

## RESEARCH ARTICLE

# Antibody-based profiling of cerebrospinal fluid within multiple sclerosis

Anna Häggmark<sup>1</sup>, Sanna Byström<sup>1</sup>, Burcu Ayoglu<sup>1</sup>, Ulrika Qundos<sup>1</sup>, Mathias Uhlén<sup>1</sup>, Mohsen Khademi<sup>2</sup>, Tomas Olsson<sup>2</sup>, Jochen M. Schwenk<sup>1</sup> and Peter Nilsson<sup>1</sup>

<sup>1</sup> SciLifeLab Stockholm, School of Biotechnology, KTH Royal Institute of Technology, Stockholm, Sweden

<sup>2</sup> Neuroimmunology Unit, Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden

Antibody suspension bead arrays have proven to enable multiplexed and high-throughput protein profiling in unfractionated plasma and serum samples through a direct labeling approach. We here describe the development and application of an assay for protein profiling of cerebrospinal fluid (CSF). While setting up the assay, systematic intensity differences between sample groups were observed that reflected inherent sample specific total protein amounts. Supplementing the labeling reaction with BSA and IgG diminished these differences without impairing the apparent sensitivity of the assay. We also assessed the effects of heat treatment on the analysis of CSF proteins and applied the assay to profile 43 selected proteins by 101 antibodies in 339 CSF samples from a multiple sclerosis (MS) cohort. Two proteins, GAP43 and SERPINA3 were found to have a discriminating potential with altered intensity levels between sample groups. GAP43 was detected at significantly lower levels in secondary progressive MS compared to early stages of MS and the control group of other neurological diseases. SERPINA3 instead was detected at higher levels in all MS patients compared to controls. The developed assay procedure now offers new possibilities for broad-scale protein profiling of CSF within neurological disorders.

Received: December 21, 2012

Revised: March 19, 2013

Accepted: April 9, 2013

## Keywords:

Antibody microarrays / Biomarker discovery / Cerebrospinal fluid / Multiplexed proteomics technology / Protein arrays / Proteome profiling



Additional supporting information may be found in the online version of this article at the publisher's web-site

## 1 Introduction

The use of affinity reagents in microarray formats has been widely described and utilized for protein profiling in the field of biomarker discovery [1]. Most studies have focused on the

analysis of blood-derived sample material since blood is a systemic and clinically preferred body fluid that can be collected with relatively low invasiveness. However, in some biological contexts, analysis of a fluid in closer proximity to the affected tissue or organ could be beneficial to reveal local changes and to overcome limited availability of tissue material. For neurological disorders the access to tissue material is particularly restricted, which increases the analytical value of cerebrospinal fluid (CSF) as the local source for acquiring additional information regarding the status of the central nervous system (CNS). One relatively common neurological disease is multiple sclerosis (MS) [2], which is generally regarded as an inflammatory neurodegenerative disease characterized by the demyelination of axons [3]. Patients are classified into one of four subtypes—relapsing remitting MS (RRMS), primary progressive MS (PPMS), secondary progressive MS (SPMS), and progressive relapsing MS—which

**Correspondence:** Professor Peter Nilsson, SciLifeLab Stockholm, School of Biotechnology, KTH Royal Institute of Technology, Stockholm SE-17121, Sweden  
**E-mail:** peter.nilsson@scilifelab.se

**Abbreviations:** CSF, cerebrospinal fluid; CTB, cholera toxin subunit B; GAP43, growth associated protein 43; HPA, Human Protein Atlas; MS, multiple sclerosis; OND, other neurological diseases; PPMS, primary progressive MS; PQN, probabilistic quotient normalization; RRMS, relapsing remitting MS; SERPINA1, serin peptidase inhibitor, clade A, member 1; SERPINA3, serin peptidase inhibitor, clade A, member 3; SPMS, secondary progressive MS

differ in regard to whether remission appears between clinical episodes and to the extent of recovery between relapses. Today, diagnosis of stages and subtypes of MS is based on certain criteria [4], but since symptoms are very diverse and only few of them are disease specific, the process of diagnosis usually takes a long time. There is therefore a great need of additional markers to assist the clinical diagnosis of MS. Furthermore, individual variation in the disease course, for example, the conversion from the RRMS subtype into the progressive stage (SPMS), is not yet clearly understood, calling for exploration of not only diagnostic markers but also the underlying disease mechanisms [5].

Apart from proteomic studies utilizing various MS-based methods [6–8], the majority of available literature on protein content of body fluids is based on information from ELISA-type assays, which for many years have been the “golden standard” in clinical routine. Today, traditional sandwich ELISA setups are increasingly being complemented by multiplexed versions to meet the growing interest in biomarker panels rather than single markers [9]. Even though sandwich assays are desired due to their sensitivity and specificity compared to a single-binder assay, the possibilities of multiplexing in terms of analytes are limited [10]. In recent years, new technologies have emerged, combining the concept of single-binder analysis and sandwich setups with, for example, flow cytometry-based methods, bringing the potential of multiplexing into the field of affinity-based biomarker discovery. To fully make use of the multiplexing capacity of such methods, the availability of large numbers of affinity reagents is of great importance [1]. One initiative aiming to produce antibodies toward all human proteins is the Human Protein Atlas (HPA) [11] where antibodies are generated, validated, and used to study protein expression in various normal and cancerous tissues and cell lines. The current version 11.0 hosts more than 18 000 antibodies to approximately 75% of the human protein encoding genes (>15 000 genes) and the results are publicly available online ([www.proteinatlas.org](http://www.proteinatlas.org)).

The study presented here aimed at exploring the possibilities for high-throughput protein profiling of CSF using an antibody-based method. Here, we describe the adaptation of an assay for CSF analysis based on an established protocol for multiplexed analysis of serum and plasma [12] and the influence of sample preparation parameters. Furthermore, we investigated CSF-derived protein profiles in a quest to discover disease-related proteins in the context of MS.

## 2 Materials and methods

### 2.1 Samples

CSF and EDTA plasma samples were collected during routine neurological diagnostic investigations at the Karolinska University Hospital, Sweden. All CSF samples were centrifuged ( $350 \times g$ ) immediately after lumbar puncture to isolate the cells and the cell-free CSF supernatant was stored at  $-80^{\circ}\text{C}$

until use. The study was approved by the regional ethical committee and a written informed consent was obtained from all patients. In total, 339 CSF and plasma samples were analyzed out of which 209 were either from patients diagnosed with primary progressive MS, RRMS, and SPMS subtypes or from patients with one episode of neurological symptoms (referred to as clinically isolated syndrome, CIS). MS cases diagnosed with RRMS were further categorized based on if the sample was collected during relapse or remission (RRMSrel and RRMSrem respectively). Samples from patients with other neurological diseases (OND) either with or without signs of inflammation (iOND and OND respectively,  $n = 130$ ) were used as controls. For the majority of samples (80%), clinical information was available on IgG index (CSF IgG/albumin ratio over serum IgG/albumin ratio) as well as IgG and albumin quotient (CSF over plasma IgG and albumin content, respectively), see Table 1 for sample demographics.

Assay development was performed using 65 CSF samples (25 OND, 30 RRMS, and 10 SPMS) and this set was for subsequent verification expanded to include all 339 samples. The relative total protein concentrations of samples in the initial subset were determined by the median of triplicate absorbance measurements (280 nm, Nanodrop instrument) of CSF diluted 1:2 in PBS (Phosphate buffered saline, pH 7.4).

### 2.2 Antibodies and coupling

A set of target proteins for the assay protocol development was chosen either based on potential disease relevance according to previous literature or by involvement in inflammatory mechanisms, resulting in 71 antibodies targeting 69 unique proteins. For analysis of the large sample set, selected antibodies from the initial set were complemented with additional targets, making up an array of 101 antibodies targeting 43 proteins, see Supporting Information Table 1 for a full representation of antibodies and targets. The majority of antibodies were provided by the HPA and immobilization onto magnetic beads (Luminex Corporation) was conducted as previously described [12]. In brief, all antibodies were diluted to a concentration of  $1.6 \mu\text{g/mL}$  in buffer (2[N-Morpholino]ethanesulphonic acid, pH 5.0) and incubated with the beads after activation of the carboxylic surface using 1-ethyl-3-[3]-dimethylaminopropyl and *N*-hydroxysulphosuccinimide. In addition to the selected protein targets, one bead identity was used for immobilization of an albumin-binding antibody (Dako), one bead identity for anti-human IgG (Jackson ImmunoResearch), one for rabbit IgG (Bethyl) and one bead identity was incubated with protein-free buffer. After incubation, the coupled beads were washed and stored in a blocking reagent for ELISA (BRE, Roche) before combining all bead identities to create a bead array in suspension.

### 2.3 Assay protocol development

In this study, the protocol for protein profiling of plasma, as previously reported by Schwenk et al. [12], was adapted for

**Table 1.** Sample demographics. The table provides information on gender, age, and the clinical values for IgG index and IgG and albumin quotient per sample group

| Sample group | N   | Female % | Age  |       | IgG index |         | Albumin quotient |       |
|--------------|-----|----------|------|-------|-----------|---------|------------------|-------|
|              |     |          | Mean | Range | Median    | Range   | Median           | Range |
| PPMS         | 9   | 78       | 50   | 35–62 | 0.9       | 0.3–1.4 | 7.2              | 4–307 |
| RRMS         | 144 | 76       | 39   | 17–70 | 0.8       | 0.4–3.3 | 5.0              | 2–17  |
| Relapsing    | 24  | 64       | 40   | 22–68 | 0.8       | 0.4–1.8 | 4.9              | 2–10  |
| Remitting    | 120 | 78       | 39   | 17–70 | 0.8       | 0.4–3.3 | 5.0              | 2–17  |
| SPMS         | 31  | 52       | 53   | 35–62 | 0.7       | 0.5–2.2 | 5.7              | 2–11  |
| CIS          | 25  | 68       | 39   | 21–60 | 0.7       | 0.4–2.1 | 4.6              | 1–8   |
| OND          | 79  | 76       | 39   | 19–68 | 0.5       | 0.3–0.8 | 4.6              | 2–11  |
| iOND         | 51  | 74       | 45   | 18–72 | 0.5       | 0.3–2.1 | 5.7              | 2–42  |

protein profiling of CSF. Parameters for sample preparation were assessed by comparing sample dilution buffers for direct labeling and two temperatures during heat treatment for epitope retrieval.

