

JoVE submission

Title

Sample Preparation for Endopeptidomic Analysis in Human Cerebrospinal Fluid

AUTHORS & AFFILIATIONS:

Hansson KT, Skillbäck T, Pernevik E, Holmén-Larsson J, Brinkmalm G, Blennow K, Zetterberg H, Gobom J.

Inst. of Neuroscience and Physiology, Dept. of Psychiatry and Neurochemistry, The Sahlgrenska Academy at University of Gothenburg, Mölndal, Sweden

Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden

Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK

Keywords

Cerebrospinal fluid, peptidomics, endogenous peptides, biomarkers, neurodegeneration, Alzheimer's disease, Parkinsonian Disorders, LC-MS/MS, pre-fractionation, multiplexed isobaric labelling

Short abstract

A method for mass spectrometric analysis of endogenous peptides in human cerebrospinal fluid (CSF) is presented. By employing molecular weight filtration, chromatographic pre-fractionation, mass spectrometric analysis and a subsequent combination of peptide identification strategies, it was possible to expand the known CSF peptidome nearly ten-fold compared to previous studies.

Long abstract

This study was performed to identify as large a part as possible of the human cerebrospinal fluid (CSF) endopeptidome, particularly peptides in the lower concentration ranges of CSF. For this purpose, a previously developed method based on molecular weight filtration and mass spectrometric analysis, was combined with an offline high-pH reverse phase HPLC pre-fractionation step.

Endogenous peptides are the result of *in vivo* proteolytic activity and their relative concentrations could hypothetically be used to monitor the function of a biological system. Since CSF is the main pathway for removal of molecules shed by cells in the central nervous system (CNS), endogenous peptides found in CSF could potentially be used as biomarkers for disorders affecting the CNS, such as neurodegeneration.

CSF is a complex mixture and has a wide dynamic range of proteins, spanning 8-9 orders of magnitude^{1,2}; two features which together make in-depths analysis complicated. Since endogenous peptides are recovered by separating them from the protein content through molecular weight filtration, it is possible to largely avoid including the few highly abundant proteins which constitute the bulk mass of CSF. Studying the endogenous peptide content, rather than the proteome through

proteolysis of the whole sample, therefore allows for use of larger sample volumes, potentially increasing the relative amount of low abundant peptides.

The filtrate complexity was addressed further by including a HPLC pre-fractionation step over an alkaline gradient. This increased protocol resolving power both by subdividing the sample as well as introducing orthogonality towards the acidic gradient used during HPLC-MS analysis³⁻⁵. By also performing a concatenation scheme where 60 fractions were pooled into 12, analysis time consumption could be reduced while still largely avoiding co-elution.

Finally, a combination of peptide identification strategies was used, increasing data analysis time considerably but also resulting in a significantly higher number of identified peptides.

Introduction

Biomarkers in cerebrospinal fluid (CSF) are currently transforming research into neurodegenerative disorders. In Alzheimer's disease, the most common neurodegenerative disorder, affecting over 60 million people worldwide{Wimo, 2017 #426;Scheltens, 2016 #398}, a biomarker triplet consisting of the peptide amyloid beta, microtubule-stabilizing protein tau, and a phosphorylated tau form, can detect the disease with high sensitivity and specificity, and has been included in the diagnostic criteria{Dubois, 2014 #428}. In other neurodegenerative diseases, such as Parkinson's disease and Multiple Sclerosis, proteomic studies have identified numerous biomarker candidates, some of which are currently under evaluation in clinical studies{Olsson, 2016 #366;Spellman, 2015 #128;Höglund, 2015 #380}.

Alongside proteins, CSF also contains an abundance of endogenous peptides⁶⁻¹¹. Constituting cleavage products of many brain-derived proteins, these peptides also represent a potentially important source of disease biomarkers. To increase the inventory of identified endogenous peptides in human CSF and enable CSF endopeptidomic analyses in clinical studies, a method was developed for sample preparation and LC-MS analysis. The application of this method in a recent study (Ref. Karl) resulted in the identification of more than 18,000 endogenous CSF peptides in pooled CSF samples from several individuals of non-specific diagnosis, expanding the known CSF endopeptidome ten-fold. The method can optionally be used in conjunction with the Tandem Mass Tag (TMT) approach for quantification.

