons the raw data without normalization was used. Combining the QTOF and Orbitrap methods we identified 233 proteins with at least 2 unique peptides each.

![Sample group](image)

**Figure 2.** A. Area under the curve of the Total Ion Chromatograms (TICs) from analysis of 0.5 μL trypsin-digested cerebrospinal fluid (CSF) from groups A-F. Univariate statistical analysis (Student’s t-test) revealed significant differences between the sample groups (A vs. E: p < 0.05; C vs. E: p < 0.05; B vs. F: p < 0.0001; D vs. F: p < 0.0001). B. TIC areas after taking the difference in TIC area into account by injecting 1 μL of trypsin-digested CSF for groups A-D and 0.2 μL for groups E and F (Univariate statistical analysis revealed no significant difference when the groups were compared with each other).

Since protein concentration (measured as Total Ion Chromatogram, TIC) in group E was highly variable, we compared this group as a whole but also divided it into a “high” TIC group (E<sub>high</sub>) (squares) and a “low” TIC group (E<sub>low</sub>) (stars) (Table 1). Figure 3A visualizes the PCA of the normalized Orbitrap data showing that groups E and F deviate from the control groups A-D even though the total ion chromatograms were normalized. Figure 3B show the corresponding PCA of raw QTOF data (normalized through injection of different volumes at sample level) confirming that groups E and F differ significantly from the control groups in their digested protein profile. In both the Orbitrap and the QTOF analysis group E<sub>high</sub> is more separated from group A-D than E<sub>low</sub>.

In order to select individual peptide peaks that show statistically significant differences in abundance between the different groups of animals, we applied the NSC classifier to the QTOF-MS and the Orbitrap-MS peak matrices using a double cross validation scheme (30, 31). For all the comparisons the NSC recorded the lowest classifications error rates for wide continuous ranges of 100
shrinkage values which means that there was a very clear separation of the groups. **Figures 4A-D** show principal component analyses of identified NSC-selected peaks for groups A vs. E (upper panels) and C vs. E (lower panels) based on the data from Orbitrap (left panels) and QTOF LC-MS analyses (right panels). The E group is divided in E<sub>high</sub> (squares) and E<sub>low</sub> (stars) (see **Table 1**). In the plots it is visible that the E<sub>high</sub> group is more separated from the control groups than the E<sub>low</sub> group, indicating differences in the protein profile between healthy controls and EAE group in part of the animals at day 10. This indication can be seen as an asymptomatic onset of the disease of some of the animals already at day 10. **Figures 5A-D** show principal component analyses of identified NSC-selected peaks for the comparison of groups B vs. F (upper panels) and D vs. F (lower panels) for the both Orbitrap- (left panels) and QTOF-derived data (right panels). The PCA plots of these comparisons show a clear to moderate tendency to group separation indicating that the trypsin-digested CSF protein profiles are different between healthy and inflammatory control groups and the EAE group at the height of disease symptoms (day 14).

**Figure 3.** **A.** PCA based on all identified peptide ions (6195 peaks) of groups A-D (●), group E<sub>low</sub> (●), E<sub>high</sub> (□) and group F (△) from the preprocessed, normalized Orbitrap LC-MS data. **B.** PCA based on the complete peak matrix (5336 peaks) of groups A-D (●), group E<sub>low</sub> (●), E<sub>high</sub> (□) and group F (△) from the in-house preprocessed QTOF LC-MS data.