CSF samples were diluted 1:2 in PBS or PBS supplemented with 0.5% w/v bovine serum albumin (BSA, Sigma) and 0.1% w/v rabbit IgG (Bethyl), denoted BIG buffer, using a liquid handler (SELMA, CyBio) and labeled at 4°C with a tenfold molar excess of biotin (NHS-PEG4-Biotin, Thermo Scientific) over total protein content (approximated to 0.5 mg/mL for all crude CSF samples, average protein size estimated to 70 kDa). The labeling was terminated after 2 h by addition of 1 M Tris-HCl in a 250-fold molar excess over biotin and the reaction was allowed to occur for 20 min. Using the liquid handler, a second dilution step of 1:8 was performed with a buffer composed of 0.5% w/v PVA, 0.8% (w/v) PVP and 0.1% w/v casein (all Sigma-Aldrich) supplemented with 10% v/v rabbit IgG, denoted PVXCas buffer. The samples were then heat treated either at 56 or 72°C for 30 min and cooled to 20°C for 15 min in a thermocycler (Techne TC-PLUS). Following this, 45 µL of sample was incubated with 5 µL of the bead array overnight at 20°C under permanent rotation on a shaker (Grant). The bead-sample mixture was subsequently washed with 3 × 100 µL PBST (1 × PBS, 0.05% Tween 20) using a wash station (BioTek EL406). Bound targets were cross-linked to the immobilized antibodies by addition of 50 µL 0.4% paraformaldehyde in PBS for 10 min, followed by a bead-washing step. For detection, 50 µL of a streptavidin conjugated fluorophore (R-phycoerythrin, Invitrogen) diluted to 0.5 µg/mL in PBST was added and incubated for 20 min. Finally, the beads were washed before addition of 100 µL PBST for measurement in the LX200 or FlexMap3D instrument (both Luminex Corporation). For each bead identity, at least 50 events were counted, with 34 events required for reliable results and the median fluorescence intensity was used as read-out of the assay. See Fig. 1 for a schematic overview of the assay workflow and study design.

Plasma samples were analyzed according to the protocol above with biotinylation in PBS at a 1:10 dilution, a second dilution of 1:50 in PVXCas and heat treatment at 56°C for

30 min with cooling to 20°C for 15 min. Analysis of the sample set for assay development consisting of 65 CSF samples was performed with samples labeled both in PBS and BIG buffer while the increased set of 339 samples was labeled only in BIG buffer prior to analysis. In addition to the targeted bead arrays, an undirected screening of the PBS labeled samples was made using a bead array with HPA antibodies targeting 384 different proteins that were selected only based on the antibody concentration and not by any disease relevance.

## 2.4 Limit of detection

To investigate how sample complexity and buffer composition in this direct labeling assay affected the ability to detect proteins, a LOD experiment was conducted. A sample pool containing the initial 65 CSF samples was spiked-in with cholera toxin subunit B (CTB, Sigma) prior labeling. This was made in an 11 step dilution series of 12 000–68 pg/mL and a CTB-free sample was included as a reference. The CSF sample pool was labeled both in PBS and BIG buffer and a plasma sample and a buffer control containing 0.1% BSA in PBS were also spiked-in in the same way for comparison purposes. All samples were analyzed in triplicates according to the CSF protocol with 56°C heat treatment and CTB was detected using an anti-CTB antibody (Pierce) immobilized on one bead identity. The LOD was evaluated based on a five-parametric logistic regression using the mean intensity plus 3 × SD of the nonspiked reference sample.

## 2.5 Sandwich assay

To confirm the profiles for growth associated protein 43 (GAP43) from the single binder assay with the direct labeling approach, a sandwich assay was setup utilizing all three antibodies toward this protein as capture reagents and one of the three (HPA015600) as detection antibody. Immobilization onto beads was conducted as described above and for

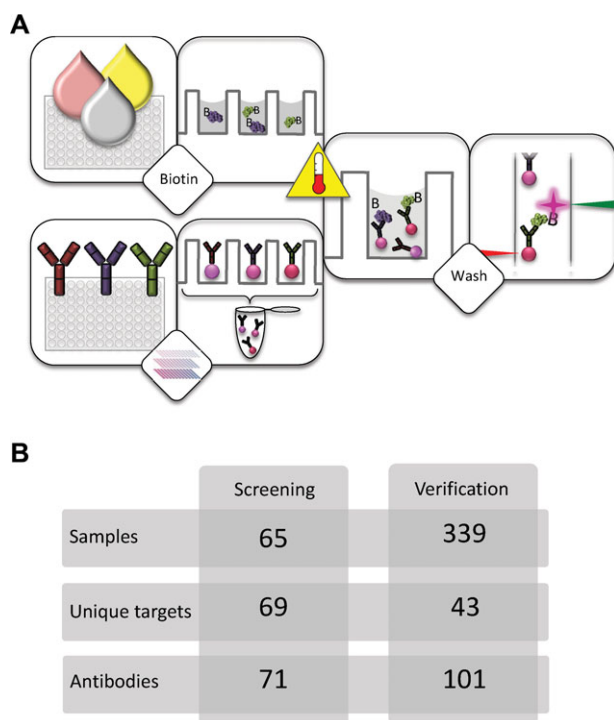


**Figure 1.** Overview of workflow and study setup. (A) The suspension bead array technology enables the analysis of directly labeled samples for multiplex protein profiling. Crude samples are distributed in a randomized fashion in 96-well plates before they are diluted and labeled with biotin. In parallel, antibodies are immobilized on magnetic, color-coded beads and all bead identities are subsequently pooled to create an antibody array in suspension. The labeled samples are heat treated before their incubation with the bead array. After this, unbound proteins are washed away and a streptavidin-conjugated fluorophore added for target detection. In the used instrumentation, one laser is for the identity of beads and the other laser for the fluorophore used for detection of captured, biotinylated target proteins. (B) In this study, a small set of samples ( $n = 65$ ) was utilized for assay development and an initial screening with 71 antibodies targeting 69 unique proteins. At a later stage, a subset of 43 targets selected based on results from the screening were additionally analyzed in the set of 339 CSF samples. At this point, multiple antibodies per target were included, resulting in a 101-plex bead array.

detection, the antibody was labeled with R-phycoerythrin conjugated Fab fragments (Zenon, Invitrogen) and diluted to 1  $\mu\text{g}/\text{mL}$  in 1% BSA w/v in PBST before addition to samples. Nine samples with high intensity in direct labeling analysis from the RRMS group together with nine with low intensity SPMS samples were diluted 1:5 in PVXCas buffer without supplement of rabbit IgG, heat treated at 56°C for 30 min and incubated with the bead array overnight. Washing of beads was performed as described for the direct labeling procedure and incubation with 50  $\mu\text{L}$  of detection reagent allowed for 30 min followed by a new round of washing and addition of 100  $\mu\text{L}$  PBST prior to analysis in the LX200 instrument.

## 2.6 Data analysis

Data evaluation and statistical analysis was performed using R [13]. For analysis during assay development, both unprocessed data as well as data normalized by the probabilistic quotient normalization (PQN) method were used [14, 15]. This normalization method accounts for sample dilution effects, thus facilitating comparability of relative protein abundances between one type of sample collected from different donors. CVs were used as estimates of technical performance of experiments and calculated for each bead identity based on samples analyzed in triplicates. Sample outliers were identified through robust principle component analysis [16] and excluded from further data analysis. Comparisons between sample groups were performed on normalized data using Wilcoxon rank-sum test or Kruskal-Wallis test and  $p$ -values  $< 0.05$  were regarded as significant. Receiver operating characteristic (ROC) curve analyses were performed using the R-package ROCR.

## 3 Results

To explore the possibilities to utilize massive antibody-based protein profiling of CSF we developed an assay protocol for direct sample labeling and analysis with the suspension bead array technology, see Fig. 1. Based on the previously reported plasma profiling protocol [12], dilution factors were adjusted to compensate for the lower total protein concentration in CSF as compared to plasma. In addition, different sample dilution buffers were tested for the direct labeling procedure as well as two heat treatment temperatures for epitope retrieval. The developed assay was subsequently applied to profile 339 CSF samples from MS patients and controls with 101 antibodies targeting 43 different proteins. Examples of representative CSF and plasma sample profiles are presented in Supporting Information Fig. 1. Except for the two most abundant proteins, albumin and IgG, other proteins such as complement C3 and complement C5 were also measured with high intensity in both fluids. The profiles were largely similar, although a few targets were found in a relatively higher amount in plasma as compared to CSF.

### 3.1 Effect of total protein content

In this study, the antibody suspension bead array technology was applied to create protein profiles of CSF. The method utilizes biotin for direct labeling of proteins via their primary amines. In a first attempt to analyze CSF samples with a protocol adapted from serum analysis and the undirected selection of 384 antibodies on beads, systematic differences in the comparison of MS and OND samples were observed in raw data. Also, when the same samples were analyzed with the selected set of 71 antibodies, a similar result was observed. Figure 2 shows a representation of fold changes based on the

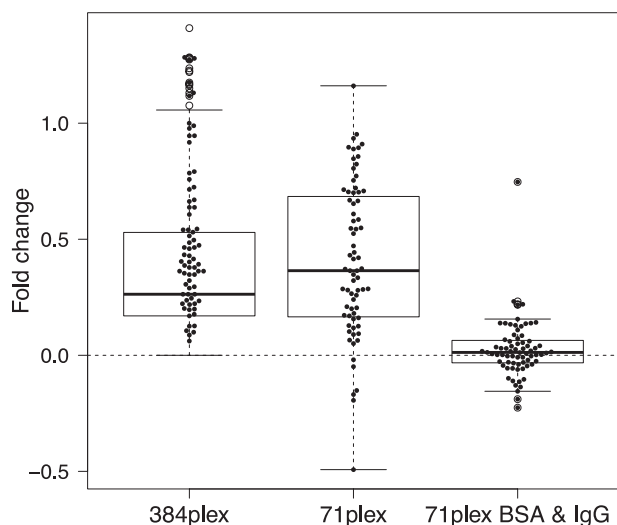


**Figure 2.** Systematic differences between cases and controls. Fold changes for the comparison of all MS patients compared to controls ( $y$ -axis) were calculated for antibodies in the undirected 384-plex array as well as the selected set of 71 antibodies. For samples labeled in PBS, almost all antibodies showed higher mean intensity levels in MS patients compared to the control group. The dotted line represents a log2 fold change of zero, meaning no difference between groups.

comparison of all MS cases to controls among the 65 samples using raw data. Samples labeled in PBS buffer were analyzed with the undirected set of 384 antibodies and samples labeled both in PBS and BIG buffer with the set of 71 antibodies. When CSF was labeled in PBS, essentially all antibodies, both in the undirected as well as the directed array, showed higher signal intensities in the MS group compared to control groups, indicating a systematic effect independent of the specificity of the antibody immobilized on the bead. Supplementing the samples with an excess of albumin and IgG during labeling, however, diminished these effects and the resulting raw data showed fold changes in both directions. In addition to fold change,  $p$ -values were calculated for the detected differences and found to be significant for the majority of comparisons with PBS labeling ( $p$ -values  $< 0.05$ ), while samples labeled in BIG showed significant differences for only a few targets, Supporting Information Fig. 2.