Sample Preparation

In the case of biomarker discovery in complex biological samples there is a potential advantage in selectively removing parts of the material. Since the main source of protein mass in CSF is plasma constituents (e.g. albumin and immunoglobulins) passing over the blood brain barrier^{1,13}, proteolytic digestion results in highly abundant proteins dominating the peptide content. Furthermore, endogenous peptides can be readily separated from the high-abundant proteins such as albumin and immunoglobulins that make up the bulk of the CSF protein contents^{1,13}, allowing a significantly larger volume of CSF peptide extract to be used for LC-MS analysis, thereby enabling detection of lower-abundant peptides.

In the protocol presented here, molecular weight filtration was used to separate the CSF peptides from the protein fraction, a method that has been used in several previous studies{Berven, 2007 #416;Hölttä, 2016 #427;Hölttä, 2015 #404;Hölttä, 2012 #317;Yuan, 2005 #418;Zougman, 2007 #415}. The filtration step was followed by an offline reverse-phase (RP) HPLC pre-fractionation step performed over a high-pH mobile phase gradient. By performing two RP-HPLC steps in tandem, with pH being the main distinction, the difference in selectivity between the two steps results mainly from altered peptide retention as a consequence of different peptide charge states. The application of high-pH peptide pre-fractionation prior to LC-MS under acidic conditions has proven efficient in increasing

peptide identification^{4,5}, and even to be superior for this purpose in complex biological samples compared to more orthogonal separation modes¹⁷, such as strong cation exchange (SCX) and RP³. To shorten the analysis time, a concatenation scheme was used, pooling every 12th fraction (e.g., fractions 1, 13, 25, 37, 49 and 61), which due to the high resolving power of RP-HPLC still largely avoids co-elution of peptides from different fractions in the LC-MS step^{3,18}.

Peptide identification

Peptide identification in peptidomic studies differs from that of proteomic studies in that no enzyme cleavage can be specified in the database search, and as a consequence, identification rates are usually lower^{Hölttä, 2012 #317}. A recent study^{Hansson, 2017 #414} showed that the identification rates for endogenous peptides obtained with Sequest and Mascot were substantially improved when the default scoring algorithm of the respective software program was modified using the adaptive scoring algorithm Percolator, indicating that optimal scoring algorithms for endogenous peptides differ from that of tryptic peptides. In that study, identification based on automatic peptide *de novo* sequencing using the software PEAKS (BSI) was found to be complementary to the two fragment ion fingerprinting-based search engines, resulting in a significantly larger set of identified peptides.

Highlighted protocol

1. Extraction of human CSF through lumbar puncture.

1.1. Immediately after extraction cell debris and other non-soluble material is removed through centrifugation. The CSF is subsequently checked for blood contamination and moved on to clinical analysis or storage at -80°C.

2. Pre-treatment of CSF (1.5 ml sample volume, no quantification):

2.1. Protein degradation: A chaotropic agent, guanidinium hydrochloride (GdnHCl), is added to the sample in order to denature higher protein structures. GdnHCl both affects solvent viscosity and interacts with the polypeptide chain which results in protein unfolding being energetically favourable²⁴. In short, the chaotropic agent dissolves protein aggregates and increases recovery of endogenous peptides during subsequent filtration.

2.2. Reduction and alkylation: Inter and intra disulphide bridges between cysteine residues in peptides and proteins are first reduced (broken) and subsequently alkylated (capped with an alkyl group). This results in further disruption of protein aggregates and ensures that peptides do not spontaneously form new disulphide bridges in later stages of the protocol.

3. Pre-treatment of CSF (10 x 150 µl sample volume, TMT-quantification)

3.1. Degradation, reduction and alkylation are performed just as in the non-quantifiable protocol steps [2.1-2.2].

3.2. Each of the 10-plex Tandem Mass Tag (TMT₁₀) reagents are dissolved in AcN and added to their corresponding samples. TMT includes an NHS-ester group which reacts with the primary amines present at peptide N-termini and Lysine residues.

