Discovery and verification of multiple sclerosis candidate biomarkers using iTRAQ and SID-SRM

Ann C. Kroksveen [1,2,3], Elise Aasebø [1,2], Heidrun Vethe [1], Vincent Van Pesch [4], Diego Franciotta [5], Charlotte Teunissen [6], Rune J. Ulvik [3,7], Christian Vedeler [8,9], Kjell-Morten Myhr [8,9,10], Harald Barsnes [1,11], and Frode S. Berven [1,10*]*

1 Proteomics Unit (PROBE), Department of Biomedicine, University of Bergen, Bergen, Norway.

2 Department of Clinical Medicine, University of Bergen, Norway.

3 Institute of Medicine, University of Bergen, Bergen, Norway.

4 Neurology Department UCL, Université Catholique de Louvain, Brussels, Belgium.

5 Laboratory of Neuroimmunology, IRCCS, "C. Mondino" Neurological Institute, Pavia, Italy.

6 Department of Clinical Chemistry, Neurological Laboratory and Biobank, VU University Medical Center, Amsterdam, The Netherlands.

7 Laboratory of Clinical Biochemistry, Haukeland University Hospital, Bergen, Norway.

8 Department of Neurology, Haukeland University Hospital, Bergen, Norway.

9 The KG Jebsen Centre for MS-research, Department of Clinical Medicine, University of Bergen, Bergen, Norway.

10 The Norwegian Multiple Sclerosis Competence Centre, Haukeland University Hospital, Bergen, Norway.

11 Computational Biology Unit, Uni Computing, University of Bergen, Bergen Norway.

* Corresponding authors:

Ann Cathrine Kroksveen, Proteomics Unit (PROBE), Department of Biomedicine, University of Bergen, Bergen, Norway, phone: +4755586378, Email: ann.kroksveen@med.uib.no

Frode S. Berven, Proteomics Unit (PROBE), Department of Biomedicine, University of Bergen, Bergen, Norway, phone: +4799163866, Email: frode.berven@biomed.uib.no

ABSTRACT

In the present study, we aimed to discover and verify proteins with differential abundance in cerebrospinal fluid (CSF) from patients with early multiple sclerosis compared to controls. iTRAQ and Orbitrap MS was used to compare the CSF proteome of patients with clinically isolated syndrome (CIS) (n = 5), patients with relapsing-remitting multiple sclerosis that had CIS at the time of lumbar puncture (n = 5), and controls with other inflammatory neurological disease (n = 5). Of more than 1200 identified proteins, five were selected as biomarker candidates. Selected reaction monitoring (SRM) was used to verify the biomarker candidates in a larger patient and control cohort (n = 132). We also included proteins reported as differentially abundant in multiple sclerosis in the literature for SRM verification. We found differential abundance of 11 proteins after verification, of which the five proteins alpha-1-antichymotrypsin, contactin-1, apolipoprotein D, clusterin, and kallikrein-6 show potential as diagnostic markers for multiple sclerosis. This study forms the basis for further biomarker verification studies in even larger sample cohorts, to determine if these proteins have clinical relevance as biomarkers for multiple sclerosis.

KEYWORDS

Multiple sclerosis, cerebrospinal fluid, diagnostic biomarker, quantitative mass spectrometry, selected reaction monitoring.

INTRODUCTION

Multiple sclerosis is a chronic inflammatory demyelinating disease of the CNS, characterized by recurrent attacks of neurological dysfunction. The clinical course is diverse and normally initiated by a single demyelinating event known as clinically isolated syndrome (CIS) [1]. Currently, diagnosis relies on magnetic resonance imaging (MRI) and cerebrospinal fluid (CSF) analysis together with disease history and clinical examination. Studies have shown that CIS patients with white matter lesions on MRI, and CSF oligoclonal bands have a higher risk of developing multiple sclerosis [2-8], but these abnormalities are only suggestive and not specific for the disease [2-4, 9]. Axonal damage and tissue destruction may be clinically silent, and neurological damage can therefore be present before the first symptoms occur [10-13]. Early treatment can slow down disease progression, prevent axonal damage, and reduce the frequency of relapses making early diagnosis important [11, 14-18].

Several biomarker candidates have been suggested for multiple sclerosis using proteomics (reviewed in [19]), but the gap between the number of proposed and verified biomarker candidates is large. Selected reaction monitoring (SRM) is a targeted proteomic technology that eventually may cover this gap. SRM allows for multiplexing and enables relative and absolute quantification of predefined proteins and peptides [20]. This technique uses the unique feature of two-step mass filtering on triple quadruopole mass spectrometers (QQQ), giving increased sensitivity, specificity, and dynamic range in the MS analysis.

In this study, we first aimed to detect CSF proteins that could have potential as early diagnostic biomarkers for multiple sclerosis. We did an in-depth proteomics discovery experiment comparing patients that had stable CIS over a mean follow-up period of 32 months, patients that had CIS at the time of lumbar puncture but later developed relapsing-remitting multiple sclerosis (RRMS) (CIS-MS), and controls with other inflammatory neurological disease (OIND). Secondly, we aimed to verify the findings with SRM in crude CSF in a larger cohort of patients (n = 67) and controls (n = 65). This larger sample cohort also included established RRMS patients and controls with other neurological diseases (OND). Furthermore, we included proteins that have been reported as differentially abundant in multiple sclerosis in the literature for SRM verification. Previous verification of some of these proteins was limited, and others were interesting to include in an additional verification study including other patient and control groups. To our knowledge, this is the first report of a larger SRM verification study of biomarkers for multiple sclerosis.

EXPERIMENTAL PROCEDURES

Patients and controls for sample collection

CSF samples were obtained from three different locations: (a) Department of Neurology, Haukeland University Hospital, Bergen, Norway, (b) Neurology Department UCL, Université Catholique de Louvain, Brussels, Belgium, and (c) Laboratory of Neuroimmunology, IRCCS, "C. Mondino" Neurological Institute, Pavia, Italy. All are members of the BioMS*eu* network for biomarkers in multiple sclerosis (www.bioms.eu). The CSF samples were collected according to the recommended consensus protocol for CSF collection and biobanking [21]. The study was approved by the institutional ethical committees, and written informed consent was obtained from all patients and controls.

Patients and controls included six categories: (i) CIS-MS; i.e. patients with RRMS that had CIS at the time of lumbar puncture, (ii) CIS; i.e. patients with CIS that did not develop multiple sclerosis in the follow-up period, (iii) patients with established RRMS, (iv) controls with OIND, (v) controls with OND, and (vi) spinal anesthesia subjects (SAS), i.e. controls without any neurological symptoms that underwent surgical procedures for lower extremity conditions. Supplementary Table 1 lists all the patients and controls included in this study, including diagnosis for the OIND and OND controls. The individual patient and control information is summarized in Table 1. The iTRAQ discovery study included CIS-MS (n = 5), CIS (n = 5), OIND (n = 5), and a pooled reference sample with CSF from SAS controls (n = 12). The mean follow-up period was 13.2 and 32.4 months for the CIS-MS and CIS patients, respectively (Table 1A). The SRM verification included CIS-MS (n = 16), CIS (n = 15), RRMS (n = 36), OIND (n = 33), and OND (n = 32). The mean follow-up period was 8.9 and 25.1 months for the CIS-MS and CIS patients, respectively (Table 1B).

Chemicals

N-octyl-β-D-glycopyranoside (NOG) was from Anatrace. Iodacetamide, urea, calcium hydrochloride, methylamine, and DTT were from Sigma. Tris(hydroximetyl)aminometan was from Merck. Trypsin was from Promega. Ammonium formate was from Fluka. Water, ACN, and formic acid (FA) was MS grade.

Sample preparation and abundant protein depletion prior to iTRAQ labeling

The CSF protein concentration was measured by Qubit™ fluorometer (Invitrogen, Carlsbad, CA). 200 µg CSF from each patient and control was protein depleted using a human Multiple Affinity Removal System (MARS Hu-14) 4.6 mm x 50 mm LC column (Agilent Technologies). This column deplete CSF of albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin, fibrinogen, alpha-2-macroglobulin, alpha-1-acid glycoprotein, IgM, apolipoprotein AI, apolipoprotein AII, complement C3, and transthyretin. The MARS column was connected to an Agilent 1100 series LC system (Agilent Technologies, Palo Alto, CA). For the first 16 minutes, 100% MARS Buffer A with a flow of 0.125 mL/min was used and the flow-through with protein depleted CSF was collected between 4.5-15 minutes. The bound fraction was collected with 100% MARS Buffer B from 16-21 minutes with a flow of 1 mL/min. MARS Buffer A was used for regeneration of the column before the next sample was injected. The protein depleted CSF was concentrated using 3 kDa ultracentrifugation filters (Amicon Ultra-4, Millipore, Bedford, MA), which were pre-rinsed with 0.1% NOG. Each sample was freeze dried prior to protein digestion and iTRAQ labeling.

Protein digestion and iTRAQ labeling

The entire amount of protein in each depleted sample was reduced, alkylated, digested with trypsin, and iTRAQ labeled according to the manufacturer's protocol using the reagents provided (Applied Biosystems, Foster City, CA). The reduced and S-methylmethanethiosulfonate (MMTS) treated proteins were digested to peptides over night at 37 °C using 2.5 µg trypsin as protease. The peptides were iTRAQ labeled (4-plex) and combined.