For a protein to be considered as discriminatory we have used highly stringent criteria. Proteins identified with only one peptide were discarded from further analysis and at least 80% of the detected peptides belonging to one protein had to be significantly discriminatory ($p < 0.01$) by ANOVA comparison with Bonferroni correction between diseased animals and control (A vs. E, C vs. E, B vs.
Proteome profiling in a rat EAE model

F and D vs. F). Because of the large differences in total amount of protein in the EAE diseased animals (groups E and F) compared to the control groups, only proteins that exceeded the increase of total protein amount (five times) by at least twofold. To determine the fold increase the average of the sum of all peptides from a protein in one group was compared it to the identical protein in another group. For proteins that decreased in the diseased groups compared to control a factor of two was considered as discriminatory. The protein level in samples from group E and F was already present at a factor of five lower than in the raw, native samples for QTOF data (because of the normalized injection). This means that in the QTOF data a two fold increase and a tenfold decrease was considered as discriminatory. In the Orbitrap data that was not normalized at either sample or data level but instead compared at a native relative level the increase had to be a factor of ten in group E and F compared to control, while the decrease had to be twofold to be considered as significant.

**Figure 4.** PCA of NSC-selected peaks from the Orbitrap (A and C) and QTOF data (B and D). A. A (▼) vs. E: E_{low} (▼) E_{high} (□) Orbitrap data: PCA based on 3 identified NSC-selected peaks at shrinkage 8. B. A vs. E QTOF data: PCA based on 70 identified NSC-selected peaks.
peaks at shrinkage 4. C. C (▼) vs. E: $E_{\text{low}}$ (•) $E_{\text{high}}$ (□) Orbitrap data: PCA based on 1950 identified NSC-selected peaks at shrinkage 1. D. C vs. E QTOF data: PCA based on 17 identified NSC-selected peaks at shrinkage 7.

![Figure 5](https://example.com/figure5.png)

**Figure 5.** PCA of NSC-selected peaks from the Orbitrap (A and C) and QTOF data (B and D). A. B (▼) vs. F(•) Orbitrap data: PCA based on 96 identified NSC-selected peaks at shrinkage 4. B. B vs. F QTOF data: PCA based on 23 identified NSC-selected peaks at shrinkage 7. C. D (▼) vs. F (•) Orbitrap data: PCA based on 583 identified NSC-selected peaks at shrinkage 2. D. D vs. F QTOF data: PCA based on 31 identified NSC-selected peaks at shrinkage 7.

Figure 6 A-D shows box and whisker plots of the average sum of the peak areas of all detected peptides from four different proteins. The serum protein paraoxonase/arylesterase 1 (P55159) (average of 4 peptides) is clearly increased in the $E_{\text{high}}$ group (84 times increased at native comparison level [no normalization] in Orbitrap data) and decreases relative to group $E_{\text{high}}$ to an 8 fold increase in group F relative to healthy and inflammatory control groups (see Figure 6A and Table 3). This indicates that animals react differently to induction of EAE in time.
Proteome profiling in a rat EAE model

and that the onset of disease, as detected by proteomics, is not synchronized exactly on Day 10. It is noteworthy that none of the animals in group E had neurological symptoms (the neurological score was zero). The fact that this protein was increasing strongly at the onset of the disease while decreasing to a lower level at the climax of the disease course, indicates that this protein is elevated as a result of an acute early disease process or the start of disruption of the blood brain barrier. Lysozyme C1 (P00697) (average of 4 peptides) was dramatically increased in group F (11 times increased in the QTOF data [normalized at sample level]) while it remained almost at the level of the control groups in group E (see Figure 6B, Table 3). This protein showed the same behavior in the Orbitrap data as well but did not fulfill our criteria to be considered as significantly discriminatory. Lysozyme C1 here show a very disease specific pattern with a clear increase only at the apex of the disease, this indicates that this protein might be involved in processes that are more related to CSF membrane disruption. Complement C3 (P01026) (figure based on QTOF data, normalized at sample level, average of 11 peptides) showed yet another pattern in that it was increased in group E<sub>high</sub> as well as the F group (3 times increased in both groups compared to control in QTOF analysis, see Figure 6C, Table 3). This protein was 14 times increased in the F vs. B comparison and 26 times increased in the E<sub>high</sub> vs. A comparison in the Orbitrap analysis (non-normalized). The pattern of an increased level already at the onset of the disease and with a maintained high level at the apex makes it an early onset candidate marker as well as a marker for one of the pathological processes in the developing disease. Vitamin D binding protein (P04276) (figure based on QTOF data, normalized at sample level, average of 8 peptides) also increased significantly already in the E<sub>high</sub> group (3 times increased in QTOF analysis and 21 times increased in Orbitrap analysis), in group F the increase was also significant but lower (2 times increased in QTOF analysis and 7 times increased in Orbitrap analysis) compared to control (Figure 6D, Table 3). Also this protein has a stronger effect in the early stage of the disease making it a potential early candidate marker.