3.3. Since the separately labelled samples are to be combined it is vital to ensure the labelling reaction is quenched. By addition of an abundance of amine groups in the form of hydroxylamine, the remaining TMT reagent is allowed to react and is thus rendered inert.

3.4. The content of each vial is combined into a single multiplexed sample.

4. Molecular weight cut-off (MWCO) ultrafiltration

4.1. Recovery of endogenous peptides through application of a 30 kDa MWCO cellulose filter which retains proteins and larger polypeptides whilst peptides pass through.

5. Solid phase extraction – Sample cleaning/de-salting

- 5.1. Performed in order to remove salts and various unwanted CSF components. The C₁₈-cartridge used has similar retention properties to the HPLC separation column used in the fractionation step.
- 5.2. The sample is loaded onto the cartridge; peptides are retained by hydrophobic interaction with the C₁₈ packing material, and components which do not adhere are removed out during subsequent washing steps.
- 5.3. As a last step the peptides are eluted by addition of organic solvent (acetonitrile in this case), which then is rapidly evaporated in a vacuum concentrator.

6. Offline high-pH reverse phase HPLC sample fractionation

- 6.1. The sample pre-fractionation over an alkaline mobile phase gradient resulted in de-convolution of the sample and simultaneously introduced orthogonality towards the second, acidic, HPLC gradient. Peptide retention is dependent on pH, primarily through charge distribution, and the required amount of organic solvent to elute them will therefore shift slightly depending on the pH of the mobile phase.
- 6.2. In order to reduce the total analysis time in the following HPLC-MS/MS step, fractions were concatenated. Every 12th fraction was pooled which over a 60 minute gradient results in 12 concatenated fractions, each containing 5 sub-fractions. The sub-fractions were collected with a sufficiently large concentration difference of organic solvent to avoid co-elution during MS-analysis.

7. Reverse phase HPLC coupled high resolution MS/MS analysis

- 7.1. 12 fractions from an original sample of endogenous peptides from 1.5 mL CSF were re-dissolved and consecutively loaded on the HPLC and separated over a 210 min acidic gradient.
- 7.2. HPLC-MS/MS analysis was performed in data-dependent acquisition mode, the 10 most intense peaks from a full MS scan were sent on to HCD-fragmentation and subsequent MS/MS detection.

8. Peptide Identification

- 8.1. In order to address the issue of identifying endogenous peptides two separate search strategies were employed:
 - 8.1.1. Fragment ion fingerprinting – performed in proteomics software Proteome Discoverer 2.0 (Thermo Fisher Scientific) by search engines Mascot v2.4 (Matrix Science) and Sequest HT (Thermo)
 - 8.1.2. *De novo*-sequencing – performed by software/search engine PEAKS v7.5 (Bioinformatic Solutions Inc.)
- 8.2. The main differences between the three search engines involve mode of PSM generation and PSM-validation. In the case of the two fragment ion fingerprinting-based search engines, PSMs are identified by direct comparison of MS/MS data to a library containing theoretical peptide information. In the case of *De novo*-sequencing the search engine first goes over the spectrum information of each MS/MS scan, annotates the mass difference between peaks (which corresponds to one or several AAs), adds a peptide sequence tag to the MS/MS information and submits the result to a database.
- 8.3. PSM-validation: PSMs returned from each database search are subsequently compared against a second database containing decoy peptide sequences (e.g. peptides with scrambled AA sequences compared to the PSMs). The search engines apply slightly different strategies for this step but at the end the process allows for a measure of statistical significance to be added to the identification.