Strong cation exchange chromatography and Oasis® HLB µElution cleanup

The combined iTRAQ labeled peptide mixtures were separated by strong cation exchange (SCX) chromatography on an Agilent 1100 series LC system (Agilent, Palo Alto, CA) using a BioBasic SCX column (250 mm x 2.1 mm, 5 µm) (Thermo Scientific). The duration of the SCX fractionation was 120 minutes with the following buffers: SCX loading buffer (25% ACN, pH 3.0) and SCX elution buffer (250 mM ammonium formate / 25% ACN, pH 3.0). The gradient was as follow: Hold at 1% SCX elution buffer for loading and washing from 1-10 minutes. From 10-64 minutes 1-61% SCX elution buffer, from 64-84 minutes 61-100% SCX elution buffer. From 84-88 minutes hold at 100% SCX elution buffer, 88-90 minutes

then ramp from 100-1% SCX elution buffer and 1% SCX elution buffer was kept constant from 90-120 minutes. The flow was 0.2 mL/min. A total of 68 SCX fractions were collected with 1.3 minute intervals and the fractions were frozen and freeze dried. Each freeze dried SCX fraction was resuspended in Oasis Solution A (0.1% FA) and desalted using reverse phase Oasis® HLB µElution Plate 30 µm (Waters, Milford, MA). Briefly, the plates were washed once with Oasis Solution B (80% ACN / 0.1% FA) and thereafter washed twice with Oasis Solution A. The dissolved samples were added to the µElution plate and washed thrice using Oasis Solution A, before the peptides were eluted twice with Oasis Solution B.

Orbitrap mass spectrometry

Thirty (30) SCX fractions from each iTRAQ experiment were analyzed on an LTQ-Orbitrap XL (Thermo Scientific) coupled to an Agilent 1100 series LC system (Agilent, Palo Alto, CA). These were SCX fractions 31-60, collected between 39.10-78.10 minutes (SCX elution buffer 37-88%). SCX fractions 1-30 and 61-68 were excluded due to previous experience that these fractions contain low number of peptides. The samples were resuspended in 3% ACN / 5% FA and loaded (1 µg) onto a PicoFrit column (New Objective, Woburn, MA) with an inner diameter of 75 µm packed with 12-14 cm of ReproSil-Pur C_{18} 3 µm particles (Dr. Maisch GmbH). Analysis was of 120 minute total duration, using the following mobile phases: Mobile phase A (0.1% FA) and mobile phase B (0.1% FA / 90% ACN). The gradient used was as follows: Hold at 3% B at 0.6 µL/min from 0-13 minutes then reduce flow to 0.2 µL/min from 13-15 minutes. From 15-80 minutes 3-40% B, from 80-95 minutes 60-100% B. From 95-108 minutes hold at 100% B at a flow 0.6 µL/min, 108-110 minutes then ramp from 100-3% B, and re-equilibrate column at 3% B from 110-130 minutes.

The MS run time was 122 minutes with 11 scan events of which the first was a full FTMS scan over the 300-1600 *m/z* range with 60 000 resolution. The source voltage was 2 kV. For each of the five most intense ions both a FTMS higher-energy C-trap dissociation (HCD) and an ITMS CID event were triggered. For both events, a minimum signal intensity of 500 was required. The isolation width was set to 3 *m/z* and the activation time to 30 ms. The collision energy for the HCD event was 55% and the resolution was 7500. For CID the collision energy was 28% with an activation q of 0.25. Dynamic exclusion was enabled with a repeat count of 1 and exclusion duration of 25 s. Lock mass was not enabled. All 30 SCX fractions were analyzed by LC-MS/MS for each of the four iTRAQ experiments and searched using the Spectrum Mill software package v4.0 (Agilent Technologies, Santa Clara, CA) with the

following settings: Data extraction fixed modifications iTRAQ (N-term, K), MMTS. MS/MS spectral features MH+ 300-4000, scan time range 0-300 minutes. Sequence tag length: 1. Merge scans with same precursor *m/z* +/- 0 seconds +/- 1.4 *m/z*. Merge MS^2 and MS^3 from the same precursor. Ignore spectra with fragmentation mode HCD. Minimum signal to noise: 25 and find ^{12}C checked. MS/MS data was searched against the International Protein Index (IPI) human protein database version 3.45. The search parameters were: A maximum of two missed cleavages, precursor mass tolerance 0.035 Da, product mass tolerance 0.7 Da, MMTS treatment of cysteines and iTRAQ (N-term, K) as fixed modifications. Allowed variable modifications were oxidized methionines. Identities interpreted for individual spectra were automatically designated as valid by applying the scoring threshold criteria provided below to all spectra derived from a particular experiment in a two step process. First, protein mode was used which requires two or more matched peptides per protein and allowing a range of medium to excellent scores for each peptide. Second, peptide mode was applied to the remaining spectra allowing for excellent scoring peptides that are detected as the sole evidence for particular proteins. Protein mode thresholds: Protein score >20, peptide (score, Scored Percent Intensity, delta rank1-rank2) peptide charge +2: (> 8, > 65%, > 2) peptide charge +3: (> 9, > 65%, > 2) peptide charge +4: (> 9, > 70%, > 2) peptide charge +5: (> 6, > 90%, > 1). Peptide mode thresholds for all charge states: > 13, > 70, > 2, respectively. iTRAQ intensities from linked HCD spectra were integrated into Spectrum Mill after validation. The false discovery rate (FDR) was < 1.2%.

iTRAQ data analysis and biomarker candidate selection

In the iTRAQ discovery study, five CIS-MS, five CIS, and five controls (OIND) were grouped into five iTRAQ experiments (experiment 1-5). Each experiment included a pooled reference sample (12 SAS, labeled with iTRAQ 114). The three patient/control samples in each experiment were labeled as shown in Table 2. A Spectrum Mill (Agilent Technologies) protein summary report was generated for all five experiments (Supplementary Table 2) and the iTRAQ 114 channel was used as reference in all experiments. To normalize the data, we re-centered the protein ratios of each individual sample versus the reference sample by subtracting the median in each iTRAQ channel from each ratio so that the distribution was symmetrical around a $\log_2$ ratio of 0. This follows the assumption that most proteins are unchanged compared to the control group (the reference sample). For all comparisons, the individual ratios across the five samples for each group were averaged to generate the overall

group average (i.e. one ratio for CIS-MS, one ratio for CIS, and one ratio for OIND), finally resulting in a fold change difference per protein between the three groups.

For all comparisons, we removed various versions of keratin and applied the following filter settings: The biomarker candidate was required to (i) have quantitative information in all five experiments, (ii) be identified with two or more unique peptides in two or more experiments, (iii) show a *p*-value ≤ 0.05. After applying the filter settings we further removed biomarker candidates that did not differ in relative abundance from the reference sample (i.e. remove proteins with fold change < 1.20 compared to the reference sample). The number of quantified proteins, number of proteins that remained after applying the filter settings, the number of biomarker candidates, and the fold change criteria are listed in Table 3. This table also shows the standard deviation fold changes for each comparison.

Stable Isotope Dilution SRM-MS analysis

Signature peptides for the target proteins were defined and their uniqueness examined with a BLAST search against all human proteins in the NCBI database. Only unique peptide sequences were selected for the assays. Stable isotope–labeled internal standards (SISs) corresponding to the signature peptides were purchased in crude quality from Thermo Scientific and JPT technologies. Their C-terminal Lysine or Arginine was labeled with ^{13}C and ^{15}N. Each peptide was manually optimized by direct infusion on a Q-Trap 5500 (AB SCIEX), and the three most intense fragments were included in the final SRM-assays.

Each CSF sample (10 µg protein) was reduced, alkylated, and digested with trypsin according to the following protocol http://www.uib.no/filearchive/in-solution-proteindigestion.pdf. SISs were spiked into each sample prior to desalting with C18 StageTip (3M Empore™). The SIS spike-in was added in the approximately same amount as the corresponding endogenous peptide in the sample, using the response curves generated as described below.

A Q-Trap 5500 (AB SCIEX) coupled to a Dionex Ultimate NCS-3500RS LC system (Sunnyvale, CA) was used for LC SRM-MS analysis. Each sample was resuspended in 3% ACN / 5% FA, and 1 µL (corresponding to 1 µg CSF protein digest) was loaded onto the pre-column (Dionex, Acclaim PepMap Nano Trap column, C18, 75 µm i.d. x 2 cm, 3 µm) followed by separation on the analytical column (Dionex, Acclaim PepMap100 RSLCnano column, 75 µm x 15 cm, C18, 2 µm). The LC analysis was of 70 minute total duration with the following mobile phases: Mobile phase A (0.1% FA) and mobile phase B (0.1% FA / 90% ACN). The gradient used was as follows: 5-10% B from 0-3 minutes, 10-45% B from 3-45

minutes and 45-90% B from 45-45.5 minutes, and hold at 90% B from 45.5-51.5 minutes. From 51.5-55 minutes ramp from 90-5% B and re-equilibrate column at 5% B from 55-70 minutes. The flow rate was 250 nL/min. The MS run time was 68 minutes. The peptides were analyzed using scheduled SRM, with the target scan time set to 1 second and the detection window 240 seconds. Three transitions per peptide were monitored. The transition used for quantification, Q1 and Q3 values for the endogenous and SIS peptide, and the collision energy used for SRM analysis is in Table 4.