To investigate the origin of these skewed trends of fold changes, clinical data from CSF samples such as albumin and IgG quotients as well as IgG index available for the majority of samples were taken into account. To enable a direct comparison of assay results with these clinical parameters, intensities from the bead carrying IgG from nonimmunized rabbit was considered as background. A positive correlation between clinical data, especially IgG-quotient, and measured intensities from the rabbit IgG bead for samples labeled in PBS and heat treated at 72°C was observed (Spearman's  $\rho = 0.59$ ). This result suggests that the content of high-abundant proteins were related to the measured intensities in this direct labeling assay. In addition, measurements of total pro-

tein concentrations in all samples included in the study also implied a relation between background intensity and total protein amount (Spearman's  $\rho = 0.50$ ). For samples labeled in a buffer containing BSA and IgG, lower correlation to IgG-quotient and total protein concentration was observed ( $\rho = 0.26$  and  $-0.01$  respectively), Supporting Information Fig. 3.

A more global way to analyze the effect of varying levels of total protein concentrations is by utilization of the normalization factor calculated within the PQN method. This value, called the PQN denominator, is a sample-specific normalization factor that accounts for different dilution effects introduced due to sample handling [15] or fluctuations caused by circadian effects (as seen in urine) [14]. The distribution of these normalization factors within a sample group and subsequent comparison of distributions between sample groups implied that there was a general difference between the intensity levels of the samples in the different groups when samples were labeled in PBS. As shown in Supporting Information 4A and B, this holds true for both the assay with the undirected and the selected set of antibodies (both with  $p$ -value  $< 0.01$ ). Moreover, the PQN denominators were also positively correlated ( $\rho = 0.69$ ) between the two assays, indicating a sample specific consistency. However, when the samples were labeled in presence of BSA and IgG as buffer supplements, there was no longer a sample group related difference ( $p$ -value = 0.5), see Supporting Information Fig. 4C.

To investigate if the LOD of the assay was altered by the addition of BSA and IgG, a spike-in experiment was performed using a nonhuman protein CTB. A dilution series ranging from 12 ng/mL to 68 pg/mL CTB was prepared in 1% BSA, diluted plasma as well as a diluted CSF sample pool labeled in both PBS and BIG buffer. In CSF, the assays revealed LODs of approximately 1 ng/mL, regardless of which buffer the sample was labeled in. A similar result was obtained also for the plasma sample while a twofold lower LOD value was calculated for the buffer control, as shown in Fig. 3.

In summary, the total protein concentration affected the generated protein profiles globally. The supplementation of BSA and IgG prior labeling reduced this effect and enabled analysis that better reflected the sample content.

### 3.2 Protein profiling within MS

To evaluate the effect of heat treatment temperature on assay results, 65 CSF samples labeled in the presence of albumin and IgG were heated at either 56 or 72°C. To enable comparison of the two separate assays the data was median scaled and the 72/56°C ratio was calculated for the median signal for each of the 71 antibodies. As presented in Fig. 4, the effect of heating appeared to be antibody-dependent, similarly to what has previously been reported for plasma [17]. The technical coefficient of variation was comparable for the two temperatures with median values of 8 and 10%, respectively.

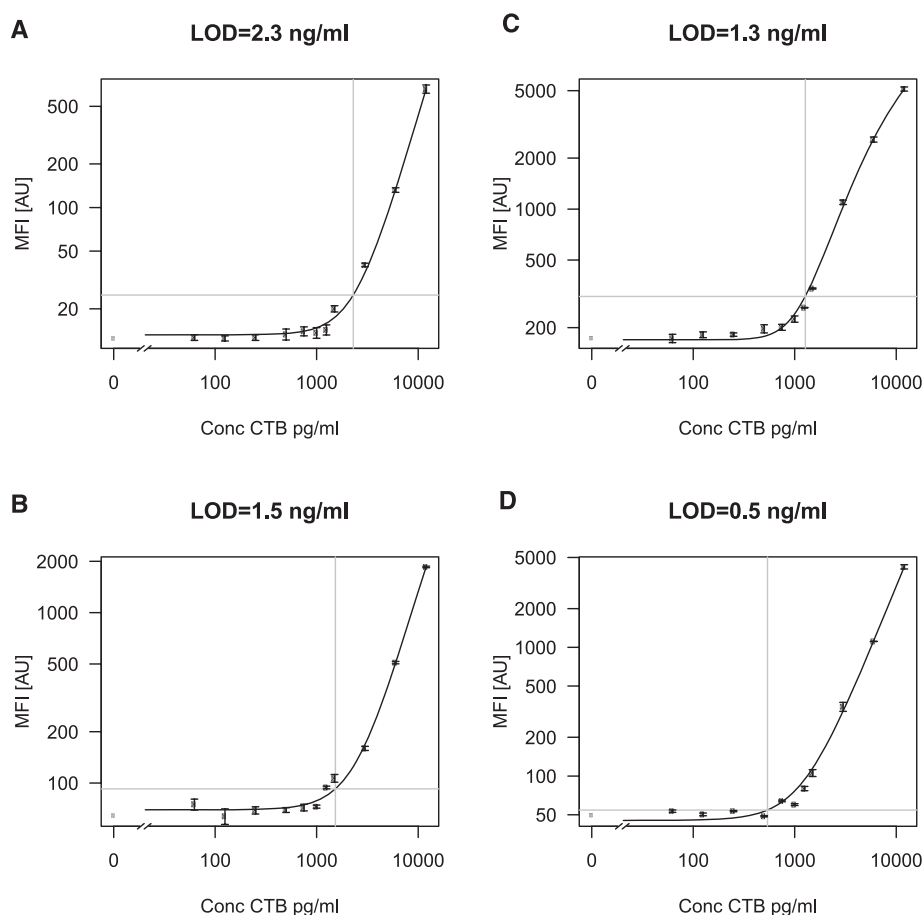


**Figure 3.** Relative LODs. CTB was spiked into samples before labeling in an eleven step dilution series with end CTB concentrations between 12 000 and 68 pg/mL. A nonspiked sample was used as a reference. As spike-in matrices, a CSF sample pool diluted in both (A) PBS and (B) BIG (BSA + IgG) buffer, (C) a plasma sample and (D) a buffer with 0.1% BSA in PBS were used. An apparent LOD was calculated from a five-parametric logistic model using the mean intensity plus three SDs of the triplicate reference sample. In this analysis, LODs in the biological samples were similar, around 1 ng/mL, while a twofold lower LOD was calculated for the BSA buffer.

In the analysis of this sample set, results from group comparisons revealed an interesting difference for the antibody targeting growth associated protein 43 (GAP43, HPA013392), with lower intensity levels determined in SPMS samples compared to both ONDs and RRMS patients. Similar trends were observed with samples treated at both temperatures although only significant with 72°C sample treatment ( $p$ -values 0.004 and 0.07, respectively).

Based on the findings from the initial study, the whole set of 339 CSF samples was profiled with heat treatment at both 56 and 72°C. The set of 71 proteins was reduced based on indications of differences found in the initial data and further complemented with additional targets from previous

indications in a separate plasma study (Byström et al, unpublished), making up a total of 43 target proteins for the verification stage. In this bead array the majority of proteins were represented with two to six antibodies in order to increase the confidence of potential findings, see Supporting Information Table 1 for a full list of antibodies included for verification. To assess the concordance between antibodies targeting the same protein, correlations between antibody pairs were investigated and the three antibodies included targeting GAP43 showed a very high correlation at 56°C ( $\rho > 0.95$ ) but not at 72°C ( $\rho < 0.41$ ), see Fig. 5A. Other antibody pairs however showed high correlation at both temperatures as exemplified by three out of the in total five included

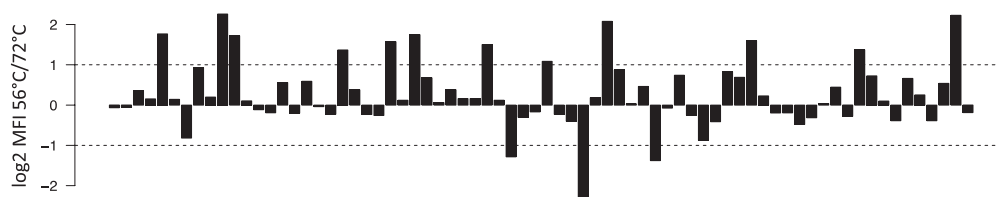


**Figure 4.** Analysis of heat treatment on CSF protein profiles. The CSF samples used for assay development were analyzed with the heat treatment temperatures of 56 or 72°C. The ratios of median intensities per antibody (y-axis) indicated that the heat treatment induced effect was highly antibody dependent.

antibodies toward serpin peptidase inhibitor, clade A, member 1 (SERPINA1) in Fig. 5B. To enable a comparison of the correlation coefficients obtained for the paired antibodies, pairwise coefficients  $\rho$  values of all antibodies included in the verification were calculated regardless of the corresponding target protein. The distribution of these coefficients was centered at zero (Supporting Information Fig. 5).

In concordance with results from the 65 samples initially analyzed, signal intensities for GAP43 were lower in SPMS samples compared to RRMS and OND samples also in the analysis of all 339 CSF samples treated at 56°C ( $p < 0.003$  for all the three antibodies). In this larger sample material, the additional sample groups iOND and CIS were represented and also these groups showed higher levels of GAP43 compared to the SPMS group ( $p < 0.001$ ), see Fig. 6A and B. When heat-treating the samples at 72°C, not only the correlation but also the discrimination ability was altered for the anti-GAP43 antibodies, with only one of them (HPA015600) showing a similar trend of lower intensities in the SPMS group. The classification performances of the three antibodies are visualized as ROC curves (Fig. 6C) with the false positive rate plotted against the true positive rate for the SPMS and RRMS groups. As indicated by the area under curve (AUC), only one antibody was able to discriminate between these sample groups equally well at 72°C as at 56°C. To verify the results obtained through direct labeling a bead-based sandwich assay was developed using parallel analysis with all three antibodies toward GAP43 as capture and HPA015600 as detection reagent. The discrimination of the two patient groups with GAP43 sandwich assays were supportive and correlated well (Spearman's  $\rho > 0.93$ ), see Supporting Information Fig. 6. To investigate whether the differences in GAP43 detection could also be observed in blood, plasma samples from the same 339 patients were analyzed using the same bead array. In these samples, no significant differences were observed between the RRMS and SPMS groups for either of the antibodies targeting GAP43, see Fig. 6D. Another protein with differences between sample groups and with a high antibody correlation was SERPINA3. Interestingly, it was found to be present at higher levels not only in all MS subtypes, but also in the inflammatory OND when analyzed at both temperatures, see Fig. 7.

## 4 Discussion

Here, we have described the adaptation of a protein profiling protocol for high-throughput and antibody-based exploration of CSF. By supplementing the buffer used during labeling of samples, we could reduce the effect of differences in total protein concentration without notably compromising the performance of the assay. With this setup, a multiplexed analysis of up to 384 protein targets in CSF became feasible and a simultaneous analysis of many samples could be achieved.