Detailed protocol

List of materials

- 1 M Triethylammonium bicarbonate (TEAB) (Fluka, Sigma-Aldrich; 17902-100ML)
- 8 M Guanidinium hydrochloride (GdnHCl) (Sigma-Aldrich; G7294-100ML)
- Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (Pierce; 20490)
- Iodoacetamide (IAA) (SIGMA; I1149-5G)
- Hydroxylamine 50% (Sigma-Aldrich; 457804-50ML)
- Acetonitrile (AcN), Far UV, HPLC gradient grade (Sigma-Aldrich; 271004-2L)
- Formic acid (FA) (Fluka, Sigma-Aldrich; 56302-1mL-F)
- Trifluoroacetic acid (TFA) (Sigma-Aldrich; T6508-10AMP)
- Ammonium hydroxide solution (NH₄OH) (Sigma-Aldrich 30501-1L-1M)
- Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-30 membrane (Merck Millipore, UFC903024)
- Sep-Pak C18, 100 mg (Waters, WAT023590)
- Resprep 12-port SPE Manifold (Restek, 26077)
- TMT10plex™ Isobaric Label Reagent Set (Thermo Fisher Scientific, 90110)

1. Extraction of human cerebrospinal fluid (CSF):

- 1.1. CSF is extracted through lumbar puncture, performed by a trained physician, using a standardized protocol (ref.). Cell debris and other non-soluble material is removed through centrifugation at 2500 g for 20 min, the supernatant is recovered and stored at -80°C.

2. Pre-treatment of CSF (1.5 ml sample volume, no quantification):

- 2.1. Thaw 1.5 mL CSF aliquots at room temperature (RT), transfer the contents to 10 mL polypropylene tubes and add 80 µL 1 M TEAB as a buffering agent.
- 2.2. Add 1.5 mL 8 M GdnHCl and vortex gently at RT for 10 min.
- 2.3. Add 60 µL 200 mM of aqueous TCEP and incubate at 55 °C for 1 h to reduce cysteine disulphides.
- 2.4. Add 60 µL 400 mM IAA and incubate at RT in darkness for 30 min to alkylate cysteines.
- 2.5. Add 2.8 mL de-ionised water to dilute the sample prior to filtration

3. Pre-treatment of CSF (10 x 150 µl sample volume, TMT-quantification):

- 3.1. Thaw 150 µL CSF aliquots from 10 individuals at RT and subsequently transfer the contents to individual 1.5 mL Lo-bind Eppendorf tubes and add 8 µL 1 M TEAB.
- 3.2. Add 50 µL 8 M GdnHCl and vortex gently at RT for 10 min.
- 3.3. Add 6 µL 200 mM of aqueous TCEP and incubation at 55 °C for 1 h to reduce cysteine disulphides.
- 3.4. Add 6 µL 400 mM aqueous IAA and incubate at RT and darkness for 30 min to alkylate to alkylate cysteines.
- 3.5. Preparation of the TMT-labelling kit – Allow the TMT-reagent vials to reach RT prior to opening, add 41 µL HPLC-grade AcN and dissolve by gentle agitation for 5 min
- 3.6. Transfer 30 µL of TMT-reagent solution to the corresponding sample and incubate for 1h at RT under gentle agitation.
- 3.7. Add 8 µL 5% hydroxylamine and shake gently at RT for 20 min to quench the labelling reaction
- 3.8. Combine the contents of each individually labelled sample in a single 15 mL polypropylene tube, dilute with 6.4 mL de-ionized water and vortex briefly

4. MWCO filtration

- 4.1. Condition the filters by loading 10 mL aqueous 2.4 M GdnHCl, 25 mM TEAB and centrifuging for 15 min. All centrifugation steps were performed at 2500 x g at RT. Discard the flow-through.
- 4.2. Load the samples and centrifuge for 30 min.