To monitor the linearity of the SRM-assays, a CSF test sample (i.e. pooled sample with CSF from multiple patients with various neurological conditions) was spiked with different concentrations of SIS. The amount of CSF loaded onto the column was kept constant (1 µg digested CSF proteins for each SIS dilution), and the concentration of SIS varied at least three orders of magnitude. We calculated the area ratio by dividing the area of SIS / area of endogenous CSF peptide since the concentration of endogenous peptides were kept constant. This area ratio was plotted against the different SIS concentrations for generation of the response curves, as shown in Supplementary Figure 1.

The SRM data was analyzed using MultiQuant 2.0.2 from AB SCIEX. SISs and endogenous CSF peptides were automatically integrated using Gaussian smooth width and peak splitting 2.0. The output from MultiQuant can be found in Supplementary Table 3. The data was manually evaluated and transitions with visible interference, wrong retention time, poor integration, and S/N < 10 were not included for further analysis. Furthermore, we used AuDIT [22] to detect interference in transitions by analyzing a CSF test sample that was included as a quality control in the SRM analysis. The output from AuDIT is shown in Supplementary Table 4. The most abundant transition free of interference was used for quantification. The fold change between the groups was calculated using the median SRM area ratio for each group (e.g. median area ratio OIND / median area ratio RRMS) and a Student's *t*-test was used to calculate the *p*-value. A p-value ≤ 0.05 was considered significant. The S/N was evaluated based on the general noise level for the SRM analysis. The noise level was below 100 and we therefore included transitions with intensity above 1000 in the quantitative analysis. This ensured that the S/N was at least 10 for all included transitions, which were defined as the LOQ.

SID-SRM quality control

Quality control samples were run to investigate the quality of the SID-SRM analysis. This quality control consisted of 21 CSF samples analyzed as processed replicates.

Using this quality control we could monitor the digestion, spike-in of SIS, and that the instrument performance were optimal. A quality control sample were run at the beginning and the end of the SRM analysis, and the 19 other quality control samples were run in between the 132 patient and control samples. The CV values for the quality controls are listed in Table 5.

RESULTS

Discovery of differentially abundant proteins using iTRAQ

In the iTRAQ discovery study (Figure 1A), we compared the CSF proteome of CIS patients (CIS-MS n = 5, CIS n = 5) with OIND controls (n = 5) to discover proteins with potential as early diagnostic markers for multiple sclerosis. A total of 1291 proteins were identified, of which 913 with two or more peptides. More than 600 proteins were quantified in each comparison (Table 3). We selected biomarker candidates with fold change ≥ 2 and *p*-value ≤ 0.05 when OIND were compared to CIS-MS and CIS (Table 6). No proteins had a *p*-value ≤ 0.05 in combination with a fold change ≥ 1.50, which were the selection criteria for biomarker candidates between CIS-MS and CIS. In total, five biomarker candidates were selected. Alpha-1-antichymotrypsin (ACT), leucine-rich alpha-2-glycoprotein (LRG), and lysozyme C were selected between CIS-MS and OIND, and ACT, LRG, complement component C5, and carboxypeptidase N subunit 2 were selected between CIS and OIND (Table 6). Bar charts with the individual iTRAQ ratios (log_2) for all biomarker candidates and patients and controls are shown in Supplementary Figure 2. All proteins with significant *p*-value from the discovery study are listed in Supplementary Table 5.

Verification of biomarker candidates using SRM

SRM-assays were built to verify the biomarker candidates from the discovery study, and we also built SRM-assays for more than 30 proteins that have been reported with differential abundance in multiple sclerosis in the literature. Differential abundance of the literature derived proteins has been reported in various group comparisons, and we wanted to examine if their differential abundance could be detected in a novel patient and control cohort. Furthermore, we wanted to investigate if these proteins could have potential as early diagnostic markers. The SRM workflow is shown in Figure 1B. Endogenous peptides were detected in crude CSF for 18 literature derived proteins and the two biomarker candidates ACT and LRG. The final SRM study therefore included 20 proteins. Table 4 lists the 20 protein biomarker candidates, and refers to a selection of studies where their differential abundance have been reported and their regulation in that study. Furthermore, the table lists their accession number, signature peptides, transition used for quantification, and the Q1 and Q3 masses for both the SIS and endogenous peptides. The MultiQuant output for the quantified proteins can be found in Supplementary Table 3.

Response curves were used to evaluate the linearity of the SRM-assays, and to determine the approximate spike-in amount of SIS peptide based on the amount of endogenous peptide present in a CSF test sample. All response curves are shown in Supplementary Figure 1. Figure 1B show the response curve for apolipoprotein D (apo D) and demonstrate that the SIS spike-in for the apo D peptide NILTSNNIDVK is close to 1:1 in relation to the endogenous level of this peptide.

The verification study included CSF from 67 CIS or multiple sclerosis patients (16 CIS-MS, 15 CIS, 36 RRMS) and 65 controls (33 OIND, 32 OND). Clinical and demographic data for the included patients and controls are listed in Supplementary Table 1 and summarized in Table 1. Five CIS-MS patients, four CIS patients, and five OIND controls in the verification cohort were also analyzed in the iTRAQ discovery study.

SRM analysis of alpha-1-antichymotrypsin and leucine-rich alpha-2-glycoprotein

ACT and LRG were significantly increased in OIND compared to CIS-MS and CIS in the iTRAQ discovery study (Table 6). Comparison of the iTRAQ and SRM quantitative results for ACT and LRG in the discovery cohort samples showed that the reported abundance for each patient and control correlated well between the two methods (Supplementary Figure 3). One CIS patient was not analyzed due to lack of material for SRM analysis. The *p*-value for ACT (peptide EQLSLLDR) was statistically significant (*p*-value 0.0398) between CIS-MS and OIND. No statistically significance was found for LRG (data not shown).

SRM analysis of the larger sample cohort confirmed the significant abundance difference of ACT (peptide EIGELYLPK) for CIS compared to OIND (Table 6), but the significance was just outside the significance threshold (*p*-value 0.0517) for CIS-MS compared to OIND (Supplementary Table 6). ACT was also significantly decreased in the RRMS compared to OIND (peptide EQLSLLDR) (Table 6). SRM analysis of LRG in the larger sample cohort did not show significant abundance difference between any of the compared groups (Supplementary Table 6). Box plots for ACT and LRG is shown in Figure 2.

SRM analysis of biomarker candidates selected from the literature

Significant abundance difference (p-value ≤ 0.05) was found with SRM for 10 of the 18 proteins with reported differential abundance in multiple sclerosis in the literature (Table 6). None of these proteins were significantly different between CIS-MS and CIS, indicating homology between these two groups.

In the comparison between CIS-MS and controls, contactin-1, apo D, clusterin, and kallikrein-6 showed significant abundance difference. Apo D was significantly less abundant in CIS-MS compared to OIND and OND, and in RRMS and CIS compared to OND (Table 6). Furthermore, the *p*-values were close to significant between RRMS and OIND (*p*-value 0.061) (Supplementary Table 6). Contactin-1, monitored by the two peptides FIPLIPIPER and DGEYVVEVR, was significantly decreased in CIS-MS compared to RRMS and OND (Table 6), and in CIS compared to OND. FIPLIPIPER was decreased in all afore-mentioned comparisons, whilst DGEYVVEVR was decreased in CIS-MS and CIS compared to OND. Kallikrein-6 and clusterin also showed potential as early diagnostic markers. Kallikrein-6 was significantly decreased in CIS-MS and CIS compared to OND, and clusterin was significantly increased in CIS-MS and CIS compared OIND (Table 6). Box-plots for apo D, contactin-1, clusterin, and kallikrein-6 are shown in Figure 3. Alpha-1-acid glycoprotein, alpha-2-macroglobulin, and complement C3 were significantly decreased in RRMS compared to OIND (Table 6). Furthermore, chromogranin A and cystatin-C showed decreased abundance in CIS compared to OND, and serotransferrin was increased in CIS compared to RRMS. Supplementary Table 6 contains a summary of the SRM results for all proteins and comparisons (*p*-value and fold change).

DISCUSSION

Patient selection

In this study, we aimed to detect protein biomarker candidates that could serve as early indicators of multiple sclerosis. We hypothesized that such proteins could be found by comparing the CSF proteome of patients with CIS-MS and CIS to controls with OIND. The CIS-MS patients had CIS at the time of lumbar puncture and developed RRMS in the follow-up period, and can therefore be regarded as early RRMS patients. However, the clinical and biological onset of the disease may not coincide, as the clinical symptoms will depend on the affected site in the CNS. The clinical onset is therefore unpredictable, and the disease may be more advanced than presumed from clinical symptoms in the CIS patients. Patients with established RRMS were included in the SRM verification study to examine if the early biomarker candidates also were significantly and differentially abundant at a later stage of the disease. The included CIS patients did not convert to clinically definite (CD) multiple sclerosis during follow-up. However, all but one CIS patient had CSF oligoclonal bands, and thus has an increased risk of developing CD multiple sclerosis [27]. A 30 year follow-up of patients with CIS (optic neuritis) reported that 49% of the patients with CSF oligoclonal bands would develop CD multiple sclerosis, but that most patients converted within three years [8]. The risk of developing CD multiple sclerosis is further increased if silent lesions are detected with MRI [28]. It is therefore likely that some of the CIS patients will develop CD multiple sclerosis after the end of the follow-up. Differentially abundant proteins in the CIS patients may however indicate slow conversion rate to CD multiple sclerosis or, potentially, no conversion at all.