The protein composition of CSF is similar to that of blood since the majority of the protein content is derived from the

filtration of blood through the blood CSF barrier [18]. The two most abundant proteins, albumin and IgG are present at varying levels in CSF depending on the blood-brain barrier integrity and the level of inflammation in the CNS. In active brain lesions of MS patients, there is strong evidence for both neurodegeneration and inflammation [3]. Furthermore, the presence of oligoclonal IgG bands in CSF derived from an intrathecal production of immunoglobulins in about 90% of these patients also suggest an ongoing inflammatory process [5]. Even though not all neurological diseases have such clear signs of inflammation, this is by no means unique to the disease of MS. In addition, disruption of the blood-brain barrier has been suggested to play a role in MS by enabling the infiltration of immune cells into the CNS. This is however also observed for other brain-related diseases (as reviewed in [19]) and thereby, similarly to inflammation, it is not exclusive for MS. Taking this together, the composition of high abundant proteins can vary in CSF depending on disease status and such proteins can potentially cause biases in group comparison studies.

For high-throughput screening approaches it is desirable not to adjust the input volume for each sample individually. This preparative procedure would consume additional resources and samples for assessing their overall protein concentrations. It would also primarily affect the levels of other proteins than albumin and IgG, thus potentially masking disease-related differences. It became apparent that for CSF, the difference in total protein concentration with respect to the abundant proteins was an important factor systematically influencing the assay results when all samples were diluted with the same dilution factor in PBS. Because of the relatively large number of samples and analytes, we could observe these systematic effects that otherwise might have been less obvious and perhaps misinterpreted as biologically important.

Many technologies employ or require depletion of high abundant proteins in order to detect the lower abundant portion of the proteome [20]. This pre-treatment of samples inevitably entails the risk of also removing desired protein targets, which therefore might change the original protein content of a sample [21]. Here, as an alternative and new approach, instead of removing proteins, an excess of BSA and IgG was added to all samples prior sample biotinylation in order to adjust for the relative differences in total protein amounts between samples and to create a more similar ratio of protein content and biotin in all samples. The results indicate that this general adjustment successfully reduced systematic differences, thereby providing an improved representation of the biological content and composition of samples. Importantly, the ability to detect the desired protein targets with the developed assay appeared not to be affected by neither the presence nor the amount of high abundant proteins in both samples and buffers when taking spike-in experiments into account. Estimated LODs in the complex sample environment was in the low ng/mL to high pg/mL range, in line with both previous findings based on plasma analysis [17] and as suggested for single-binder assays [22].

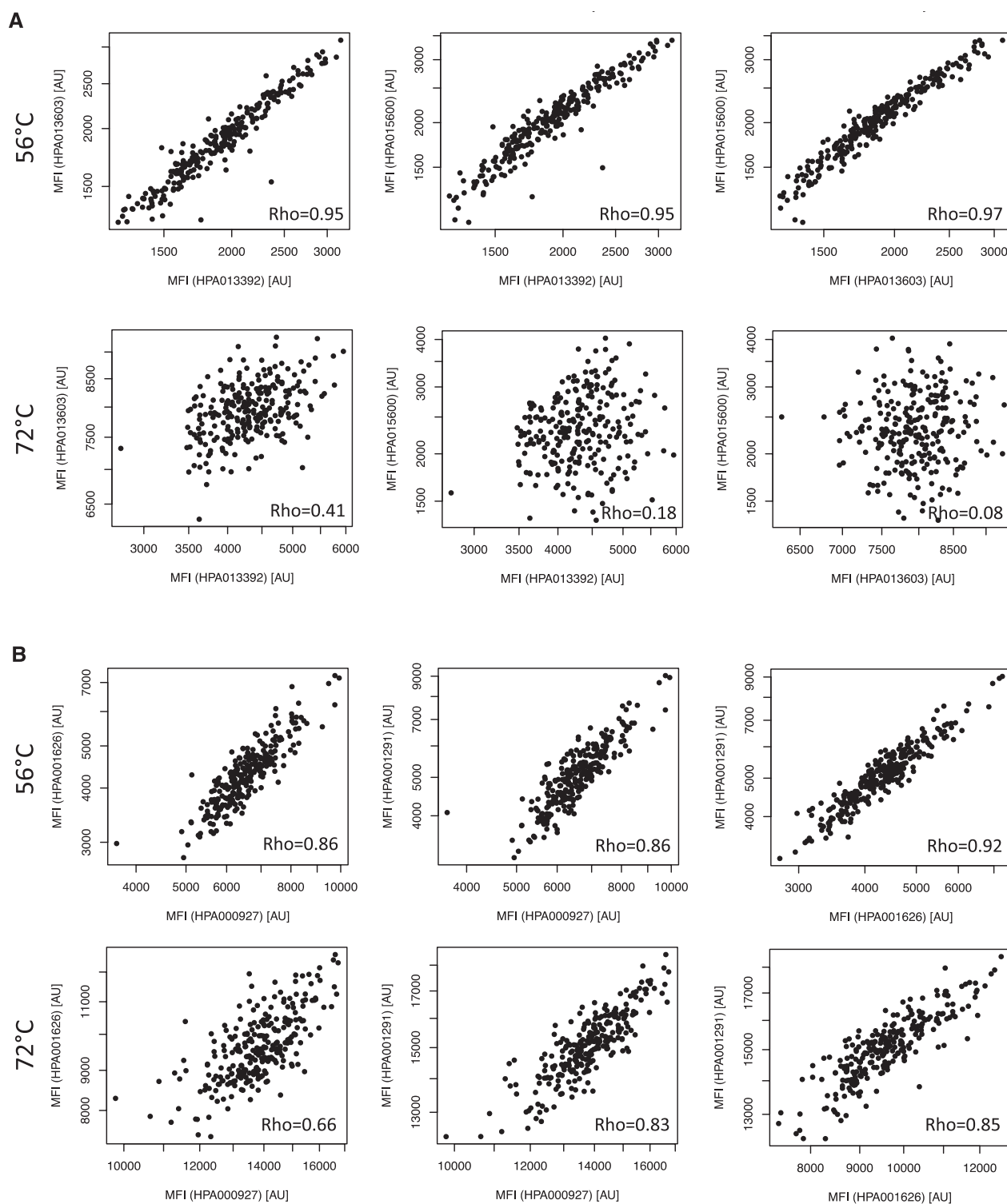


**Figure 5.** Correlation between paired antibodies. Results from multiple antibodies targeting the same protein were compared through correlation of intensity profiles. (A) For the protein target GAP43, three antibodies were available out of which two were generated against the same protein fragment (HPA013392 and HPA013603) while the third antibody was generated toward a different part of the protein (HPA015600). All three antibodies showed a Spearman correlation of  $>0.95$  with 56°C heat treatment (upper panel) while the positive correlation was lost when samples were treated at 72°C (lower panel). (B) Correlations of three antibodies targeting SERPINA1 are visualized as an example of interactions that were less affected by the sample treatment temperature (upper versus lower panel).

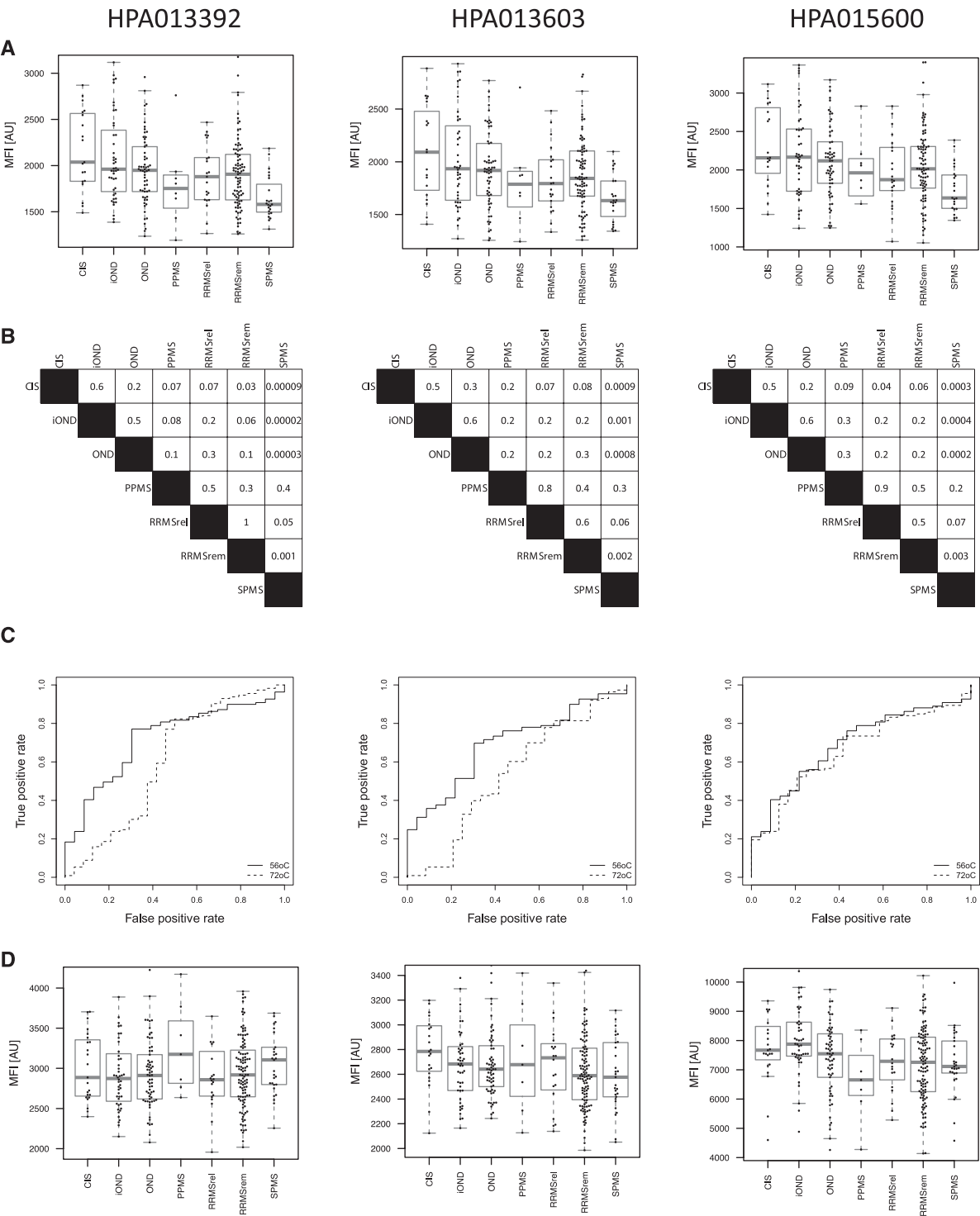


**Figure 6.** Protein levels of GAP43 in CSF and plasma. (A) All three antibodies targeting GAP43 revealed similar profiles for the sample groups when treated at 56°C with significantly lower intensities in the SPMS samples compared to all other groups except primary progressive MS and RRMSrel. (B) The resulting *p*-values from two-group comparisons using Wilcoxon rank-sum test are summarized with significant differences highlighted in bold. (C) To evaluate the performance for classification of the SPMS and RRMS samples based on these antibodies, ROC curves were generated for both sample treatment temperatures. One of the antibodies (HPA015600, rightmost panel) showed similar discriminatory power while the other two appeared to be more heat sensitive with reduced differences between groups at higher temperature. HPA013392 (AUC 56°C = 0.72, AUC 72°C = 0.60); HPA013603 (AUC 56°C = 0.69, AUC 72°C = 0.54); HPA015600 (AUC 56°C = 0.69, AUC 72°C = 0.66). (D) In addition to the CSF profiling, plasma samples from the same individuals were analyzed with the GAP43 antibodies. The lower intensities for the SPMS group were not observed in this sample material.