- 4.3. Load 5 mL de-ionised water and centrifuge for 15 min.
5. **De-salting and sample clean-up by SepPak C₁₈ solid phase extraction (SPE)**
- 5.1. Non-TMT-labelled samples: acidify the samples by addition of 880 µL 1% Trifluoroacetic acid (TFA)
- 5.2. TMT-labelled samples: add 20 mL 0.1% TFA to reduce pH and lower AcN concentration from 3% to 1%
- 5.3. (If pH > 3) Titrate the sample with 20% phosphoric acid until sample pH is <3
- 5.4. Condition the SPE cartridges by addition of 1 mL 84% AcN, 0.1% FA, discard the flow-through. Repeat once.
- 5.5. Equilibrate the cartridge by addition of 1 mL 0.1% TFA, discard flow-through. Repeat once.
- 5.6. Load the sample and let the flow-through run into waste
- 5.7. Wash the sample by loading 1 mL 0.1% TFA, discard flow-through. Repeat once.
- 5.8. Place 1.5 mL Lo-bind Eppendorf tubes under the cartridge and elute the sample by addition of 1 mL 84% AcN, 0.1% FA
- 5.9. Remove the solvents from the flow-through by evaporation in a vacuum centrifuge and store at -80°C
6. **Offline high-pH reverse phase HPLC sample fractionation**
- 6.1. Re-dissolve the sample in aqueous solution of 16 µL 2% AcN, 2.5 mM ammonium hydroxide (NH₃OH) by gentle agitation for 20 min
- 6.2. 15 µL of the sample was injected into an Ultimate 3000 (rapid separation liquid chromatography) RSLC system equipped with an internal fraction collector for 96 deep-well plates, based on an HPLC-setup by Batth *et al*³ with minor alterations. Fractionation was performed at a flow of 100 µL/min over an XBridge BEH130 C18 3.5 µm, 2.1 mm x 250 mm separation column (Waters), collecting one fraction per minute over a 60 min gradient.
- 6.3. Mobile Phases:
- 6.3.1. Buffer A) Pure water
- 6.3.2. Buffer B) 84% AcN (Aqueous)
- 6.3.3. Buffer C) 25 mM NH₄OH (Aqueous)
- 6.3.4. Loading buffer) 2.5 mM NH₄OH, 2% AcN (Aqueous)
- 6.4. Gradient:

t [min]	0	4	50	64	68	69	76
Buffer B [%]	2	2	50	90	90	2	2
Buffer C [%]	10	10	10	10	10	10	10
Fraction collection		Start		Stop			

- 6.5. Fractions were collected repetitively in 12 wells, concatenating peptides spaced apart by 12 minutes - resulting in 12 fractions, each containing 5 concatenated sub-fractions.
- 6.6. Remove the solvents from the flow-through by evaporation in a vacuum centrifuge and store at -80°C
7. **Reverse phase HPLC coupled high resolution MS/MS analysis**
- 7.1. Each of the 12 fractions were re-dissolved by addition of 6 µL aqueous 0.05% TFA, 2% AcN (Loading Buffer) and shaking at RT for 20 min
- 7.2. 5 µL was injected into an Ultimate 3000 RSLC nano system (Thermo), operating in trap column configuration (trap column: Acclaim PepMap 100, 75 µm x 2 cm, C₁₈, 100 Å pore size, 3 µm particle size; separation column: Acclaim PepMap C18, 75 µm x 500 mm, 100Å pore size, 2 µm particle size; both separation and trap columns were acquired from Thermo).
- 7.3. Mobile phase buffers:
- 7.3.1. Buffer A) 0.1% FA (Aqueous)

- 7.3.2. Buffer B) 0.1% FA, 84% AcN (Aqueous)
- 7.3.3. Loading Buffer, LB) 0.05% TFA, 2% AcN (Aqueous)
- 7.4. Gradient:

t [min]	0	10	11	100	170	175	181	210
Buffer B [%]	2	2	7	26	45	80	2	2

- 7.5. The HPLC was connected to an Orbitrap Fusion Tribrid mass spectrometer (Thermo) via a FlexiSpray nano-ESI interface (Thermo). Full scan spectra were recorded in MS mode at a resolution setting of 120,000 (2.0e5 AGC target) over the m/z range 350-1400. The mass spectrometer was operated in the data-dependent acquisition mode, recording MS/MS spectra from the top ten most intense peaks with $m/z > 150$ and within the intensity range 1.0e4-1.0e5 were selected for fragment ion analysis. Precursor ions were isolated using a quadrupole isolation window of 3 m/z . Dynamic exclusion was used, with an exclusion time of 15 s and an m/z tolerance of ± 10 ppm. Fragmentation was performed in the higher-collision energy dissociation (HCD) cell (29% collision energy) and MS/MS acquisitions were recorded in the orbitrap at a resolution setting of 30,000 (5.0e4 AGC target value).