Biomarker candidates from the iTRAQ discovery study

In the iTRAQ discovery study, five proteins with fold change ≥ 2.0 and p-value ≤ 0.05 were selected as biomarker candidates (Table 6). Out of these five proteins, ACT and lysozyme C have previously been associated with multiple sclerosis [24, 29-32]. In our discovery data, LRG and ACT were found to be decreased in CIS-MS and CIS compared to OIND, and may have potential as early diagnostic markers for multiple sclerosis. The fold change for the proteins that had significant p-value between CIS-MS compared to CIS was small (< 1.5) (Supplementary Table 5), suggesting that these groups are more similar in terms of their quantified CSF proteins.

Verification of biomarker candidates using SRM

For verification of biomarker candidates we successfully developed SRM-assays for 20 proteins, including ACT and LRG from the iTRAQ discovery study. SRM-assays were built for a larger number of proteins both from the discovery study and literature, but endogenous peptides were not detected in crude CSF for these proteins. In the discovery experiment, CSF was protein depleted and SCX fractionated prior to MS analysis. It is therefore likely that the remaining biomarker candidates are present below the LOD for SRM measurements in crude CSF. Protein depletion and limited SCX fractionation prior to SRM could eventually allow verification of lower abundance biomarker candidates [20, 33]. The aim of the present study was, however, to verify biomarker candidates in crude CSF and not verification of lower abundant proteins.

In the SRM verification, we found ACT increased in OIND compared to CIS-MS, CIS, RRMS, and OND (Table 6), with significant *p*-value for all comparisons except for CIS-MS compared to OIND (*p*-value 0.0517) (Supplementary Table 6). We therefore confirmed our findings from the iTRAQ discovery study for CIS compared to OIND in the larger sample cohort. As described in the results section, two peptides were monitored for ACT, and the result for these peptides was different in terms of significance between and within the compared groups (Supplementary Table 6). This may indicate that disease related processes influence the protein sequence differently. ACT is a serine protease inhibitor which is released with other acute-phase proteins by the liver during inflammation [34], but is also secreted from astrocytes [35, 36]. It has been linked to Alzheimer's disease pathogenesis, and found in macrophages in multiple sclerosis lesions [34-38]. ACT has been reported differentially abundant in patients with multiple sclerosis also by others [24, 29, 30]. Similar to Stoop and colleagues [30] we found ACT to be increased in OIND compared to CIS and OND.

Apo D was in our SRM study found to be differentially abundant in several group comparisons (Table 6). Its abundance was significantly decreased in CIS-MS compared to OIND and OND, and in RRMS and CIS compared to OND. Furthermore, the abundance difference between the RRMS and OIND group were close to significant (Supplementary Table 6). These data suggest that apo D may be of value as an early diagnostic marker for multiple sclerosis. Apo D is a member of the lipocalin protein family (reviewed in [39]) and has been found widely expressed throughout the brain, especially in glia and neurons in the CNS [40-43]. Several roles for apo D has been proposed, involving cholesterol transport, neuroprotection, myelination, and synaptogenesis [40, 44, 45]. It has been reported that apo D

accumulate during nerve regeneration, and may play a role in both normal and regenerating neuronal tissue [46]. A study by Reindl and colleagues suggested that apo D was synthesized intrathecally in multiple sclerosis patients [47]. Furthermore, a recent study reported that apo D was increased in primary-progressive multiple sclerosis (PPMS) and RRMS compared to healthy controls [23]. Since PPMS and healthy controls were not included for SRM verification, our study is not comparable to this study from Stoop and colleagues. However, the same group have also reported increased abundance of apo D in CIS compared to OIND [48], but we did not find this abundance difference using SRM (Supplementary Table 6). A possible explanation for the discrepancy, may be that that we monitored the abundance of apo D based on the peptide NILTSNNIDVK, and differential abundance for the RRMS/PPMS and CIS patients were reported for NPNLPPETVDSLK and VLNQELR, respectively.

Contactin-1 was significantly decreased in CIS-MS compared to RRMS and OND, and in CIS compared to OND, and may have potential as an early diagnostic marker for multiple sclerosis (Table 6). However, the two peptides used to monitor the abundance of contactin-1 did not show consistent significance between and within the different groups (Supplementary Table 6). Contactin-1 is a cell-adhesion glycoprotein distributed in the brain [49], and is highly involved in the regulation of oligodendrocyte survival, maturation, and myelination [50, 51]. In the literature, increased abundance of contactin-1 have been verified in SPMS compared to controls using multiplex immunoassays [24], and decreased levels have also been found in multiple sclerosis compared to OND [30].

Differential abundance for several biomarker candidates, including LRG from our discovery study, was not found using SRM. However, these proteins should not be rejected as biomarker candidates based on this study alone. Reporting protein abundances based on single peptides will only give a representative abundance of that particular sequence of the protein. Other peptides representing different regions of the same protein can give a different abundance ratio due to for example changes in post-translational modification, various isoform distribution, splice variants, truncation products, and proteolytic processing.

CONCLUSION

In the present study, we identified five proteins from an iTRAQ biomarker discovery study that had significant abundance difference between CIS-MS and OIND and/or CIS and OIND. Differential abundance was found for ACT also in the SRM verification, and this protein may be a potential marker for early diagnosis of multiple sclerosis. From SRM verification of literature candidates for multiple sclerosis, we found that alpha-1-acid glycoprotein, alpha-2-macroglobulin, and complement C3 was differentially abundant in RRMS compared to OIND. Furthermore, apo D and contactin-1 showed significant abundance difference in CIS, CIS-MS, and/or RRMS compared to controls, and are together with ACT especially interesting as early diagnostic biomarker candidates for multiple sclerosis. Further studies in larger patient and control cohorts, in addition to different multiple sclerosis subgroups, are however needed to confirm or reject the significance of these proteins as biomarker candidates. Our study serves as a good starting point for larger verification studies of these proteins, to further reveal their diagnostic or prognostic predictive value. However, our data indicate that a panel of biomarkers is most likely needed, since none of the verified biomarker candidates individually could differentiate between multiple sclerosis and controls.

ACKNOWLEDGEMENT

The study was supported by the National Program for Research in Functional Genomics (FUGE) funded by the Norwegian Research Council, Western Norway Regional Health Authority, the Leiv Eiriksson Mobility Program, the Meltzer Foundation, the Hejes foundation, and the Kristian Gerhard Jebsen Foundation.

Supporting material available:

The supporting material includes Supplementary Table 1, 2, 3, 4, 5, and 6, and Supplementary Figures 1, 2, and 3.

Supplementary Table 1 contains all data for the patients and controls included in the iTRAQ discovery and SRM verification. Patients and controls used in the iTRAQ study are in worksheet 1 and patients and controls used for SRM verification are in worksheet 2. The information is: Sample ID, condition at lumbar puncture (LP), diagnosis at follow-up, oligoclonal bands (OCB), gender, age at LP, protein concentration, and the Kurtzke Expanded Disability Status Scale (EDSS) score. The samples are originating from three different biobanks. The various biobanks are separated under the heading Biobank with a, b, and c in this Table.

Supplementary Table 2 contains the combined Spectrum Mill output from the five iTRAQ discovery experiments.

Supplementary Table 3 contains the MultiQuant output from the SRM analysis for all successfully quantified peptides. Each worksheet contains data from one peptide and all patients and controls. The transition used for quantification is in bold.

Supplementary Table 4 show the output from AuDIT where a CSF test sample was used to evaluate interference in transitions that was not visible in MultiQuant.

Supplementary Table 5 lists the proteins with p-value ≤ 0.05 in the iTRAQ discovery experiment. The CIS-MS vs. CIS comparison is in worksheet 1, the CIS-MS vs. OIND comparison is in worksheet 2, and the CIS vs. OIND comparison are in worksheet 3.

Supplementary Table 6 shows the SRM results for the different group comparisons (fold change and p-value) for all proteins included in the final SRM analysis.

Supplementary Figure 1 shows response curves for all peptides and proteins in the final SRM analysis. The amount of CSF endogenous peptide is kept constant while the concentration of SISs is varying. The area ratio (Area SIS / Area endo) is plotted against the concentration of SIS peptide.

Supplementary Figure 2 shows the individual iTRAQ reporter ion ratio ($\log_2$) for the CIS-MS and CIS patients and OIND controls for the five biomarker candidates from the iTRAQ study.

Supplementary Figure 3 shows a line plot of the original cohort samples where the correlation between the iTRAQ ratios (black) and the SRM verification area ratio (grey) are

compared for the two biomarker candidates (A) alpha-1-antichymotrypsin and (B) leucine rich alpha-2-glycoprotein.