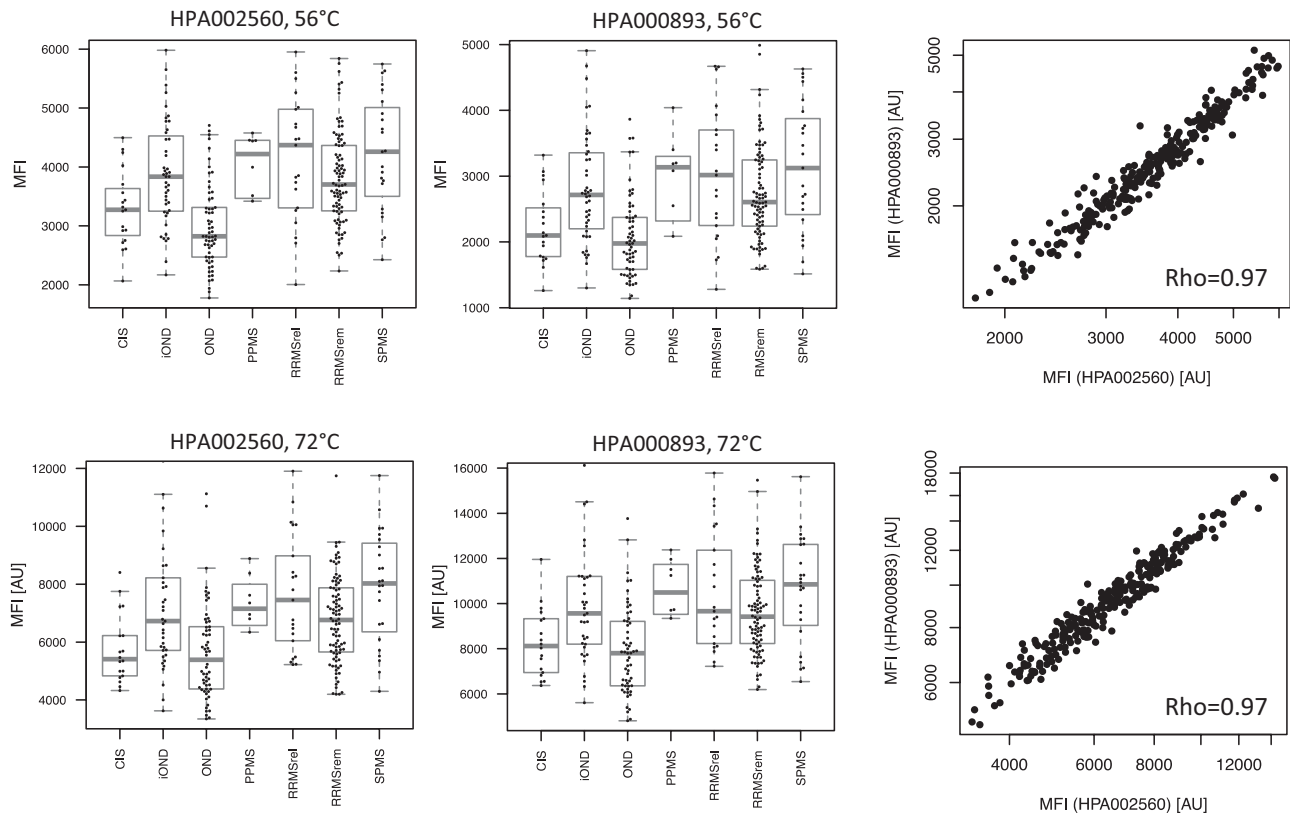


**Figure 7.** Protein levels of SERPINA3 in CSF. The two antibodies targeting SERPINA3 could detect a significant difference between the CIS and OND groups as compared to iOND and all MS subtypes ( $p < 0.0001$ ). The antibodies performed similarly at both temperatures and showed a positive sample correlation with Spearman's rho  $> 0.97$ .

These values are however highly dependent on the specific affinity reagent and varies for different targets.

The assay protocol developed here was based on an earlier described protocol for plasma profiling, in which heat treatment of samples prior to bead array incubation facilitated epitope retrieval [12, 17]. Here, we also investigated temperature as a modulator of CSF samples and in accordance to previous experience, antibody–protein interactions were affected differently for different targets. Possible explanations include both the nature of the epitope (conformational or continuous), the susceptibility of the protein to unfold at high temperatures as well as aggregation and degradation processes.

One of the interesting findings in this study was the lower level of GAP43 in SPMS samples. This protein has an important function in the process of axonal growth [23] and decreased levels of GAP43 have previously been observed in demyelinated brain lesions whereas surrounding areas and remyelinated lesions have showed increased levels, indicating a role in neurological regeneration [24]. Other studies have concluded that increased axonal damage is a marker of worse prognosis and of conversion into a progressive state of the disease, which is in line with our finding of decreased levels of GAP43 in SPMS patients. In the verification assay,

all three antibodies targeting this protein showed significant differences for samples treated at 56°C, in contrast to 72°C heat treatment which resulted in one (HPA015600) out of three antibodies significantly discriminating SPMS. This was contradictory to the initial screening where the one antibody included showed similar trends although a better group separation for the higher temperature. This altered outcome of group comparisons appears to be related to the additional individuals analyzed in the verification stage. The antibody with concordant profiles for the two temperatures (HPA015600) was generated against a different protein fragment than the other two antibodies that were generated toward the same protein sequence immunized in two animals in parallel. This might explain why their ability to detect the target protein is differently sensitive to the heat treatment temperature. The difference in levels of GAP43 observed between RRMS and SPMS in CSF samples could this far not be detected in plasma obtained from the same individuals. This finding highlights that analysis of proximal fluids are an important complement to the more easily accessible sample materials such as blood-derived samples.

Besides target abundance, the ability for an antibody to bind its intended target in the described assay is dependent on the affinity of the antibody, the accessibility of epitopes

considering secondary or tertiary structures as well as PTMs. One example of a protein with many well functioning antibodies was SERPINA1. The five antibodies targeting this protein were generated against different protein fragments spanning over the whole protein sequence and they correlated well to each other at both temperatures, see Supporting Information Fig. 7.

The second protein found to show a difference between the sample groups was SERPINA3. For this target, both antibodies included revealed higher relative intensities in all MS subtypes as well as in the control group of patients diagnosed with OND with signs of inflammation. This is an acute phase protein that is released early in response to inflammation and it has been extensively studied in the context of Alzheimer's disease [25]. It has also previously been reported to be present at higher levels in CSF from MS patients compared to non-inflammatory controls [26]. Our results suggest that, even though the levels are higher in MS patients compared to OND and CIS, the increased levels observed in the inflammatory controls could reflect a sign of general inflammation rather than a disease specific process.

In conclusion, the presented development opens up possibilities to perform broad scaled antibody-based protein profiling of CSF. Using the single binder assay setup and a direct labeling approach, protein supplementation of CSF prior labeling reduces the effect of differential total protein concentration. It thereby enables the exploration of the CSF protein content independently of albumin and IgG levels for the search of new candidate molecules needed to aid diagnostics and patient care. Furthermore, the level of the protein GAP43 in CSF has been identified as a component of potential importance within MS.

*We like to thank the entire staff of the Human Protein Atlas for their efforts. This study was funded by grants from Science for Life Laboratory Stockholm, ProNova VINN Excellence Centre for Protein Technology (VINNOVA, Swedish Governmental Agency for Innovation Systems) and the Knut and Alice Wallenberg Foundation.*

*The authors have declared no conflict of interest.*

## 5 References

- [1] Ayoglu, B., Häggmark, A., Neiman, M., Igel, U. et al., Systematic antibody and antigen-based proteomic profiling with microarrays. *Expert Rev. Mol. Diagnost.* 2011, 11, 219–234.
- [2] Hirtz, D., Thurman, D. J., Gwinn-Hardy, K., Mohamed, M. et al., How common are the "common" neurologic disorders? *Neurology* 2007, 68, 326–337.
- [3] Frischer, J. M., Bramow, S., Dal-Bianco, A., Lucchinetti, C. F. et al., The relation between inflammation and neurodegeneration in multiple sclerosis brains. *Brain* 2009, 132, 1175–1189.
- [4] Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M. et al., Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann. Neurol.* 2011, 69, 292–302.
- [5] Compston, A., Coles, A., Multiple sclerosis. *Lancet* 2008, 372, 1502–1517.
- [6] Dumont, D., Noben, J. P., Raus, J., Stinissen, P., Robben, J., Proteomic analysis of cerebrospinal fluid from multiple sclerosis patients. *Proteomics* 2004, 4, 2117–2124.
- [7] Hammack, B. N., Fung, K. Y., Hunsucker, S. W., Duncan, M. W. et al., Proteomic analysis of multiple sclerosis cerebrospinal fluid. *Mult. Scler.* 2004, 10, 245–260.
- [8] Komori, M., Matsuyama, Y., Nirasawa, T., Thiele, H. et al., Proteomic pattern analysis discriminates among multiple sclerosis-related disorders. *Ann. Neurol.* 2012, 71, 614–623.
- [9] Light, M., Minor, K. H., Dewitt, P., Jasper, K. H., Davies, S. J., Multiplex array proteomics detects increased MMP-8 in CSF after spinal cord injury. *J. Neuroinflammation* 2012, 9, 122.
- [10] Poetz, O., Henzler, T., Hartmann, M., Kazmaier, C. et al., Sequential multiplex analyte capturing for phosphoprotein profiling. *Mol. Cell. Proteomics* 2010, 9, 2474–2481.
- [11] Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E. et al., Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* 2010, 28, 1248–1250.
- [12] Schwenk, J. M., Gry, M., Rimini, R., Uhlen, M., Nilsson, P., Antibody suspension bead arrays within serum proteomics. *J. Proteome Res.* 2008, 7, 3168–3179.
- [13] Ihaka, R., Gentleman, R., R: A language for data analysis and graphics. *J. Comput. Graph. Stat.* 1996, 5, 299–314.
- [14] Dieterle, F., Ross, A., Schlotterbeck, G., Senn, H., Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal. Chem.* 2006, 78, 4281–4290.
- [15] Kato, B. S., Nicholson, G., Neiman, M., Rantalainen, M. et al., Variance decomposition of protein profiles from antibody arrays using a longitudinal twin model. *Proteome Sci.* 2011, 9, 73.
- [16] Hubert, M., Rousseeuw, P. J., Branden, K. V., ROBPCA: A new approach to robust principal component analysis. *Technometrics* 2005, 47, 64–79.
- [17] Schwenk, J. M., Igel, U., Neiman, M., Langen, H. et al., Toward next generation plasma profiling via heat-induced epitope retrieval and array-based assays. *Mol. Cell. Proteomics* 2010, 9, 2497–2507.
- [18] Regeniter, A., Kuhle, J., Mehling, M., Moller, H. et al., A modern approach to CSF analysis: pathophysiology, clinical application, proof of concept and laboratory reporting. *Clin. Neurol. Neurosurg.* 2009, 111, 313–318.
- [19] Rosenberg, G. A., Neurological diseases in relation to the blood-brain barrier. *J. Cereb. Blood. Flow. Metab.* 2012, 32, 1139–1151.
- [20] Thouvenot, E., Urbach, S., Dantec, C., Poncet, J. et al., Enhanced detection of CNS cell secretome in plasma protein-depleted cerebrospinal fluid. *J. Proteome Res.* 2008, 7, 4409–4421.