8. Peptide Identification

- 8.1. Three search engines were used in parallel for peptide identification; Mascot v2.4 (Matrix Science), Sequest HT (Thermo) and PEAKS v7.5 (Bioinformatic Solutions Inc.) The following settings were used for all three programs unless otherwise specified:
 - 8.1.1. Database: UniProt_SwissProt [version2015_11]
 - 8.1.2. Taxonomy: homo sapiens
 - 8.1.3. Enzyme: none
 - 8.1.4. Max. missed cleavages: 0
 - 8.1.5. Instrument (Mascot only): ESI-Trap
 - 8.1.6. Min. peptide length (SequestHT only): 6
 - 8.1.7. Precursor mass tolerance: 15 ppm
 - 8.1.8. Fragment mass tolerance: 0.05 Da
 - 8.1.9. Static modifications: Carbamidomethyl (C); [If labelled] TMT10plex (N-Term)
 - 8.1.10. Dynamic modifications: Oxidation (M); [If labelled] TMT10plex (K)
 - 8.1.11. Peptide-spectrum match (PSM) validator: Percolator (Mascot and Sequest HT only) or Decoy Fusion (PEAKS only)
 - 8.1.12. Target FDR: 0.01

Representative results

The method here presented has been applied and evaluated in three studies prior to the introduction of pre-fragmentation (Table 1). The first study used offline LC for spotting CSF fractions on a MALDI target plate and resulted in 730 identified endogenous peptides {Höltkä, 2012 #317}. In the two following studies, TMT labelling was employed. Primarily in a case/control study for identification and characterisation of potential biomarkers in the CSF endopeptidome and proteome simultaneously {Höltkä, 2014 #362}, and in the second study TMT was used to monitoring treatment effects *in vivo* of a γ -secretase inhibitor on the peptide expression in CSF over 36 hours {Höltkä, 2016 #427}. In the case/control study 437 endogenous peptides were identified, 64 of which significantly altered in concentration between individuals with AD and healthy controls. The third, treatment study, identified 1798 endogenous peptides, 11 of the monitored peptides could be shown to respond to the treatment.

In the fourth study, the aim was to increase the number of identified CSF peptides, particularly to identify lower-abundant peptides. Therefore, peptide pre-fractionation by HpH-RP chromatography was included and a 10-fold larger CSF sample volume was used, resulting in identification of 18,031 peptides. In this study, no TMT-labelling was performed. In addition to sample fractionation, the most recent study employed a combined peptide identification approach (Mascot, Sequest, PEAKS), whereas in the first three studies only Mascot database searching was performed, which to some extent accounts for the larger number of peptides identified.

Table 1: Compilation of recent studies performed by this group which applies molecular weight filtration and mass spectrometric analysis for identification of endogenous peptides in human CSF.

Study summary	TMT labelling (y/n)	HpH-RP fractionation (y/n)	Corresponding volume of CSF per MS-analysis (μl)	Number of identified peptides	Comment	Reference
Explorative CSF peptidome analysis	n	n	500	730	Offline LC MALDI target preparation, MALDI-MS; evaluation of MWCO filters	{Hölttä, 2012 #317}
Quantitative comparison of CSF peptides; samples from 8 AD + 8 Ctrl	y	n	200	437	HPLC-ESI MS; combined peptidomic and proteomic protocol	{Hölttä, 2014 #362}
AD gamma secretase inhibitor treatment study	y	n	300	1798	HPLC-ESI MS; CSF extracted at six time points after treatment	{Hölttä, 2016 #363}
Expanding the CSF peptidome	n	y	750-1000	18,031	HPLC-ESI MS; combination of peptide identification softwares	{Hansson, 2017 #414}

Figure and table legends

Table 1:

Discussion

The introduction of an high-pH RP HPLC pre-fractionation step to a previously developed protocol for recovery of endogenous peptides by molecular weight ultrafiltration¹⁰ reduced relative sample complexity and thereby allowed for a 5-fold larger sample volume to be studied. This, in turn, increased the concentration of the subset of peptides present in each fraction and thereby improved the chances of detecting low abundant peptides.