There were a few proteins, apolipoprotein E (P02650), angiotensinogen (P01015) and prostaglandin D synthase (PTGDS, P22057), of which the concentration did not follow the TIC. We assume that these proteins in this case are genuine CSF-derived proteins, which are not affected by a change in permeability of the blood-brain barrier (BBB). Figure 7 shows the sum of the peak areas of 4 peptides originating from PTGDS detected in the QTOF analysis. In Figure 7A the raw data from the QTOF analysis (TIC normalized at sample level, five times less injected of group E and F compared to group A-D) showing a significantly decreasing pattern of this protein in the diseased animals compared to control (<i>p</i> < 0.01) the decrease of the diseased groups compared to the control groups were however less than 4 times. Figure 7B is based on a QTOF analysis
were all samples were injected at the same volume (native relative comparison), here the protein is stable in all groups, indicating clearly that differences between groups are non-significant ($p > 0.05$).

**Figure 6.** Boxplots of the average sum of all peptides from 4 discriminatory proteins. In the figures group E is divided in two groups ($E_{\text{high}}$ and $E_{\text{low}}$) based on their TIC area; $E_{\text{high}}$: high TIC, $E_{\text{low}}$: low TIC. All values originates from raw data from the analysis (no normalization at data level). A. Serum paraoxonase/arylesterase 1 (average of 4 peptides), this protein was detected by the Orbitrap analysis only. B. Lysozyme C1 (average of 4 peptides) the protein was only significantly discriminatory in the QTOF analysis but showed the same behavior in the Orbitrap analysis. C. Complement C3 (average of 10 peptides), significantly discriminatory in both Orbitrap and QTOF analysis. D. Vitamin D binding protein (average of 8 peptides) also this protein was significantly discriminatory in both QTOF and Orbitrap analysis.

In Table 3 all discriminatory proteins that increased more than 2 times (10 times in non-normalized data) in the diseased groups $E_{\text{high}}$ and/or $F$ compared to groups A and B respectively from both analysis platforms are listed with the
relative change in abundance between groups F and B and $E_{\text{high}}$ and A, respectively. No protein showed a significantly decrease in abundance in the disease groups (E and F). The upper part of the table shows 14 proteins that were found to be discriminatory in both data sets, the middle part shows 23 proteins that were only detected by the Orbitrap, and the lower part shows 4 proteins that were discriminatory only in the QTOF data set, although these four proteins were also detected in the Orbitrap analysis. While Orbitrap data showed the same tendency as the QTOF data, the results did not fulfill our stringent criteria and the proteins were therefore not considered as significantly different (see Table S.1., supplementary material for the identification information for each peptide and protein detected by the QTOF method given in Table 3). In summary, based on our data we can discriminate 41 proteins that increase in EAE compared to healthy and inflammatory control groups, of which 28 are not described as discriminatory in EAE before.

**Figure 7.** Boxplots of the average sum of 4 peptides originating from prostaglandin D synthase detected in the QTOF analysis. Group E is divided in two groups ($E_{\text{high}}$ and $E_{\text{low}}$) based on their TIC area; $E_{\text{high}}$: high TIC, $E_{\text{low}}$: low TIC. A. Raw values from the QTOF LC-MS analysis (normalization of TIC’s at injection level; samples from group E and F injected at 5 times lower volume than group A-D). B. Raw values from QTOF LC-MS analysis where all samples were injected at the same volume (0.5 µL each).