- [21] Granger, J., Siddiqui, J., Copeland, S., Remick, D., Albumin depletion of human plasma also removes low abundance proteins including the cytokines. *Proteomics* 2005, 5, 4713–4718.
- [22] Landegren, U., Vanelid, J., Hammond, M., Nong, R. Y. et al., Opportunities for sensitive plasma proteome analysis. *Anal. Chem.* 2012, 84, 1824–1830.
- [23] Oestreicher, A. B., De Graan, P. N., Gispen, W. H., Verhaagen, J., Schrama, L. H., B-50, the growth associated protein-43: modulation of cell morphology and communication in the nervous system. *Prog. Neurobiol.* 1997, 53, 627–686.
- [24] Teunissen, C. E., Dijkstra, C. D., Jasperse, B., Barkhof, F. et al., Growth-associated protein 43 in lesions and cerebrospinal fluid in multiple sclerosis. *Neuropathol. Appl. Neurobiol.* 2006, 32, 318–331.
- [25] Licastro, F., Candore, G., Lio, D., Porcellini, E. et al., Innate immunity and inflammation in ageing: a key for understanding age-related diseases. *Immun. Ageing* 2005, 2, 8.
- [26] Ottervald, J., Franzen, B., Nilsson, K., Andersson, L. I. et al., Multiple sclerosis: Identification and clinical evaluation of novel CSF biomarkers. *J. Proteomics* 2010, 73, 1117–1132.

Supplementary table 1. Antibodies included in the verification stage

| Target protein | Antibody ID  | Ensemble ID     | Target protein | Antibody ID | Ensemble ID     |
|----------------|--------------|-----------------|----------------|-------------|-----------------|
| ALPK2          | HPA027976    | ENSG00000198796 | MANSC1         | HPA007541   | ENSG00000111261 |
| ALPK2          | HPA029801    | ENSG00000198796 | MAP2           | HPA012828   | ENSG00000078018 |
| ALPK2          | HPA028084    | ENSG00000198796 | MAP2           | HPA012558   | ENSG00000078018 |
| ALPK2          | HPA027377    | ENSG00000198796 | MAP2           | HPA008273   | ENSG00000078018 |
| ANXA1          | HPA011271    | ENSG00000135046 | METTL14        | HPA038002   | ENSG00000145388 |
| ANXA1          | HPA011918    | ENSG00000135046 | METTL14        | HPA038001   | ENSG00000145388 |
| ANXA1          | HPA011272    | ENSG00000135046 | MMP25          | HPA036376   | ENSG00000008516 |
| APEX1          | HPA002564    | ENSG00000100823 | MMP25          | HPA055640   | ENSG00000008516 |
| APEX1          | HPA000956    | ENSG00000100823 | MMP25          | HPA049236   | ENSG00000008516 |
| APEX1          | HPA001430    | ENSG00000100823 | MOG            | HPA021873   | ENSG00000204655 |
| ARG2           | HPA000663    | ENSG00000081181 | MTPN           | HPA019735   | ENSG00000105887 |
| ATP5I          | HPA035010    | ENSG00000169020 | NRCAM          | HPA012606   | ENSG00000091129 |
| BCL6           | HPA004899    | ENSG00000113916 | NRG1           | HPA010964   | ENSG00000157168 |
| BCL6           | HPA004791    | ENSG00000113916 | PLAU           | HPA008719   | ENSG00000122861 |
| BCL6           | HPA050645    | ENSG00000113916 | PLAU           | HPA008738   | ENSG00000122861 |
| C3             | HPA003563    | ENSG00000125730 | PLG            | HPA048823   | ENSG00000122194 |
| C3             | HPA020432    | ENSG00000125730 | PLG            | HPA021602   | ENSG00000122194 |
| C3a            | AF3677*      | ENSG00000125731 | PLG            | HPA053770   | ENSG00000122194 |
| CD14           | HPA002127    | ENSG00000170458 | PLG            | HPA043178   | ENSG00000122194 |
| CD14           | HPA002035    | ENSG00000170458 | PLG            | HPA055966   | ENSG00000122194 |
| CD14           | HPA001887    | ENSG00000170458 | PSMC2          | HPA049621   | ENSG00000161057 |
| CFI            | HPA001143    | ENSG00000205403 | PSMC2          | HPA019238   | ENSG00000161057 |
| CFI            | HPA024061    | ENSG00000205403 | RNF141         | HPA018133   | ENSG00000110315 |
| CKB            | HPA001254    | ENSG00000166165 | RNF141         | HPA057139   | ENSG00000110315 |
| CNDP1          | HPA016933    | ENSG00000150656 | RNF213         | HPA026790   | ENSG00000173821 |
| CNDP1          | HPA008933    | ENSG00000150656 | RNF213         | HPA008347   | ENSG00000173821 |
| CP             | HPA001834    | ENSG00000047457 | RNF213         | HPA003347   | ENSG00000173821 |
| CP             | HPA002737    | ENSG00000047457 | S100A8         | HPA024372   | ENSG00000143546 |
| DNMT3B         | HPA001595    | ENSG00000088305 | SERPINA1       | HPA000927   | ENSG00000197249 |
| FBLN1          | HPA001612    | ENSG00000077942 | SERPINA1       | HPA001626   | ENSG00000197249 |
| FBLN1          | HPA001642    | ENSG00000077942 | SERPINA1       | HPA001291   | ENSG00000197249 |
| FBLN1          | HPA001613    | ENSG00000077942 | SERPINA1       | HPA001292   | ENSG00000197249 |
| FBLN1          | HPA001643    | ENSG00000077942 | SERPINA1       | HPA000749   | ENSG00000197249 |
| GAP43          | HPA015600    | ENSG00000172020 | SERPINA3       | HPA000893   | ENSG00000196136 |
| GAP43          | HPA013392    | ENSG00000172020 | SERPINA3       | HPA002560   | ENSG00000196136 |
| GAP43          | HPA013603    | ENSG00000172020 | SERPINA4       | HPA002869   | ENSG00000100665 |
| GCH1           | HPA028612    | ENSG00000131979 | SERPINA4       | HPA003607   | ENSG00000100665 |
| GCH1           | HPA053990    | ENSG00000131979 | SH3BGR13       | HPA030848   | ENSG00000142669 |
| ICT1           | HPA003634    | ENSG00000167862 | SLC30A7        | HPA018034   | ENSG00000162695 |
| IL7            | HPA029567    | ENSG00000104432 | SOD1           | HPA001401   | ENSG00000142168 |
| IL7            | HPA042855    | ENSG00000104432 | SPP1           | HPA005562   | ENSG00000118785 |
| IL7            | HPA019590    | ENSG00000104432 | SPP1           | HPA027541   | ENSG00000118785 |
| IRF8           | HPA002531    | ENSG00000140968 | SPP1           | HPA027540   | ENSG00000118785 |
| IRF8           | HPA002267    | ENSG00000140968 | TJP1           | HPA001636   | ENSG00000104067 |
| IRF8           | 18977-1-AP** | ENSG00000140969 | TJP1           | HPA001637   | ENSG00000104067 |
| IRF8           | AF5117**     | ENSG00000140970 | TJP2           | HPA001813   | ENSG00000119139 |
| IRF8           | 39-8800***   | ENSG00000140971 | TJP2           | HPA005934   | ENSG00000119139 |
| IRF8           | AP2830C****  | ENSG00000140972 | ZFP36L1        | HPA000786   | ENSG00000185650 |
| MANSC1         | HPA007540    | ENSG00000111261 | ZFP36L1        | HPA001301   | ENSG00000185650 |
| MANSC1         | HPA007955    | ENSG00000111261 | ZFP36L1        | HPA035423   | ENSG00000185650 |
| MANSC1         | HPA007956    | ENSG00000111261 |                |             |                 |

Antibodies with ID HPAxxxxxx were provided by the Human Protein Atlas/Atlas Antibodies

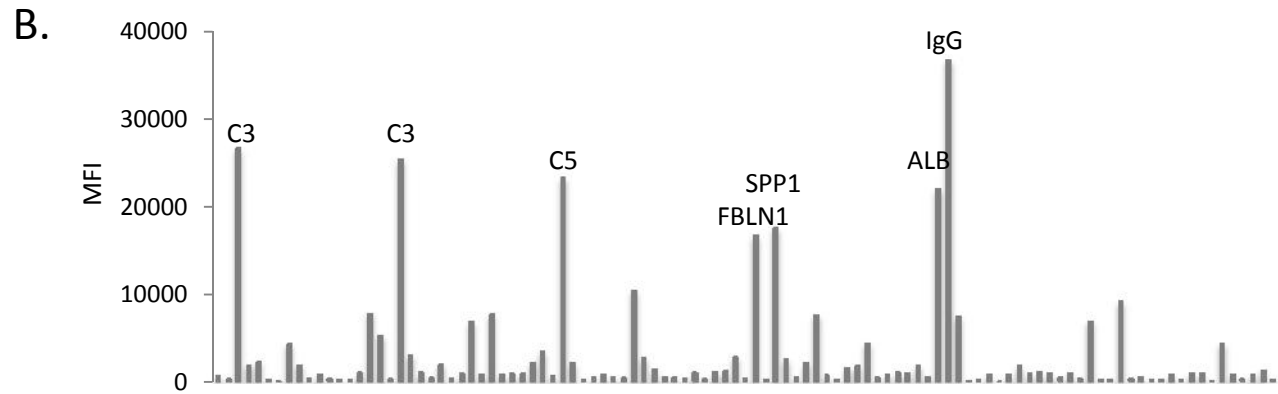
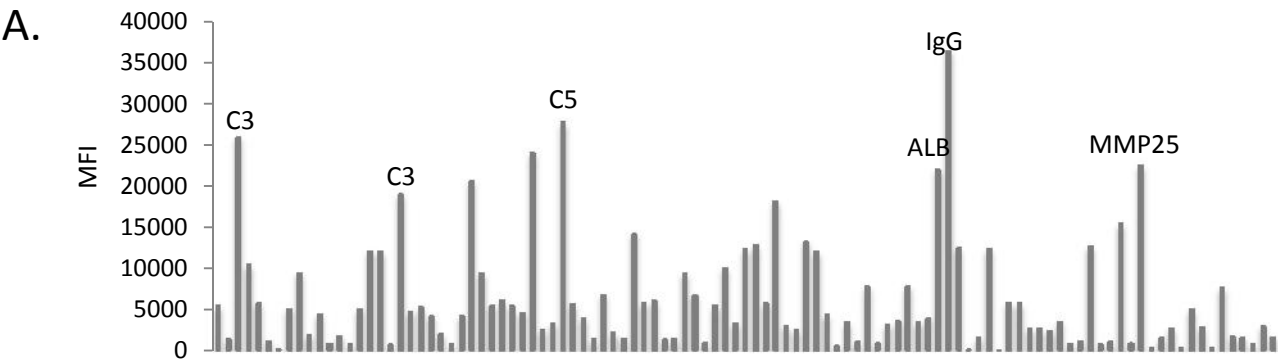
\* R&D Systems

\*\* ProteinTech

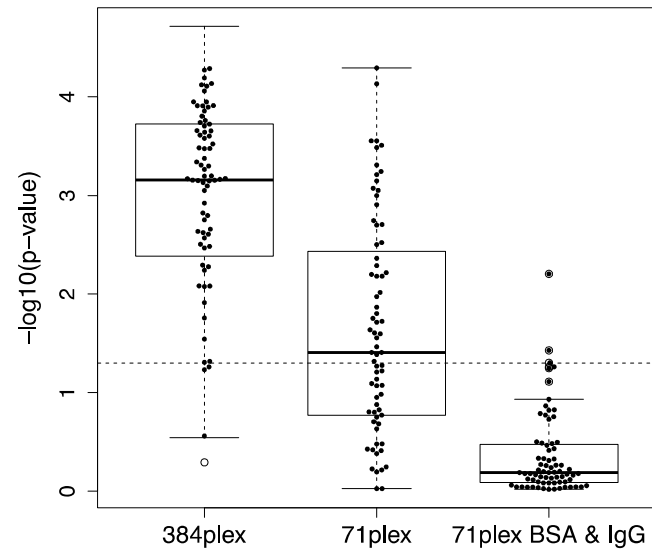
\*\*\* Zymed

\*\*\*\* Abgent

Supplementary figure 1. Representative sample profiles of plasma and CSF.

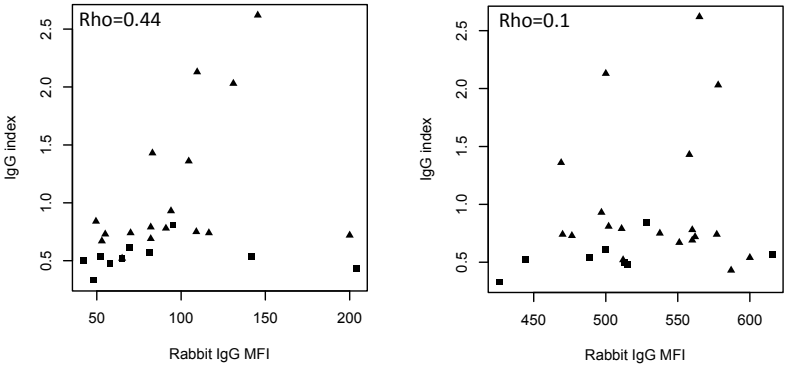


Supplementary figure 2. Significance levels of differences in raw data.

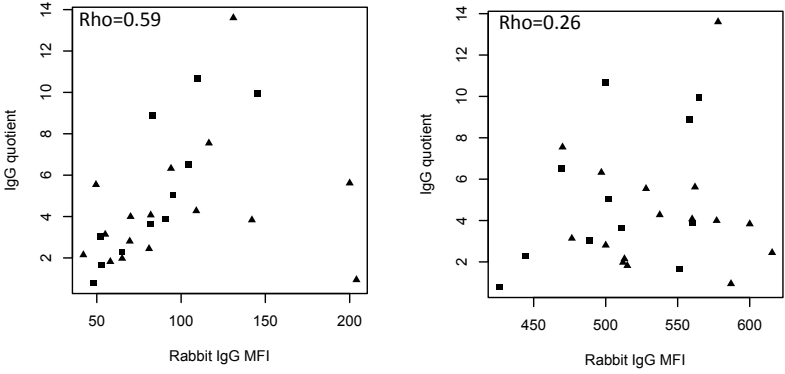


Supplementary figure 3. Correlation of measured intensities to clinical sample information.

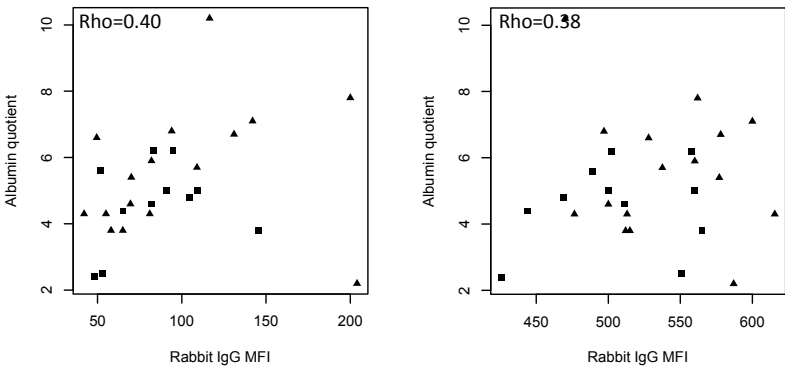
A.



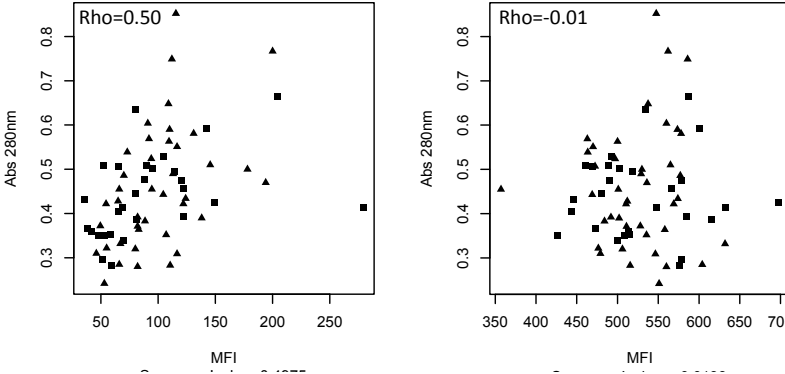
B.



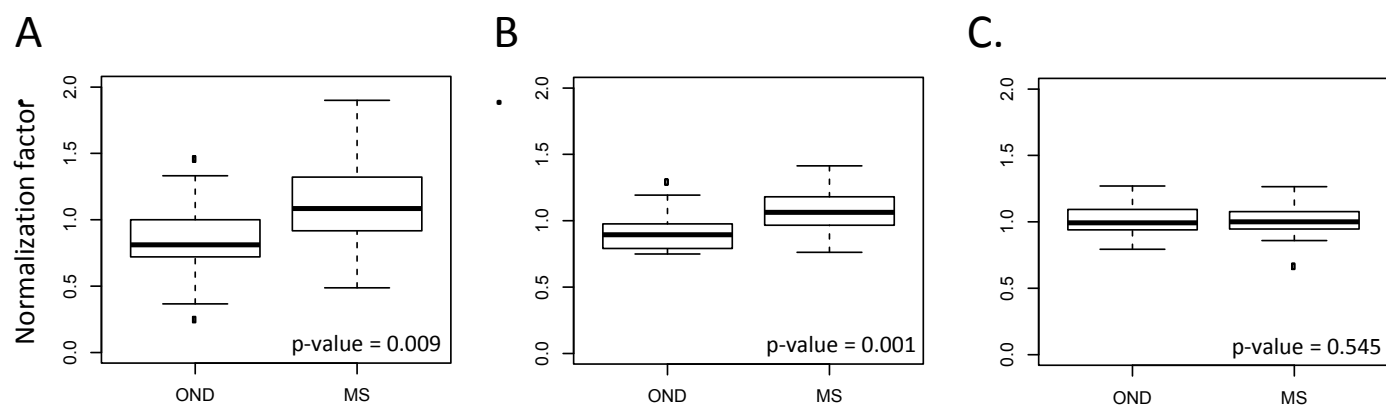
C.



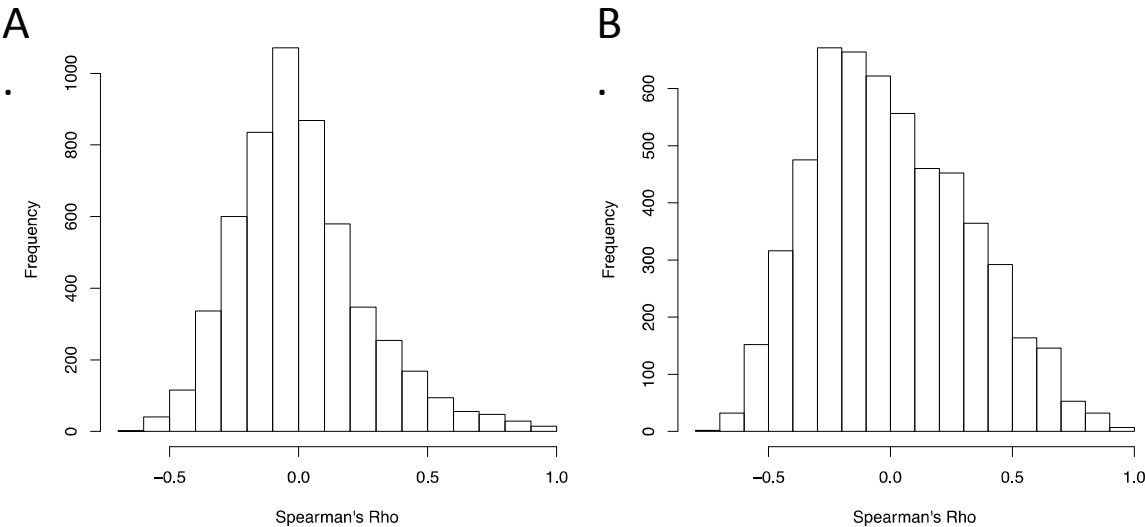
D.