REFERENCES

1. Miller, D., F. Barkhof, X. Montalban, A. Thompson, et al., *Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis.* Lancet Neurol, 2005. 4(5): p. 281-8.
2. Morrissey, S.P., D.H. Miller, B.E. Kendall, D.P. Kingsley, et al., *The significance of brain magnetic resonance imaging abnormalities at presentation with clinically isolated syndromes suggestive of multiple sclerosis. A 5-year follow-up study.* Brain, 1993. 116 (Pt 1): p. 135-46.
3. Jacobs, L.D., S.E. Kaba, C.M. Miller, R.L. Priore, et al., *Correlation of clinical, magnetic resonance imaging, and cerebrospinal fluid findings in optic neuritis.* Ann Neurol, 1997. 41(3): p. 392-8.
4. Beck, R.W., J.D. Trobe, P.S. Moke, R.L. Gal, et al., *High- and low-risk profiles for the development of multiple sclerosis within 10 years after optic neuritis: experience of the optic neuritis treatment trial.* Arch Ophthalmol, 2003. 121(7): p. 944-9.
5. Minneboo, A., F. Barkhof, C.H. Polman, B.M. Uitdehaag, et al., *Infratentorial lesions predict long-term disability in patients with initial findings suggestive of multiple sclerosis.* Arch Neurol, 2004. 61(2): p. 217-21.
6. Soderstrom, M., M. Lindqvist, J. Hillert, T.B. Kall, et al., *Optic neuritis: findings on MRI, CSF examination and HLA class II typing in 60 patients and results of a short-term follow-up.* J Neurol, 1994. 241(6): p. 391-7.
7. Tintore, M., A. Rovira, J. Rio, C. Nos, et al., *Is optic neuritis more benign than other first attacks in multiple sclerosis?* Ann Neurol, 2005. 57(2): p. 210-5.
8. Nilsson, P., E.M. Larsson, P. Maly-Sundgren, R. Perfekt, et al., *Predicting the outcome of optic neuritis: evaluation of risk factors after 30 years of follow-up.* J Neurol, 2005. 252(4): p. 396-402.
9. Paolino, E., E. Fainardi, P. Ruppi, M.R. Tola, et al., *A prospective study on the predictive value of CSF oligoclonal bands and MRI in acute isolated neurological syndromes for subsequent progression to multiple sclerosis.* J Neurol Neurosurg Psychiatry, 1996. 60(5): p. 572-5.
10. Bjartmar, C., J.R. Wujek, and B.D. Trapp, *Axonal loss in the pathology of MS: consequences for understanding the progressive phase of the disease.* J Neurol Sci, 2003. 206(2): p. 165-71.
11. Trapp, B.D., J. Peterson, R.M. Ransohoff, R. Rudick, et al., *Axonal transection in the lesions of multiple sclerosis.* New England Journal of Medicine, 1998. 338(5): p. 278-285.
12. Ferguson, B., M.K. Matyszak, M.M. Esiri, and V.H. Perry, *Axonal damage in acute multiple sclerosis lesions.* Brain, 1997. 120: p. 393-399.
13. Fu, L., P.M. Matthews, N. De Stefano, K.J. Worsley, et al., *Imaging axonal damage of normal-appearing white matter in multiple sclerosis.* Brain, 1998. 121 (Pt 1): p. 103-13.
14. Jacobs, L.D., R.W. Beck, J.H. Simon, R.P. Kinkel, et al., *Intramuscular interferon beta-1a therapy initiated during a first demyelinating event in multiple sclerosis.* New England Journal of Medicine, 2000. 343(13): p. 898-904.
15. Comi, G., M. Filippi, F. Barkhof, L. Durelli, et al., *Effect of early interferon treatment on conversion to definite multiple sclerosis: a randomised study.* Lancet, 2001. 357(9268): p. 1576-82.
16. Filippi, M., M. Rovaris, M. Inglese, F. Barkhof, et al., *Interferon beta-1a for brain tissue loss in patients at presentation with syndromes suggestive of multiple sclerosis:*

a randomised, double-blind, placebo-controlled trial. Lancet, 2004. 364(9444): p. 1489-96.
17. Kappos, L., M.S. Freedman, C.H. Polman, G. Edan, et al., *Effect of early versus delayed interferon beta-1b treatment on disability after a first clinical event suggestive of multiple sclerosis: a 3-year follow-up analysis of the BENEFIT study.* Lancet, 2007. 370(9585): p. 389-97.
18. Kappos, L., M.S. Freedman, C.H. Polman, G. Edan, et al., *Long-term effect of early treatment with interferon beta-1b after a first clinical event suggestive of multiple sclerosis: 5-year active treatment extension of the phase 3 BENEFIT trial.* Lancet Neurol, 2009. 8(11): p. 987-97.
19. Kroksveen, A.C., J.A. Opsahl, T.T. Aye, R.J. Ulvik, et al., *Proteomics of human cerebrospinal fluid: discovery and verification of biomarker candidates in neurodegenerative diseases using quantitative proteomics.* J Proteomics, 2011. 74(4): p. 371-88.
20. Keshishian, H., T. Addona, M. Burgess, E. Kuhn, et al., *Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution.* Mol Cell Proteomics, 2007. 6(12): p. 2212-29.
21. Teunissen, C.E., A. Petzold, J.L. Bennett, F.S. Berven, et al., *A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking.* Neurology, 2009. 73(22): p. 1914-22.
22. Abbatiello, S.E., D.R. Mani, H. Keshishian, and S.A. Carr, *Automated detection of inaccurate and imprecise transitions in peptide quantification by multiple reaction monitoring mass spectrometry.* Clin Chem, 2010. 56(2): p. 291-305.
23. Stoop, M.P., V. Singh, L.J. Dekker, M.K. Titulaer, et al., *Proteomics comparison of cerebrospinal fluid of relapsing remitting and primary progressive multiple sclerosis.* PLoS One, 2010. 5(8): p. e12442.
24. Ottervald, J., B. Franzen, K. Nilsson, L.I. Andersson, et al., *Multiple sclerosis: Identification and clinical evaluation of novel CSF biomarkers.* J Proteomics, 2010. 73(6): p. 1117-32.
25. Axelsson, M., C. Malmestrom, S. Nilsson, S. Haghighi, et al., *Glial fibrillary acidic protein: a potential biomarker for progression in multiple sclerosis.* J Neurol, 2011. 258(5): p. 882-8.
26. Frederiksen, J., K. Kristensen, J. Bahl, and M. Christiansen, *Tau protein: a possible prognostic factor in optic neuritis and multiple sclerosis.* Mult Scler, 2011.
27. Ignacio, R.J., P. Liliana, and C. Edgardo, *Oligoclonal bands and MRI in clinically isolated syndromes: predicting conversion time to multiple sclerosis.* J Neurol, 2010. 257(7): p. 1188-91.
28. Fisniku, L.K., P.A. Brex, D.R. Altmann, K.A. Miszkiel, et al., *Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis.* Brain, 2008. 131(Pt 3): p. 808-17.
29. Comabella, M., M. Fernandez, R. Martin, S. Rivera-Vallve, et al., *Cerebrospinal fluid chitinase 3-like 1 levels are associated with conversion to multiple sclerosis.* Brain, 2010.
30. Stoop, M.P., L.J. Dekker, M.K. Titulaer, R.J. Lamers, et al., *Quantitative matrix-assisted laser desorption ionization-fourier transform ion cyclotron resonance (MALDI-FT-ICR) peptide profiling and identification of multiple-sclerosis-related proteins.* J Proteome Res, 2009. 8(3): p. 1404-14.
31. Liu, T., K.C. Donahue, J. Hu, M.P. Kurnellas, et al., *Identification of differentially expressed proteins in experimental autoimmune encephalomyelitis (EAE) by*