4. Discussion

In human subjects, the large biological variation complicates the detection of biomarkers, especially when using a screening approach of the whole proteome. Working with an animal model makes it possible to collect both tissue samples and CSF; a better control of the disease course in the animal model enables...
collection of samples at a specific stage of the disease, which makes it possible to get a strong response also when screening complex samples. Candidate markers detected in the EAE model can then be evaluated in human subjects in a targeted fashion.

Here we have described an interlaboratory proteomic biomarker study based on CSF from an acute EAE model mimicking certain aspects of MScl. The goal was to find proteins connected to the disease. We have analyzed CSF samples using two different mass spectrometry based proteomics platforms at two independent laboratories. We have also applied two separate data processing workflows on both of the data sets.

Our study shows that induction of EAE in rats leads to a significant on average five times increase in the total protein concentration in cerebrospinal fluid (CSF). Despite this general change in protein concentration, we identified 41 proteins that increased at least by a further factor of 2 clearly discriminating EAE animals from control animals. The fact that we have also found proteins that do not change in concentration between the groups of animals demonstrates that the observed increase in protein concentration is not simply due to a decreased volume of CSF or a generic overall response of all proteins in groups E and F, but that the identified proteins increase in response is due to the induction of EAE in a disease controlled active regulated way.

There was a longitudinal component in our study, as the animals in groups E and F were both injected with MBP + CFA but in group E, CSF was collected on day 10, at the onset of EAE while CSF from group F was collected at the climax of disease symptoms on day 14. Figure 1 and Table 2 show that the animals in group F started to lose weight and showed elevated neurological scores from day 11 onwards, some of the animals in group F started off with a higher neurological score at day 11 while some of the animals still had no score at this day. The separation of the animals in group E might be a sign of which animals will develop scores earlier (E_high) than other animals (E_low), all rats in group E were asymptomatic at the day of sacrifice (day 10) despite the fact that the E_high group had clearly elevated levels of proteins related to the animals in group F at the apex of disease (see Figure 6). This indicates that protein patterns may be used as early indicators of disease onset prior to the appearance of clinical symptoms. The time difference in response to induction of EAE in an inbred strain of rats held under the same conditions must be taken into account when evaluating the effect of pharmacological interventions with newly developed drug candidates at an early stage of onset of the disease.

This study was interlaboratory in the means of the same sample set being analyzed at two separate university sites using different analysis platforms. We also combined two different data processing pipelines on the two data sets each in order to cover as much of the variation as possible. The results from the different
data processing strategies were partly overlapping showing reproducibility between the analysis strategies but the combination of the techniques also yielded an increased coverage of discriminatory features.

Taking both analyses together, we identified 233 proteins from which 41 proteins were increased in group F versus the healthy control group B and the inflammatory control group D (all collected on day 14) and group E_{high} versus the healthy control group C and the inflammatory control group A (all collected on day 10). Out of these 41 proteins, 28 have not been described as discriminatory in other EAE studies (see highlighted proteins in Table 3). Fifteen proteins were shown to be discriminatory on both LC-MS/MS platforms while another 23 proteins were only observed on the Orbitrap. Four proteins fulfilled our criteria for discriminatory proteins based on the QTOF data but not based on the Orbitrap data, although they were identified on this instrument. They showed, however, the same tendency of increasing concentration on the Orbitrap as well. This shows that analysis of trypsin-digested CSF samples from an acute rat EAE model resulted in a considerable overlap with respect to discriminatory proteins in an interlaboratory proteomics analysis.