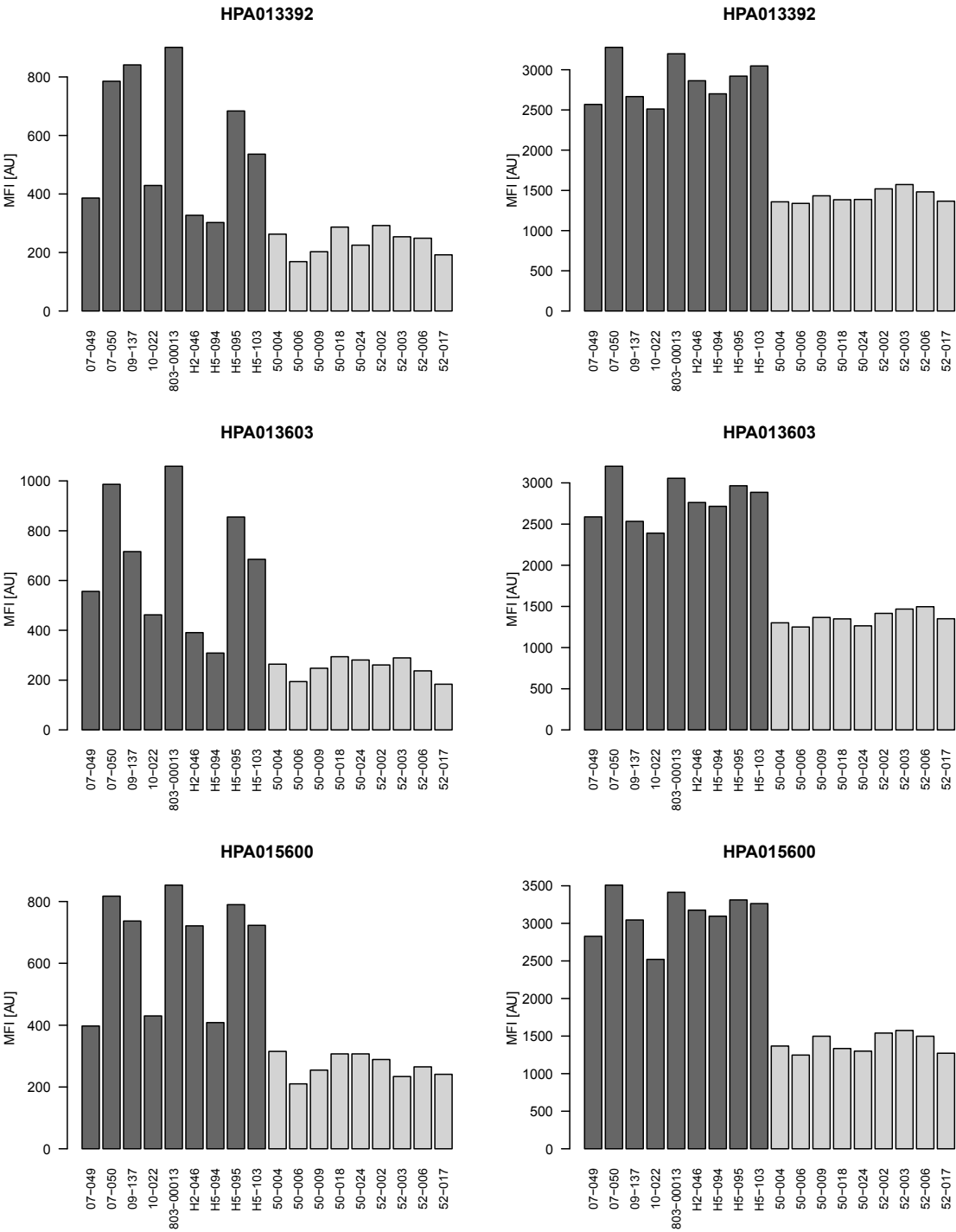
Supplementary figure 4. Normalization factors per sample group.



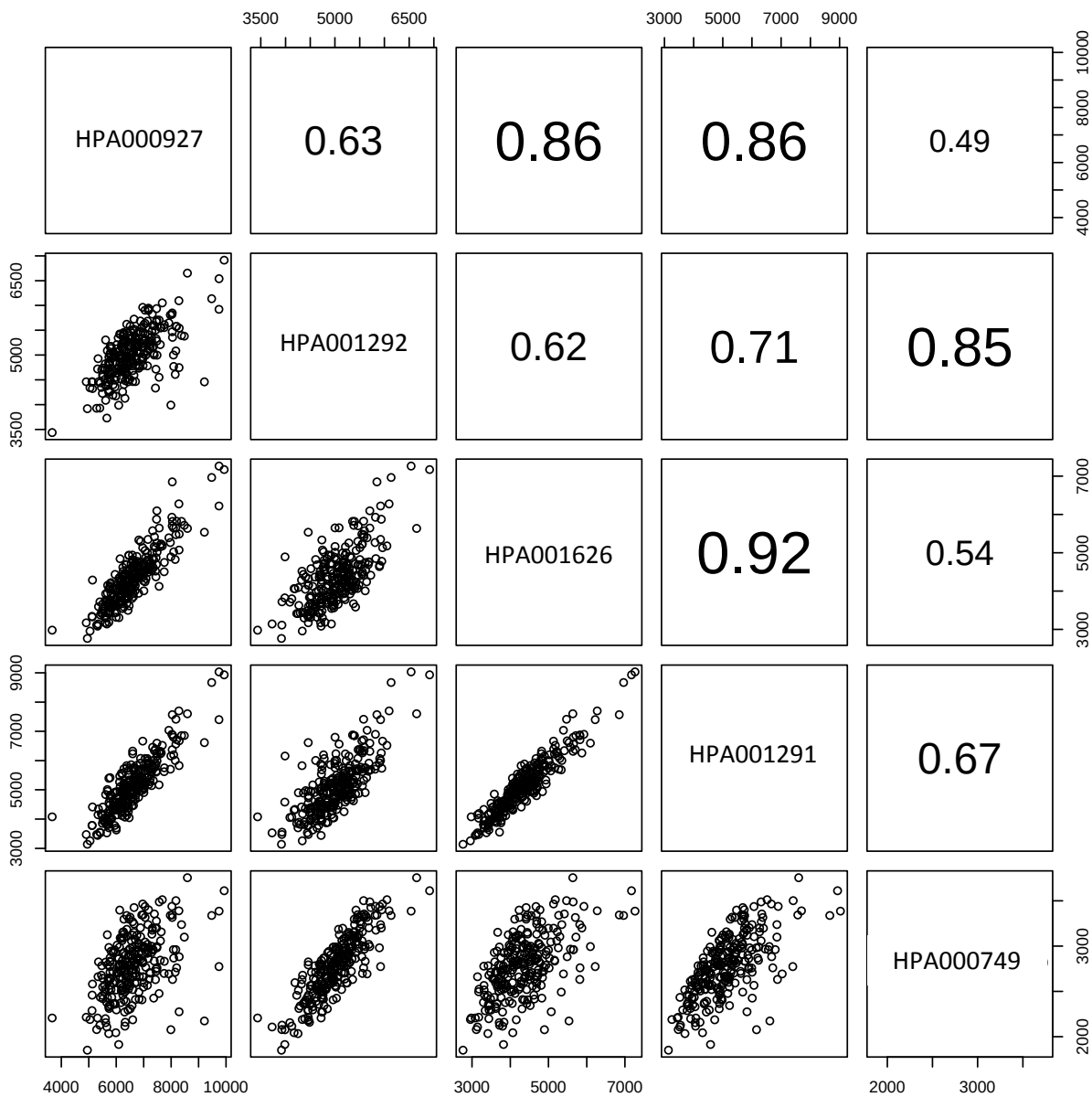
Supplementary figure 5. Distribution of correlation coefficients of all antibodies



Supplementary figure 6. Results from GAP43 sandwich assay



Supplementary figure 7. Correlations of paired antibodies targeting SERPINA1



## Supplementary information

**Supplementary table 1. Antibodies included in the verification stage.** A representation of the 101 antibodies included for the analysis of 339 CSF samples.

**Supplementary figure 1. Representative sample profiles of plasma and CSF.** Examples of representative plasma and CSF (labelled in BIG buffer) sample profiles with intensity values (y-axis) shown for all included antibodies in the 101-plex array targeting 43 proteins (x-axis). Protein targets with high signal intensities are indicated with names. The profiles were overall similar with albumin and IgG, known high abundant proteins, among the highest intensities. Note that intensities do not represent absolute quantification and that the measured signal depends on the analyte abundance and epitope availability as well as antibody affinity.

**Supplementary figure 2. Significance levels of differences in raw data.** For the set of 384 randomly selected antibodies as well as the 71-plex bead array, a skewed appearance of differences was found for the raw data of samples labelled in PBS while the same was not observed for samples labelled in the presence of excess BSA and IgG. Calculation of p-values for MS vs control samples for these three assays revealed differences at high significance level for samples labelled in PBS while samples labelled in BIG showed only a few significant differences. The dotted line represents a p-value of 0.05.

**Supplementary figure 3. Correlation of measured intensities to clinical sample information and total protein concentration.** Clinical data on IgG index, IgG quotient and albumin quotient was available for the majority of samples included in the study. This information was correlated to the intensities generated by the bead with immobilized non-specific rabbit IgG to relate assay results to sample composition. The correlations were visualized for the initial set of samples labelled in PBS (left panel) or BIG (right panel) with information on IgG index (A), IgG quotient (B), albumin quotient (C) and the result from the absorbance measurement representing the total protein concentration (D). The result implied that the addition of an excess of albumin and IgG to a sample prior labelling brings the samples to a more equal level in terms of albumin and IgG content and helps to reduce the effect of uneven sample compositions on the assay results. MS patients are depicted as triangles and controls as squares.

**Supplementary figure 4. Normalisation factors per sample group.** Sample normalisation factors were visualised to assess differences between disease related intensity levels. For both of the assays with an undirected selection of antibodies (A) and potential MS-related targets (B) and samples labelled in PBS buffer, the normalisation factors were higher in the MS subtypes compared to the control group. This reflects the collectively higher intensity level of the protein profile in CSF from MS patients compared to the control group ( $p < 0.01$ ). Adding an excess of two highly abundant proteins to the sample buffer prior labelling, a difference was no longer observed (C).

**Supplementary figure 5. Distribution of correlation coefficients for all antibodies.** As a quality control and to enable comparison of the correlation coefficients of paired antibodies, all

antibodies included in analysis of the 339 CSF samples were correlated, regardless of the specific protein target. The distributions are centred around zero for both 56 (A) and 72°C (B), indicating that there was no systematic correlation of antibodies included in the analysis.

**Supplementary figure 6. Triplex sandwich assay for GAP43.** Sandwich assays targeting GAP43 were setup to confirm the results from the direct labelling procedure. Three different capture antibodies HPA013392 (A), HPA013603 (B) and HPA015600 (C) together with HPA015600 as detection antibody composed three well-performing sandwich pairs that all supported the differences observed between the high intensity RRMS (dark grey) and low intensity SPMS samples (light grey) selected for analysis. Results from the sandwich assay are shown in the left panel and the corresponding results from the direct labelling intensities to the right.

**Supplementary figure 7. Correlation of paired antibodies targeting SERPINA1.** The five antibodies targeting SERPINA1 were generated against different protein fragments together covering the whole protein sequence. Intensity profiles from all five antibodies correlated positively in CSF (Spearman's  $\rho > 0.49$ ) for sample treatment at both temperatures, with results from samples treated at 56°C shown in the figure. One sample's intensity level deviated and was therefore omitted from all correlations.