proteomic analysis of the spinal cord. Journal of Proteome Research, 2007. 6(7): p. 2565-75.
32. Noben, J.P., D. Dumont, N. Kwasnikowska, P. Verhaert, et al., *Lumbar cerebrospinal fluid proteome in multiple sclerosis: Characterization by ultrafiltration, liquid chromatography, and mass spectrometry.* Journal of Proteome Research, 2006. 5(7): p. 1647-1657.
33. Anderson, L. and C.L. Hunter, *Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins.* Mol Cell Proteomics, 2006. 5(4): p. 573-88.
34. Licastro, F., M. Mallory, L.A. Hansen, and E. Masliah, *Increased levels of alpha-1-antichymotrypsin in brains of patients with Alzheimer's disease correlate with activated astrocytes and are affected by APOE 4 genotype.* J Neuroimmunol, 1998. 88(1-2): p. 105-10.
35. Abraham, C.R., D.J. Selkoe, and H. Potter, *Immunochemical identification of the serine protease inhibitor alpha 1-antichymotrypsin in the brain amyloid deposits of Alzheimer's disease.* Cell, 1988. 52(4): p. 487-501.
36. Koo, E.H., C.R. Abraham, H. Potter, L.C. Cork, et al., *Developmental expression of alpha 1-antichymotrypsin in brain may be related to astrogliosis.* Neurobiol Aging, 1991. 12(5): p. 495-501.
37. Mucke, L., G.Q. Yu, L. McConlogue, E.M. Rockenstein, et al., *Astroglial expression of human alpha(1)-antichymotrypsin enhances alzheimer-like pathology in amyloid protein precursor transgenic mice.* Am J Pathol, 2000. 157(6): p. 2003-10.
38. Esiri, M.M. and J. Booss, *Comparison of methods to identify microglial cells and macrophages in the human central nervous system.* J Clin Pathol, 1984. 37(2): p. 150-6.
39. Flower, D.R., *The lipocalin protein family: structure and function.* Biochem J, 1996. 318 (Pt 1): p. 1-14.
40. Smith, K.M., R.M. Lawn, and J.N. Wilcox, *Cellular localization of apolipoprotein D and lecithin:cholesterol acyltransferase mRNA in rhesus monkey tissues by in situ hybridization.* J Lipid Res, 1990. 31(6): p. 995-1004.
41. Provost, P.R., L. Villeneuve, P.K. Weech, R.W. Milne, et al., *Localization of the major sites of rabbit apolipoprotein D gene transcription by in situ hybridization.* J Lipid Res, 1991. 32(12): p. 1959-70.
42. Hu, C.Y., W.Y. Ong, R.K. Sundaram, C. Chan, et al., *Immunocytochemical localization of apolipoprotein D in oligodendrocyte precursor-like cells, perivascular cells, and pericytes in the human cerebral cortex.* J Neurocytol, 2001. 30(3): p. 209-18.
43. Navarro, A., J. Tolivia, A. Astudillo, and E. del Valle, *Pattern of apolipoprotein D immunoreactivity in human brain.* Neurosci Lett, 1998. 254(1): p. 17-20.
44. Ong, W.Y., C.P. Lau, S.K. Leong, U. Kumar, et al., *Apolipoprotein D gene expression in the rat brain and light and electron microscopic immunocytochemistry of apolipoprotein D expression in the cerebellum of neonatal, immature and adult rats.* Neuroscience, 1999. 90(3): p. 913-22.
45. Do Carmo, S., H. Jacomy, P.J. Talbot, and E. Rassart, *Neuroprotective effect of apolipoprotein D against human coronavirus OC43-induced encephalitis in mice.* J Neurosci, 2008. 28(41): p. 10330-8.
46. Boyles, J.K., L.M. Notterpek, and L.J. Anderson, *Accumulation of apolipoproteins in the regenerating and remyelinating mammalian peripheral nerve. Identification of apolipoprotein D, apolipoprotein A-IV, apolipoprotein E, and apolipoprotein A-I.* J Biol Chem, 1990. 265(29): p. 17805-15.

47. Reindl, M., G. Knipping, I. Wicher, E. Dilitz, et al., *Increased intrathecal production of apolipoprotein D in multiple sclerosis.* J Neuroimmunol, 2001. 119(2): p. 327-32.
48. Stoop, M.P., L.J. Dekker, M.K. Titulaer, P.C. Burgers, et al., *Multiple sclerosis-related proteins identified in cerebrospinal fluid by advanced mass spectrometry.* Proteomics, 2008. 8(8): p. 1576-85.
49. Berglund, E., T. Stigbrand, and S.R. Carlsson, *Isolation and characterization of a membrane glycoprotein from human brain with sequence similarities to cell adhesion proteins from chicken and mouse.* Eur J Biochem, 1991. 197(2): p. 549-54.
50. Hu, Q.D., B.T. Ang, M. Karsak, W.P. Hu, et al., *F3/contactin acts as a functional ligand for Notch during oligodendrocyte maturation.* Cell, 2003. 115(2): p. 163-75.
51. Laursen, L.S., C.W. Chan, and C. ffrench-Constant, *An integrin-contactin complex regulates CNS myelination by differential Fyn phosphorylation.* J Neurosci, 2009. 29(29): p. 9174-85.
52. Rithidech, K.N., L. Honikel, M. Milazzo, D. Madigan, et al., *Protein expression profiles in pediatric multiple sclerosis: potential biomarkers.* Multiple Sclerosis, 2009. 15(4): p. 455-464.
53. Qin, Z., Y. Qin, and S. Liu, *Alteration of DBP levels in CSF of patients with MS by proteomics analysis.* Cell Mol Neurobiol, 2009. 29(2): p. 203-10.
54. Hammack, B.N., K.Y.C. Fung, S.W. Hunsucker, M.W. Duncan, et al., Proteomic analysis of multiple sclerosis cerebrospinal fluid. Multiple Sclerosis, 2004. 10(3): p. 245-260.

TABLES

Table 1. Sample information details for the patients and controls included in the (A) iTRAQ discovery and (B) selected reaction monitoring (SRM) verification study.

	(A) Groups for iTRAQ discovery				(B) Groups for SRM verification				
Diagnosis	**CIS-MS**	**CIS**	**OIND**	**Pooled reference sample**	**CIS-MS**	**CIS**	**RRMS**	**OIND**	**OND**
Number of patients/ctr.	5	5	5	12	16	15	36	33	32
Male/Female ratio	2/3	2/3	2/3	7/5	4/12	3/12	8/28	12/21	18/14
Age (years) at LP	34.6 (7.6)	29.0 (6.3)	38.2 (7.5)	54.5 (24.5)	35.3 (11.0)	31.9 (11.7)	36.8 (9.4)	46.3 (14.9)	52.0 (13.5)
CSF protein concentration (µg/mL)	456.2 (67.9)	414.6 (6.6)	593.6 (235.3)	553.7 (145.1)	483.8 (175.9)	422.5 (66.9)	426.2 (97.0)	590.3 (260.4)	468.9 (125.2)
EDSS score at LP	1.6 (1.0-2.0)	1.7 (0.0-3.0)	N/A	N/A	1.8 (0.0 - 3.5)	1.8 (0.0 - 3.5)	1.6 (0.0 - 4.0)	N/A	N/A
% positive for oligoclonal IgG bands	80%	100%	40%	N/A	93.3%	92.9%	97.1%	36.4% *	0%
Follow-up (months)	13.2 (18.6)	32.4 (17.8)	N/A	N/A	8.9 (11.0)	25.1 (10.7)	N/A	N/A	N/A

Age, protein concentration, and follow-up values are mean. Standard deviations are in brackets. The Kurtzke Expanded Disability Status Scale (**EDSS**) score is mean with EDSS score range in brackets. **LP** = Lumbar puncture. **CIS** = Clinically isolated syndrome. **RRMS** = Relapsing-remitting multiple sclerosis. **CIS-MS** = Patients with RRMS that had CIS at the time of LP. **OIND** = Other inflammatory neurological diseases. **OND** = Other neurological diseases. *Information about oligoclonal bands is not available for 3 OIND controls.

Table 2. Experimental design for the iTRAQ biomarker discovery study.

	iTRAQ 114	**iTRAQ 115**	**iTRAQ 116**	**iTRAQ 117**
iTRAQ experiment 1	Reference	CIS-MS 1	CIS 1	OIND 1
iTRAQ experiment 2	Reference	OIND 2	CIS-MS 2	CIS 2
iTRAQ experiment 3	Reference	CIS 3	OIND 3	CIS-MS 3
iTRAQ experiment 4	Reference	CIS-MS 4	CIS 4	OIND 4
iTRAQ experiment 5	Reference	OIND 5	CIS-MS 5	CIS 5

CIS = Clinically isolated syndrome. **CIS-MS** = Patients with relapsing-remitting multiple sclerosis that had CIS at the time of lumbar puncture.
OIND = Other inflammatory neurological diseases.

Table 3. Data analysis information from the iTRAQ discovery study. The table shows the number of quantified proteins, the number of biomarker candidates, the accepted fold changes, and the standard deviation fold changes for each group comparison.

Group comparison	**# quantified proteins**	**Remaining proteins after filtration**	**# biomarker candidates**	**Fold change**	**- 2SD**	**-1SD**	**+1SD**	**+2SD**
CIS-MS vs. CIS	617	7	0	1.50	-1.60	-1.27	1.27	1.60
CIS-MS vs. OIND	613	90	3	2.00	-2.25	-1.49	1.53	2.31
CIS vs. OIND	614	96	4	2.00	-2.04	-1.39	1.52	2.20

C**IS** = Clinically isolated syndrome. **CIS-MS** = Patients with relapsing-remitting multiple sclerosis that had CIS at the time of lumbar puncture. **OIND** = Other inflammatory neurological diseases.

Table 4. Proteins and peptides used for SRM verification.