Table 3. Discriminatory proteins listed with name, accession number, number of peptides, the quotient of F/B and E_{high}/A (non-normalized peak areas). The values are based on the average of the sum of all peptides identified as originating from that protein. The first part of the table consists of proteins that were discriminatory in both the Orbitrap and QTOF data, while the mid part contains proteins that were detected as discriminator by the Orbitrap analysis only and the lower part contains proteins that were discriminatory in the QTOF data only. The orbitrap data originates from an equal sample injection of all samples from all groups (no-normalization) and the data from the QTOF analysis originates from a normalization of the TIC at sample level (group E and F injected at five times lower volume than group A-D). Proteins that have not been described as discriminatory in the EAE model before are marked in bold. All proteins are present in plasma as well.
<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession</th>
<th>MCO</th>
<th>CCO</th>
<th>MSCO</th>
<th>SCCO</th>
<th>PCCO</th>
<th>ICCO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha-1-macroglobulin</td>
<td>Q63041</td>
<td>63</td>
<td>6</td>
<td>22</td>
<td>8</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Ceruloplasmin</td>
<td>P13635</td>
<td>37</td>
<td>9</td>
<td>19</td>
<td>7</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Complement C3</td>
<td>P01026</td>
<td>94</td>
<td>14</td>
<td>26</td>
<td>10</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Fetiin-B</td>
<td>Q9QX79</td>
<td>19</td>
<td>8</td>
<td>15</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Haptoglobin</td>
<td>P06866</td>
<td>21</td>
<td>13</td>
<td>35</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Hemopexin</td>
<td>P20059</td>
<td>28</td>
<td>8</td>
<td>21</td>
<td>12</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Ig gamma-2A chain C region</td>
<td>P20760</td>
<td>9</td>
<td>13</td>
<td>17</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Murinoglobulin-1</td>
<td>Q03626</td>
<td>54</td>
<td>14</td>
<td>28</td>
<td>10</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Murinoglobulin-2</td>
<td>Q6IE52</td>
<td>6</td>
<td>17</td>
<td>44</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>T-kininogen 1</td>
<td>P01048</td>
<td>19</td>
<td>21</td>
<td>34</td>
<td>6</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>T-kininogen 2</td>
<td>P08932</td>
<td>18</td>
<td>30</td>
<td>38</td>
<td>7</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Vitamin D-binding protein</td>
<td>P04276</td>
<td>24</td>
<td>7</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

### Discriminatory proteins Orbitrap only

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession</th>
<th>MCO</th>
<th>CCO</th>
<th>MSCO</th>
<th>SCCO</th>
<th>PCCO</th>
<th>ICCO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha-1-antiproteinase</td>
<td>P17475</td>
<td>22</td>
<td>7</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apolipoprotein A-I</td>
<td>P04639</td>
<td>11</td>
<td>6</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apolipoprotein M</td>
<td>P14630</td>
<td>2</td>
<td>12</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complement C1q subcomponent subunit A</td>
<td>P31720</td>
<td>5</td>
<td>6</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complement C4</td>
<td>P08649</td>
<td>61</td>
<td>8</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrinogen alpha chain</td>
<td>P06399</td>
<td>13</td>
<td>30</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrinogen beta chain</td>
<td>P14480</td>
<td>19</td>
<td>6</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrinogen gamma chain</td>
<td>P02680</td>
<td>18</td>
<td>1</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glutathione peroxidase 3</td>
<td>P23764</td>
<td>2</td>
<td>17</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ig gamma-1 chain C region</td>
<td>P20759</td>
<td>6</td>
<td>14</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ig gamma-2B chain C region</td>
<td>P20761</td>
<td>4</td>
<td>12</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ig gamma-2C chain C region</td>
<td>P20762</td>
<td>7</td>
<td>21</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ig kappa chain C region B allele</td>
<td>P01835</td>
<td>6</td>
<td>10</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ig lambda-2 chain C region</td>
<td>P20767</td>
<td>4</td>
<td>9</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insulin-like growth factor-binding protein complex acid labile chain</td>
<td>P35859</td>
<td>3</td>
<td>12</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter-alpha-trypsin inhibitor heavy chain H3</td>
<td>Q63416</td>
<td>19</td>
<td>11</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kininogen-1</td>
<td>P08934</td>
<td>7</td>
<td>8</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasminogen</td>
<td>Q01177</td>
<td>32</td>
<td>5</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Z-dependent protease inhibitor</td>
<td>Q62975</td>
<td>2</td>
<td>10</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serine protease inhibitor A3K</td>
<td>P05545</td>
<td>13</td>
<td>7</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serine protease inhibitor A3M</td>
<td>Q63556</td>
<td>9</td>
<td>13</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum paraoxonase/arylesterase 1</td>
<td>P55159</td>
<td>4</td>
<td>8</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zinc-alpha-2-glycoprotein</td>
<td>Q63678</td>
<td>5</td>
<td>13</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Discriminatory proteins QTOF only