By performing an identification strategy for endogenous peptides which employed three software programs in parallel, it was possible to expand the known CSF endopeptidome more than 10-fold. A total of 18,031 endogenous peptides were identified in a preliminary trial comprising three separate HPLC-MS/MS rounds, and two CSF sample materials. Among the identifications were a large number of endogenous peptides derived from proteins previously noted in the context of neurodegenerative disorders.

Due to the small overlap observed in identified peptides between the three search engines, an attempt at evaluating correctness of identification was made¹. By studying two search engines at-a-time and focusing on the subset of MS/MS spectra used by both to generate PSMs it was possible to give a value on identification discrepancy. At an FDR-level of 1% the number of deviating PSMs between two search engines is expected to amount to between 0 and 2%, higher than 2% would indicate that combining the identified peptides artificially augments results. The results from the identification evaluation performed here observed that the highest number of deviating PSMs generated from the same MS/MS spectra was 1.93%, noted between search engines PEAKS and Mascot.

Acknowledgements

Many thanks to Tanveer Batth and colleagues for advice in setting up the pre-fractionation method.

This work was supported by funding from the Swedish Research Council, the Wallström and Sjöblom Foundation, the Gun and Bertil Stohne Foundation Stiftelse, the Magnus Bergwall Foundation, the Åhlén Foundation, Alzheimerfonden, Demensförbundet, Stiftelsen för Gamla Tjänarinnor, the Knut and Alice Wallenberg Foundation, Frimurarestiftelsen, and FoU-Västra Götalandsregionen.

The main recipients of funding for this project were Kaj Blennow, Henrik Zetterberg and Johan Gobom.

¹ Since peptide identification was performed based on plausibility (scoring algorithm and subsequent comparison to decoys, FDR), rather than on comparison to known standards spiked into the sample, correctness can only be estimated.

References

- 1 Gulbrandsen, A. *et al.* In-depth Characterization of the Cerebrospinal Fluid (CSF) Proteome Displayed Through the CSF Proteome Resource (CSF-PR). *Molecular & Cellular Proteomics*. **13** (11), 3152-3163, doi:10.1074/mcp.M114.038554, (2014).
- 2 Hühmer, A. F., Biringer, R. G., Amato, H., Fonteh, A. N. & Harrington, M. G. Protein Analysis in Human Cerebrospinal Fluid: Physiological Aspects, Current Progress and Future Challenges. *Disease Markers*. **22** (1-2), 3-26, doi:10.1155/2006/158797, (2006).
- 3 Batth, T. S., Francavilla, C. & Olsen, J. V. Off-Line High-pH Reversed-Phase Fractionation for In-Depth Phosphoproteomics. *Journal of Proteome Research*. **13** (12), 6176-6186, doi:10.1021/pr500893m, (2014).
- 4 Cao, Z., Tang, H.-Y., Wang, H., Liu, Q. & Speicher, D. W. Systematic Comparison of Fractionation Methods for In-depth Analysis of Plasma Proteomes. *Journal of Proteome Research*. **11** (6), 3090-3100, doi:10.1021/pr201068b, (2012).
- 5 Chiu, C.-W., Chang, C.-L. & Chen, S.-F. Evaluation of peptide fractionation strategies used in proteome analysis. *Journal of Separation Science*. **35** (23), 3293-3301, doi:10.1002/jssc.201200631, (2012).
- 6 Stark, M., Danielsson, O., Griffiths, W. J., Jörnvall, H. & Johansson, J. Peptide repertoire of human cerebrospinal fluid: novel proteolytic fragments of neuroendocrine proteins. *Journal of Chromatography B: Biomedical Sciences and Applications*. **754** (2), 357-367 (2001).
- 7 Yuan, X. & Desiderio, D. M. Human cerebrospinal fluid peptidomics. *Journal of mass spectrometry*. **40** (2), 176-181 (2005).
- 8 Berven, F. S. *et al.* Preal., F. S.(2005).ence on the low molecular weight cerebrospinal fluid proteome. *Proteomics-Clinical Applications*. **1** (7), 699-711 (2007).
- 9 Zougman, A. *et al.* Integrated analysis of the cerebrospinal fluid peptidome and proteome. *Journal of proteome research*. **7** (01), 386-399 (2007).
- 10 Hölttä, M. *et al.* Peptidome analysis of cerebrospinal fluid by LC-MALDI MS. *PLoS One*. **7** (8), e42555 (2012).
- 11 Hölttä, M. *et al.* An integrated workflow for multiplex CSF proteomics and peptidomics-identification of candidate cerebrospinal fluid biomarkers of Alzheimer's disease. *J Proteome Res*. **14** (2), 654-663, doi:10.1021/pr501076j, (2015).
- 12 Blennow, K., Hampel, H., Weiner, M. & Zetterberg, H. Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nat Rev Neurol*. **6** (3), 131-144, doi:10.1038/nrneurol.2010.4, (2010).
- 13 Kroksveen, A. C., Opsahl, J. A., Aye, T. T., Ulvik, R. J. & Berven, F. S. Proteomics of human cerebrospinal fluid: Discovery and verification of biomarker candidates in neurodegenerative diseases using quantitative proteomics. *Journal of Proteomics*. **74** (4), 371-388, doi:<http://dx.doi.org/10.1016/j.jprot.2010.11.010>, (2011).

- 14 Gilar, M., Olivova, P., Daly, A. E. & Gebler, J. C. Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *Journal of Separation Science*. **28** (14), 1694-1703, doi:10.1002/jssc.200500116, (2005).
- 15 Hao, P., Qian, J., Ren, Y. & Sze, S. K. Electrostatic Repulsion-Hydrophilic Interaction Chromatography (ERLIC) versus Strong Cation Exchange (SCX) for Fractionation of iTRAQ-Labeled Peptides. *Journal of Proteome Research*. **10** (12), 5568-5574, doi:10.1021/pr2007686, (2011).
- 16 Ritorto, M. S., Cook, K., Tyagi, K., Pedrioli, P. G. A. & Trost, M. Hydrophilic Strong Anion Exchange (hSAX) Chromatography for Highly Orthogonal Peptide Separation of Complex Proteomes. *Journal of Proteome Research*. **12** (6), 2449-2457, doi:10.1021/pr301011r, (2013).
- 17 Gilar, M., Olivova, P., Daly, A. E. & Gebler, J. C. Orthogonality of Separation in Two-Dimensional Liquid Chromatography. *Analytical Chemistry*. **77** (19), 6426-6434, doi:10.1021/ac050923i, (2005).
- 18 Yang, F., Shen, Y., Camp, D. G. & Smith, R. D. High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. *Expert review of proteomics*. **9** (2), 129-134 (2012).
- 19 Brosch, M., Yu, L., Hubbard, T. & Choudhary, J. Accurate and Sensitive Peptide Identification with Mascot Percolator. *Journal of Proteome Research*. **8** (6), 3176-3181, doi:10.1021/pr800982s, (2009).
- 20 Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature methods*. **4** (3), 207-214 (2007).
- 21 Fitzgibbon, M., Li, Q. & McIntosh, M. Modes of Inference for Evaluating the Confidence of Peptide Identifications. *Journal of Proteome Research*. **7** (1), 35-39, doi:10.1021/pr7007303, (2008).
- 22 Spivak, M., Weston, J., Bottou, L., Käll, L. & Noble, W. S. Improvements to the Percolator Algorithm for Peptide Identification from Shotgun Proteomics Data Sets. *Journal of Proteome Research*. **8** (7), 3737-3745, doi:10.1021/pr801109k, (2009).
- 23 Zhang, J. *et al.* PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Molecular & Cellular Proteomics*. **11** (4), M111. 010587 (2012).
- 24 Möglich, A., Krieger, F. & Kiefhaber, T. Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *Journal of molecular biology*. **345** (1), 153-162 (2005).

Fortfarande att fixa med manus:

Beskriv MS analys i Introduction

Inkludera kort avsnitt om kvantitativ inmärkning och TMT

Bättre retention + jonisering mha TMT

Bilder och tabeller

Tabel 1 – Time-points for both HPLCs