Protein ID.	Acc. no	Reported differential abundant	Peptide used in SRM	Transition used for quantification	Endogenous peptide		SIS peptide		Collision energy
					Q1	Q3	Q1	Q3	
Alpha-1-acid glycoprotein	P02763	↑ MS vs. healthy ctr [52] ↓ MS vs. non MS ctr [32]	TEDTIFLR	y6	498.0	764.0	503.0	774.0	48
			WFYIASAFR	Interference in transitions	N/A	N/A	N/A	N/A	N/A
Alpha-1-antichymotrypsin	P01011	↑ MS vs. OND [30] ↑CIS-MS vs. CIS-CIS [29] ↑ RRMS, PPMS, SPMS vs. ctr [24] iTRAQ discovery	EQLSLLDR	y5	487.0	603.0	492.0	613.0	38
			EIGELYLPK	y5	531.0	633.0	535.0	641.0	40
			ITLLSALVETR	Interference in all transitions	N/A	N/A	N/A	N/A	N/A
Alpha-2-macroglobulin	P01023	↑ RRMS, PPMS, SPMS vs. ctr [24] ↑ MS vs. OND [30]	NEDSLVFVQTDK	Interference in all transitions	N/A	N/A	N/A	N/A	N/A
			AIGYLNTGYQR	y6	628.2	738.3	633.2	748.3	37
Apolipoprotein A-I	P02647	↑ CIS-MS vs. CIS-CIS [29]	DLATVYVDVLK	Interference in all transitions	N/A	N/A	N/A	N/A	N/A
			THLAPYSDELR	SIS spike in to low	N/A	N/A	N/A	N/A	N/A
Apolipoprotein D	P05090	↑ RRMS and PPMS vs. ctr [23]	NILTSNNIDVK	y8	615.9	890.4	619.9	898.4	46
Ceruloplasmin	P00450	↓ MS vs. OND [30] ↑ CIS-MS vs. CIS-CIS [29]	GAYPLSIEPIGVR	y8	686.8	870.5	691.8	880.5	46
Chromogranin-A	P10645	↑ MS vs. OND [30, 48]	ELQDLALQGAK	y6	593.3	587.4	597.3	595.4	25
			GLSAEPGWQAK	y6	572.2	686.4	576.2	694.4	25
Clusterin	P10909	↑ MS vs. CIS and OND [30, 48] ↑ MS vs. healthy ctr. [52]	IDSLLENDR	y5	537.6	646.3	542.6	656.3	35
Complement C3	P01024	↑ MS vs. OIND and OND [30, 48] ↑ EAE [31]	ISLPESLK	y5	444.0	573.3	448.0	581.3	32
			GLEVTITAR	Interference in all transitions	N/A	N/A	N/A	N/A	N/A
Complement C4	P0C0L4	↓ RRMS vs. OND [53] ↑ MS vs. OIND [48]	LELSVDGAK	y7	466.0	689.4	470.0	697.4	40
Contactin-1	Q12860	↓ MS vs. OND [30] ↑ SPMS vs. ctr [24]	DGEYVVEVR	y6	533.4	764.5	538.4	774.5	14
			FIPLIPIPER	y8	597.8	934.4	602.8	944.4	17
Cystatin C	P01034	↓ RRMS vs. OND [53] ↑ EAE [31]	LVGGPMDASVEEEGVR	y9	823.9	975.4	828.9	985.4	26
Kallikrein-6	Q92876	↑ MS vs. OND [54]	LSELIQPLPLER	y6	704.3	724.4	709.3	734.4	25
Leucine-rich alpha-2-glycoprotein	P02750	iTRAQ discovery	DLLLPQPDLR	y6	590.4	725.4	595.4	735.4	18
Plasminogen	P00747	↑ CIS-MS vs. CIS-CIS [29]	EAQLPVIENK	y6	570.6	699.4	574.6	707.4	34
Prostaglandin D2 synthase	P41222	↑ MS vs. CIS, OIND, and OND [30]	AQGFTEDTIVFLPQTDK	Interference in all transitions	N/A	N/A	N/A	N/A	N/A
			WFSAGLASNSSQLR	Interference in all transitions	N/A	N/A	N/A	N/A	N/A
Secretogranin II	P13521	↓ CIS-MS vs. CIS [29]	IILEALR	y5	415.0	601.0	420.0	611.0	38
			VLLEYNQEK	Interferece in transition y6, endogenous peptide low for most patients	N/A	N/A	N/A	N/A	N/A
Semaphorin-7A	O75326	↓ CIS-MS vs. CIS-CIS [29]	VYLFDFPEGK	y8	607.7	952.4	611.7	960.4	20
Serotransferrin	P02787	↑ EAE [31] ↑ MS cs. Non-MS ctr [32]	YLGEEYVK	y6	501.0	724.3	505.0	732.3	16
Serum albumin	P02768	↓ RRMS vs. OND [53] ↓ MS vs. CIS [30]	LVNEVTEFAK	y8	575.8	937.5	579.8	945.5	34
Transthyretin	P02766	↓ RRMS vs. OND [53] ↓ MS vs. OND [30] ↑ MS vs. healthy ctr. [52]	AADDTWEPFASGK	y8	697.7	921.4	701.7	929.4	45
Vitronectin	P04004	↑ CIS-MS vs. CIS-CIS [29]	FEDGVLDPDYPR	y6	711.9	762.3	716.9	772.3	40

Table 5. Biomarker candidates that showed significant abundance difference in the SRM verification study.

Group Comparison[a]	Protein ID.[b]	Peptide[c]	Abundance[d]	Fold Change[e]	*p*-value[f]	CV quality control[g]
CIS-MS vs. RRMS	Contactin-1	FIPLIPIPER	↓ CIS-MS	-1.16	0.0280	19.2
CIS-MS vs. OIND	Apolipoprotein D	NILTSNNIDVK	↓ CIS-MS	-1.30	0.0266	14.1
	Clusterin	IDSLLENDR	↑ CIS-MS	1.67	0.0078	25.4
CIS-MS vs. OND	Apolipoprotein D	NILTSNNIDVK	↓ CIS-MS	-1.36	0.0033	14.1
	Contactin-1	DGEYVVEVR	↓ CIS-MS	-1.34	0.0122	12.4
	Contactin-1	FIPLIPIPER	↓ CIS-MS	-1.53	0.0002	19.2
	Kallikrein-6	LSELIQPLPLER	↓ CIS-MS	-1.56	0.0010	11.3
CIS vs. RRMS	Serotransferrin	YLGEEYVK	↑ CIS	1.12	0.0451	17.0
CIS vs. OIND	Alpha-1-antichymotrypsin	EIGELYLPK	↓ CIS	-1.40	0.0267	11.1
	Clusterin	IDSLLENDR	↑ CIS	1.29	0.0487	25.4
CIS vs. OND	Apolipoprotein D	NILTSNNIDVK	↓ CIS	-1.23	0.0101	14.1
	Chromogranin-A	GLSAEPGWQAK	↓ CIS	-1.36	0.0372	10.9
	Contactin-1	DGEYVVEVR	↓ CIS	-1.35	0.0245	12.4
	Contactin-1	FIPLIPIPER	↓ CIS	-1.62	0.0018	19.2
	Cystatin-C	LVGGPMDASVEEEGVR	↓ CIS	-1.47	0.0017	20.3
	Kallikrein-6	LSELIQPLPLER	↓ CIS	-1.39	0.0075	11.3
RRMS vs. OIND	Alpha-1-acid glycoprotein	TEDTIFLR	↓ RRMS	-1.44	0.0006	28.2
	Alpha-1-antichymotrypsin	EQLSLLDR	↓ RRMS	-1.24	0.0217	14.1
	Alpha-2-macroglobulin	AIGYLNTGYQR	↓ RRMS	-1.21	0.0410	8.3
	Complement C3	ISLPESLK	↓ RRMS	-1.20	0.0090	40.4
RRMS vs. OND	Apolipoprotein D	NILTSNNIDVK	↓ RRMS	-1.14	0.0085	14.1
OIND vs. OND	Alpha-1-antichymotrypsin	EQLSLLDR	↑ OIND	1.39	0.0069	14.1
	Alpha-1-antichymotrypsin	EIGELYPK	↑ OIND	1.22	0.0318	11.1
	Contactin-1	FIPLIPIPER	↓ OIND	-1.25	0.0067	19.2
	Cystatin-C	LVGGPMDASVEEEGVR	↓ OIND	-1.26	0.0163	20.3
	Kallikrein-6	LSELIQPLPLER	↓ OIND	-1.23	0.0124	11.3

CIS = Clinically isolated syndrome. **RRMS** = Relapsing-remitting multiple sclerosis. **CIS-MS** = Patients with RRMS that had CIS at the time of lumbar puncture. **OIND** = Other inflammatory neurological diseases. **OND** = Other neurological diseases.
[a] The group comparison where the biomarker candidate showed significant abundance difference. [b] Protein name for the biomarker candidate. Accession number are in Table 4. [c] The signature peptide used for quantification. [d] The abundance of the biomarker candidate in the group comparison. ↑ indicate increased abundance and ↓ indicates decreased abundance in that patient/control group. [e] The fold change between the compared groups, calculated from the median area ratio from SRM analysis. [f] Student's t-test *p*-value. *p*-value $\leq$ 0.05 is considered significant. [g] CV value for the quality control used for evaluation of the SID-SRM analysis. The CV value is calculated from 21 CSF samples analyzed as processed replicates.

Table 6. Proteins with differential abundance from the iTRAQ discovery experiment.

Group comparison	Protein ID.	Acc. no.	Fold Change	Abundance	# Peptides	% Coverage	*p*-value
CIS-MS vs. OIND	Alpha-1-antichymotrypsin	P01011	-2.15	↓ CIS-MS	27	62	0.0006
	Leucine-rich alpha-2-glycoprotein	P02750	-2.13	↓ CIS-MS	15	47	0.0213
	Lysozyme C	P61626	-2.33	↓ CIS-MS	10	66	0.0402
CIS vs. OIND	Alpha-1-antichymotrypsin	P01011	-2.00	↓ CIS	27	62	0.0017
	Carboxypeptidase N subunit 2	P22792	-2.83	↓ CIS	4	13	0.0217
	Complement component C5	P01031	-2.47	↓ CIS	44	28	0.0102
	Leucine-rich alpha-2-glycoprotein	P02750	-2.31	↓ CIS	15	47	0.0161

CIS = Clinically isolated syndrome. **CIS-MS** = Patients with relapsing-remitting multiple sclerosis that had CIS at the time of lumbar puncture. **OIND** = Other inflammatory neurological diseases.

FIGURE LEGENDS

Figure 1. Experimental workflow for (A) iTRAQ discovery experiment and (B) SRM verification. (A) Five CIS patients, five CIS-MS patients, and five OIND controls were grouped into five iTRAQ experiments. The CSF was protein depleted, digested, and iTRAQ labeled. The samples in each experiment were pooled and SCX fractionated prior to Orbitrap MS/MS analysis. All five experiments were combined for final data analysis. **(B)** Signature peptides for biomarker candidates both from the discovery study and literature were defined, and stable isotope labeled internal standards (SISs) were optimized on a Q-trap 5500 mass spectrometer. Response curves were generated to examine the assay linearity and determine the correct spike-in of SIS, here illustrated by apolipoprotein D (apo D). As shown, the SIS spike-in for the apo D peptide is approximately 1:1 to the endogenous peptide. CSF from 132 individuals were digested and spiked with SIS. The abundance of 20 biomarker candidates were analyzed with scheduled SRM.

Figure 2. Box-plots showing SRM results for alpha-1-antichymotrypsin (ACT) and leucine-rich alpha-2-glycoprotein (LRG). * indicate p-value ≤ 0.05, and ** indicate p-value ≤ 0.01 between the groups where the line is drawn. The box-plots represent the 75^{th} and 25^{th} percentiles, where the median divides the box horizontally. Whiskers are drawn to the minimum and maximum value in the data for that patient/control group. Abbreviations: CIS: Clinically isolated syndrome RRMS: Relapsing-remitting multiple sclerosis (RRMS).. CIS-MS: Patients with RRMS that had CIS at the time of lumbar puncture. OIND: Other inflammatory neurological disease. OND: Other neurological disease.

Figure 3. **Box-plots showing SRM results for the literature derived proteins apolipoprotein D (apo D), contactin-1, kallikrein-6, and clusterin.** * indicate p-value ≤ 0.05, and ** indicate p-value ≤ 0.01 between the groups where the line is drawn. The box-plots represent the 75^{th} and 25^{th} percentiles, where the median divides the box horizontally. Whiskers are drawn to the minimum and maximum value in the data for that patient/control group. Abbreviations: CIS: Clinically isolated syndrome. RRMS: Relapsing-remitting multiple sclerosis (RRMS). CIS-MS: Patients with RRMS that had CIS at the time of lumbar puncture. OIND: Other inflammatory neurological disease. OND: Other neurological disease.

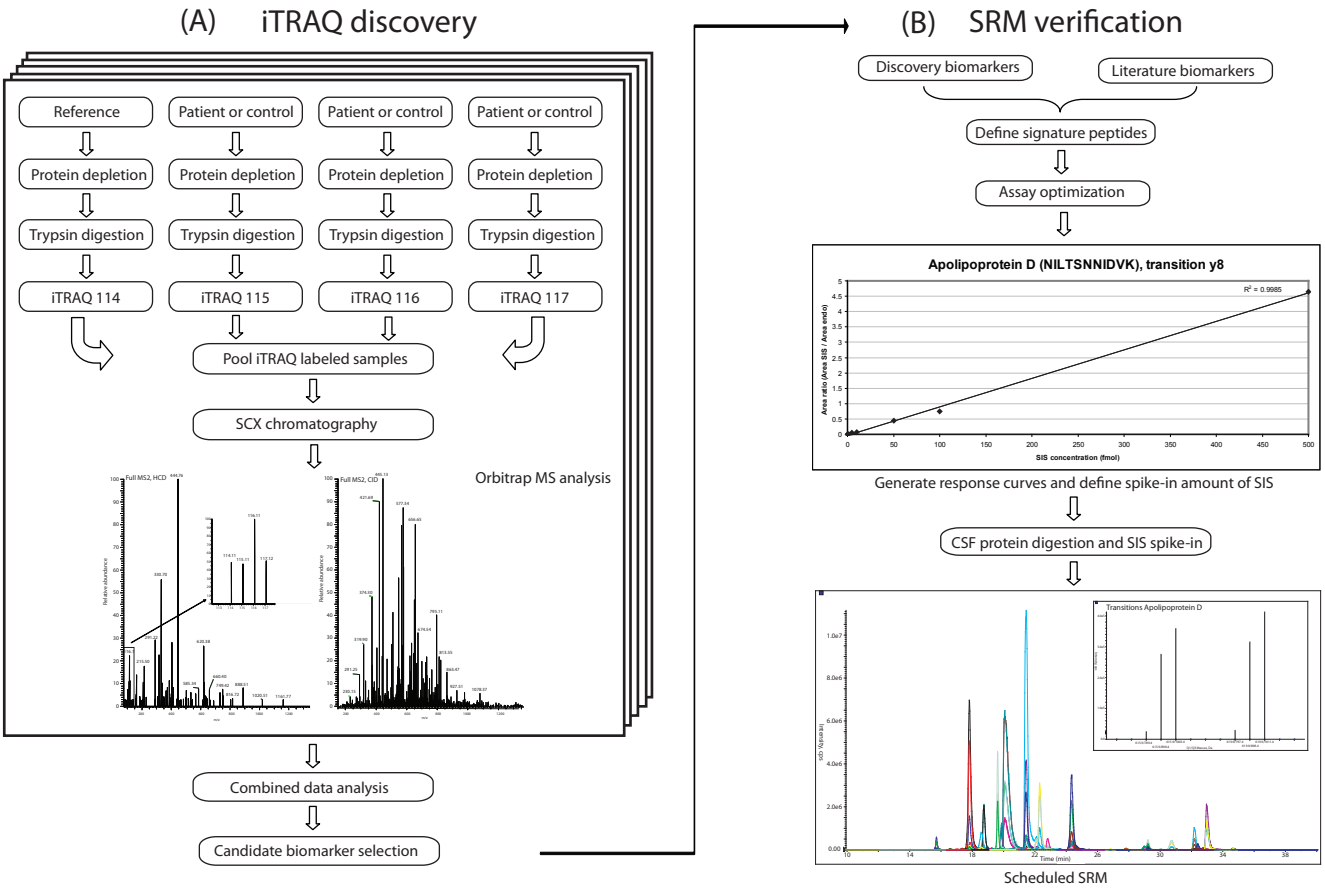
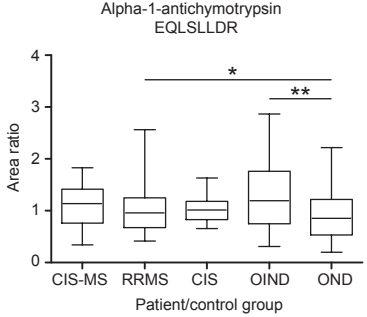
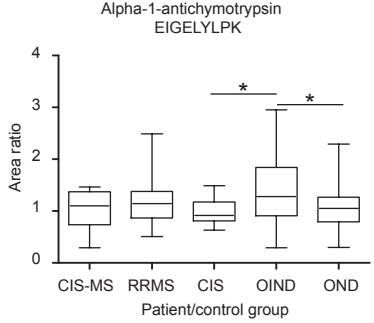

(A) iTRAQ discovery
Reference
Patient or control
Patient or control
Patient or control
Protein depletion
Protein depletion
Protein depletion
Protein depletion
Trypsin digestion
Trypsin digestion
Trypsin digestion
Trypsin digestion
iTRAQ 114
iTRAQ 115
iTRAQ 116
iTRAQ 117
Pool iTRAQ labeled samples
SCX chromatography
Orbitrap MS analysis
Combined data analysis
Candidate biomarker selection
(B) SRM verification
Discovery biomarkers
Literature biomarkers
Define signature peptides
Assay optimization
Apolipoprotein D (NILTSNNIDVK), transition y8
R² = 0.9985
Area ratio (Area SIS / Area endo)
SIS concentration (fmol)
Generate response curves and define spike-in amount of SIS
CSF protein digestion and SIS spike-in
Transitions Apolipoprotein D
Time (min)
Scheduled SRM

Figure 2. Kroksveen *et al*.





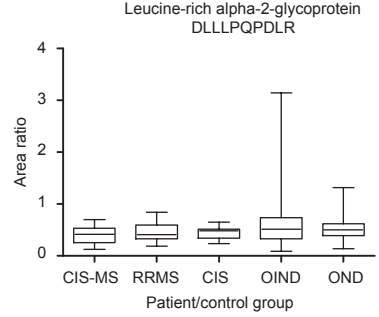


Figure 3. Kroksveen *et al*.