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession</th>
<th>MCO</th>
<th>CCO</th>
<th>MSCO</th>
<th>SCCO</th>
<th>PCCO</th>
<th>ICCO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysozyme C1</td>
<td>P00697</td>
<td>4</td>
<td>11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Proteome profiling in a rat EAE model

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha 1 acid glycoprotein</td>
<td>P02764</td>
<td>8</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Serine protease inhibitor A3N</td>
<td>P09006</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Alpha-2-HS-glycoprotein (Fetuin A)</td>
<td>P24090</td>
<td>6</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Among the 41 proteins that discriminated diseased from control groups, all are well-known, highly, moderate and low abundant plasma proteins which indicates that elevation may be partially the result of a change in the function of the blood-brain barrier (BBB). Since most of these proteins are also present in CSF under normal circumstances, discrimination between an increased intrathecal production and an increased infiltration, or both is not possible in this case.

Several of the discriminatory proteins belong to the class I acute phase proteins; e.g. alpha 1 inhibitor 3, fibrinogen, plasminogen, ceruloplasmin, hemopexin, haptoglobin, alpha 2 macroglobulin, alpha 1 antiproteinase, alpha 1 acid glycoprotein and the components of the complement system. These proteins are activated during inflammation (32-34), and up regulation of these proteins reflects the inflammatory processes that are ongoing in the EAE model as well as in MScl. Only a few of these proteins (haptoglobin, hemopexin and alpha 1 acid glycoprotein) were discriminatory in both the QTOF and the Orbitrap analysis when the inflammatory controls (C and D) where compared to the healthy controls (A and B). This shows that that the elevation of most of the proteins is not just due to a general inflammatory response to complete Freund’s adjuvant. Complement C3, involved in the activation of the complement system has been connected to both EAE and MScl in previous studies (20, 35-37), complement C1 has been discovered as discriminatory in the sense of PTM polymorphisms between EAE rats and controls (10). In this study Complement C3 were significantly increased in CSF from EAE rats in both the Orbitrap analysis and the QTOF analysis. Complement C1 was detected in the Orbitrap analysis only and was significantly increased in group E_high compared to group A. Alpha 1 inhibitor 3 is a protease inhibitor that was also found increased in the EAE model before (20) supporting our results. T-kininogen is another protein that is involved in inflammatory cascades and an important mediator of the inflammatory response (38). Earlier published data have shown that an elevated level of the kinin B1 receptor correlates with increased disease severity in MScl patients (39). Hemopexin and haptoglobin are heme-binding proteins, which have the ability to effectively inhibit the toxic effect of free heme by forming complexes (40), Heme is an iron containing porphyrin that exerts toxic effects on cells, DNA and organelles through the formation of reactive oxygen species (ROS). ROS are involved in the pathogenesis of several vascular diseases (41) and there are indications that ROS are involved in neurodegeneration and BBB disruption, which contribute to the
formation and persistence of lesions in MScl (40, 42, 43). Oxidative stress has also been connected to neurodegeneration and apoptosis in the EAE model. We found both hemopexin and haptoglobin at elevated levels in CSF of EAE compared to control. Hemopexin was also detected at elevated levels in spinal cord of EAE rats (20) and have been connected to MScl in earlier studies (44, 45). This connection between heme toxicity and MScl/EAE indicates that the increased levels of heme-binding proteins in our study could be related to the pathogenesis of EAE and ultimately MScl. Ceruloplamin is another protein involved in iron metabolism, this protein was also elevated in EAE compared to control in this study and the same behavior was observed in two earlier EAE studies (13, 20), also in MScl this protein has been found at elevated levels (46).

It is interesting to note that lysozyme C1 is strongly increased at the full-scale EAE (group F) (see Figure 6B). The primary role of lysozymes is to offer protection against bacterial infections through specific cleavage of the bacterial peptidoglycan cell wall (47). Previous studies revealed increased lysozyme levels in patients with MScl and other neurological disorders (48, 49) as well as in rheumatoid arthritis, another autoimmune disease (50). In EAE lysozyme C2 has been found at increased levels in the spinal cord of EAE rats compared to control (20). An increase in lysozyme C1 may be due to the presence of mycobacteria in complete Freund’s adjuvant but it is again noteworthy that the increase in CSF is not observed in the inflammatory control groups.

CSF of MScl patients is normally screened for an increase in immunoglobulin G (IgG) as well as for the presence of IgG bands upon isoelectric focusing (2). In agreement with this, we found that a number of immunoglobulins were elevated in CSF of EAE rats relative to inflammatory controls (see Table 3) IgG’s have also been detected at elevated levels in EAE compared to control in two previous studies (8, 20). A recent study into sequencing the complementarity determining regions (CDRs) of immunoglobulins opens new possibilities to study the antibody repertoire in more detail with ramifications for MScl and other autoimmune diseases (51, 52).

Another interesting protein that was found at elevated level is vitamin D binding protein. This protein has not been described as discriminatory in other EAE studies but elevated levels of this protein have previously been shown in CSF from patients with secondary progressive MScl compared to controls and patients with relapsing-remitting MScl (53) as well as in blood samples from pediatric MScl patients (45). Vitamin D is involved in the calcium homeostasis (54) and deficiency of this vitamin is one of the theories that is advanced in conjunction with the development of MScl (55); there are evidences that connect a reduced vitamin D level to MScl (56, 57). Another protein involved in the calcium homeostasis is the alpha 2 HS glycoprotein (also known as fetuin A) (58). This protein was found
elevated in EAE both in this study and a previous one (20) and has been found both at elevated and reduced levels connected to MScl (53, 59).

A recent study from our laboratories shows that some proteins in CSF show a very high biological variability, for example, haptoglobin, indicating that they are not well suited as disease-related biomarkers. The vitamin D binding protein on the contrary show low biological variation between healthy human subjects and therefore serves as a better biomarker candidate (60).

This interlaboratory study of CSF from an animal model of certain aspects of MScl, the rat EAE model, resulted in identification of a set of 41 proteins that discriminate EAE rats from control groups. To our knowledge, this is the first study of its kind focusing on rat CSF from the EAE model.

Part of the proteins detected in this study has been connected to EAE in previous studies supporting our data. However, our study revealed 28 additional proteins to be discriminatory between EAE and control (Table 3, highlighted). The fact that many of the proteins have also been connected to MScl before shows that it is possible to translate findings from the animal model to the human situation and that our findings may have a clinical interest in form of early disease biomarkers. In summary, we show a range of proteins that merit further study in the context of MScl (early diagnosis, prognosis, response to therapy). We are currently pursuing work focusing on analyzing these proteins in CSF in response to drug treatment in the EAE rat model.

Acknowledgements

This study was performed within the framework of Top Institute Pharma project D4-102. The work was also supported by the project BioRange 2.2.3 from the Netherlands Proteomics, The Netherlands Bioinformatics Center and Dutch Multiple Sclerosis Foundation. We thank Balaji Srinivasan (University of Groningen) for preparation of the CSF samples.
Reference List


Proteome profiling in a rat EAE model


29. (2010) In:


