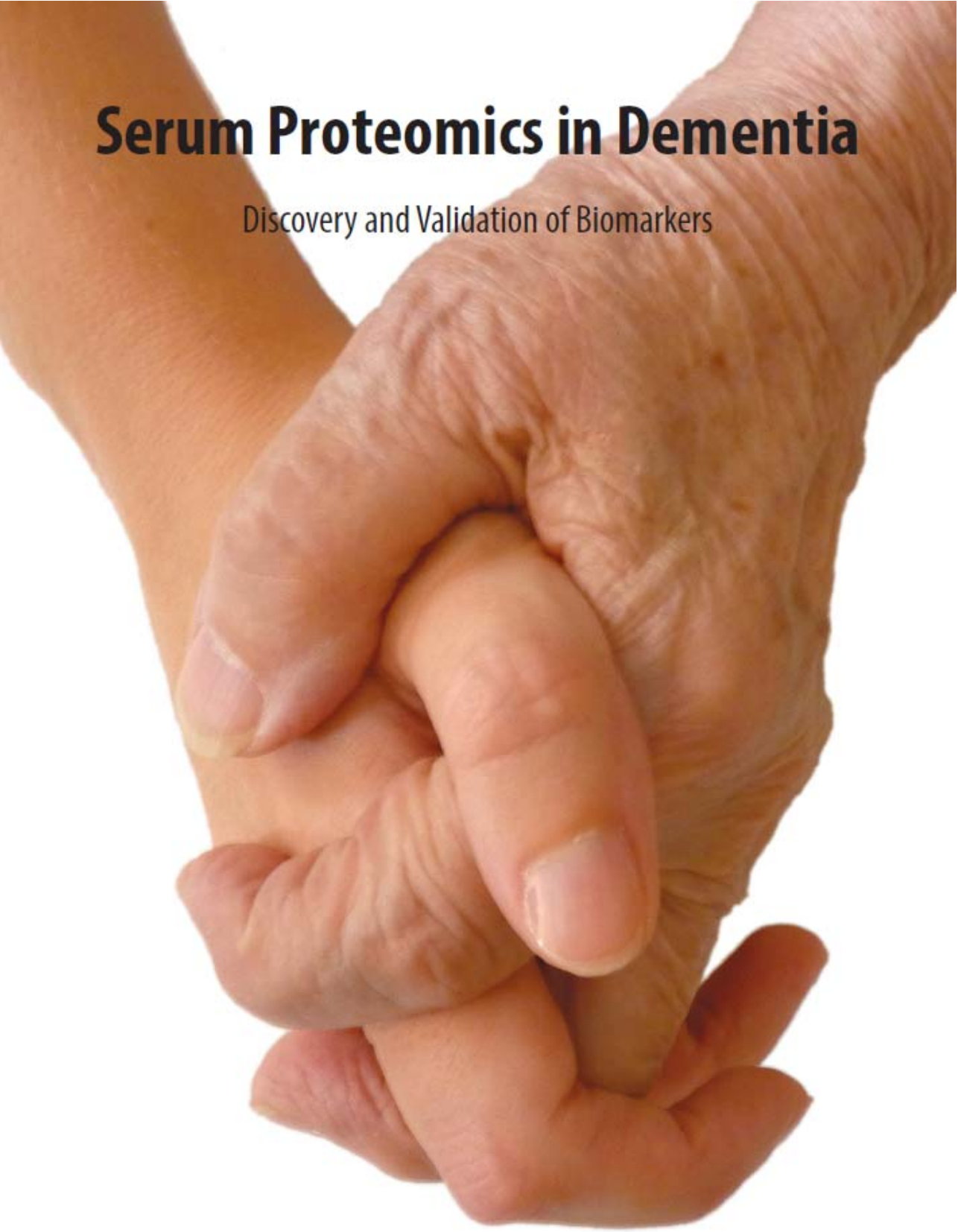
Serum Proteomics in Dementia

Discovery and Validation of Biomarkers



Linda IJsselstijn

Serum Proteomics in Dementia

Discovery and Validation of Biomarkers

Linda IJsselstijn

The studies described in this thesis were performed at the Department of Neurology of the Erasmus University Medical Center in Rotterdam, the Netherlands.

This research project was supported by Eurostars project Pepspex and Internationale Stichting Alzheimer Onderzoek. The synthetic peptides were kindly provided by Pepscan Presto B.V. (Lelystad, the Netherlands).

Financial support for the printing of this thesis was kindly provided by:

J.E. Jurriaanse Stichting

Internationale Stichting Alzheimer Onderzoek

Alzheimer Nederland

Pepscan Presto B.V.

ISBN: 978-90-8891-701-1

Cover design: Proefschriftmaken.nl || Uitgeverij BOXPress
Printed & Lay Out by: Proefschriftmaken.nl || Uitgeverij BOXPress
Published by: Uitgeverij BOXPress, 's-Hertogenbosch

Serum Proteomics in Dementia

Discovery and Validation of Biomarkers

Serum proteomics in dementie
Ontdekking en validatie van biomarkers

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de rector magnificus

prof.dr. H.G. Schmidt

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

6 november 2013 om 15.30 uur

door

Linda IJsselstijn
geboren te Rotterdam

ERASMUS UNIVERSITEIT ROTTERDAM

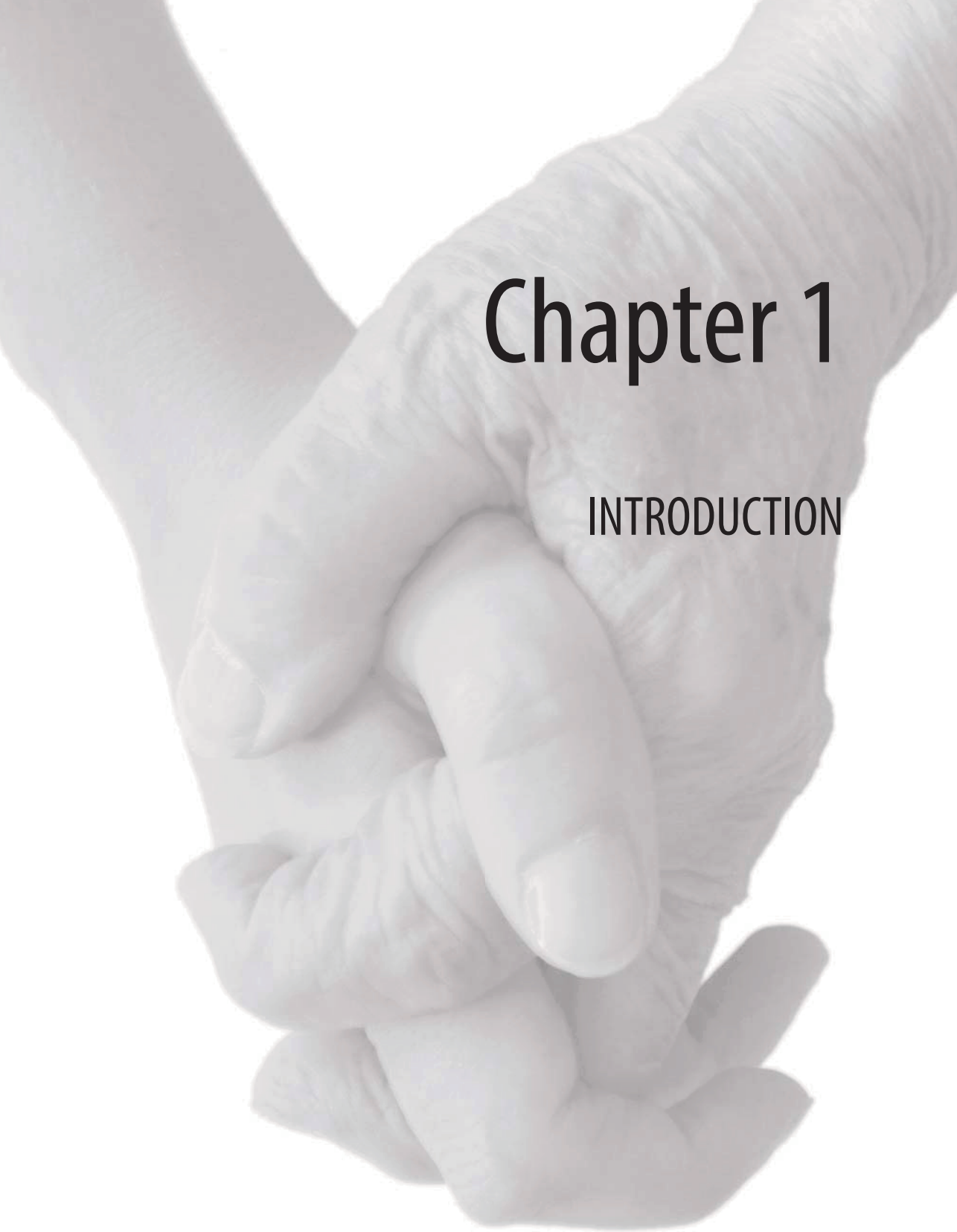
Promotiecommissie:

Promotor:	Prof. dr. P.A.E. Sillevis Smitt
Overige leden:	Prof. dr. P.J. Koudstaal Prof. dr. J.M. Kros Dr. M.A. Ikram
Copromotor:	Dr. T.M. Luider

Contents

List of abbreviations

A2MG	Alpha-2-macroglobulin
ACN	Acetonitrile
AD	Alzheimer's disease
APOE	Apolipoprotein E
Aβ	Amyloid beta
CAMDEX	Cambridge Examination for Mental Disorders of the Elderly interview
CFAB	Complement factor B
CPN2	Carboxypeptidase N subunit 2
CSF	Cerebrospinal fluid
CV	Coefficient of variation
CX3CR	Chemokine (C-X3-C motif) receptor 1
DTT	1,4-dithiothreitol
FA	Formic acid
GFAP	Glial fibrillary acidic protein
GMS	Geriatic Mental State
IgG	Immunoglobulin G
IR	Immunoreactive
LC	Liquid chromatography
LG3BP	Galectin-3 binding protein
LLOQ	Lower limit of quantification
LOD	Limit of detection
LUM	Lumican
m/z	Mass-to-chargo ratio
MCI	Mild cognitive impairment
MMSE	Mini Mental State Examination
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
PBS	Phosphate buffer saline
PZP	Pregnancy zone protein
R^2	Regression coefficient
RSS	Rotterdam Scan Study
SAA4	Serum amyloid A-4 protein
SAMP	Serum amyloid P-component
SD	Standard deviation
SFG	Superior frontal gyrus
SHBG	Sex hormone binding globulin
TFA	Trifluoric acid



Chapter 1

INTRODUCTION

DEMENTIA

Dementia is a syndrome due to disease of the brain, usually of a chronic or progressive nature, in which there is disturbance of multiple higher cortical functions, including memory, thinking, orientation, comprehension, calculation, learning capacity, language, and judgement. Consciousness is not clouded. The impairments of cognitive function are commonly accompanied, and occasionally preceded, by deterioration in emotional control, social behaviour, or motivation. This syndrome occurs in Alzheimer's disease, in cerebrovascular disease, and in other conditions primarily or secondarily affecting the brain (World Health Organization's International Classification of Diseases (ICD-10, version 2007)). It is estimated that there are currently 35.6 million people with dementia around the world and that this number will nearly double every 20 years, to 65.7 million in 2030 and 115.4 million in 2050[1]. Dementia mainly affects the elderly; after 65 years of age the prevalence doubles with every five-year increase in age. The most common types of dementia are Alzheimer's disease (AD; 50-70%), vascular dementia (VaD; 20-30%), frontotemporal dementia (5-10%) and dementia with Lewy bodies (<5%)[1]. The borders between the different types are not at all strict, most of the times a mix of various types is observed. This is particularly true for AD and VaD, and AD and dementia with Lewy bodies. In a large autopsy study of 1050 elderly demented individuals, 86% had AD related pathology. From this percentage only 43% had pure AD, 23% had mixed AD and VaD and 11% had AD with Lewy bodies[2].

Alzheimer's disease

Alzheimer's disease (AD) was first described by Alois Alzheimer in 1907[3, 4]. He described the case of a woman in her fifties who first suffered from delusions, and then rapidly deteriorated exhibiting signs of memory dysfunction, paranoia, and problems with language and behavior. After her death autopsy samples of the brain showed cerebral atrophy, intraneuronal fibrillar bundles, and numerous extracellular deposits in the cerebral cortex, but also atherosclerotic changes. Research in the 1980s, has established that the intraneuronal fibrillar bundles or tangles contained hyperphosphorylated tau protein filaments[5], and that the extracellular deposits were composed mainly of β-amyloid peptides[6], nowadays usually referred to as amyloid plaques. The definite diagnosis of AD can only be made after death, when neuropathological examination shows the presence of neurofibrillary tangles and amyloid plaques in the brain of a person with dementia symptoms. Until this neuropathological examination is done, diagnosis is either "possible" or "probable" AD.

AD can be divided in two types: familial AD and late-onset AD. Familial AD represents less than 1% of all AD cases and is characterized by early onset (<60 years) and autosomal

dominant inheritance of mutations in one of three genes: amyloid precursor protein (*APP*)[7] and presenilins 1 and 2 (*PSEN1*, *PSEN2*)[8, 9]. Most mutations in *APP*, *PSEN1* and *PSEN2* increase the production of β-amyloid peptides. Late-onset AD is characterized by later onset (>60 years) and no genetic linkage, although the apolipoprotein E gene (*APOE*) is known as a risk factor[10, 11]. Although familial AD and late-onset AD differ in age of onset, both forms of the disease are defined by the same pathological features.
At present, five drugs have been approved by the United States Food and Drugs Administration (US FDA) to treat AD[12]. Four of these are acetylcholinesterase inhibitors and the fifth is an NMDA antagonist. However, these drugs can only delay symptom progression and not cure AD.

Mild cognitive impairment

Mild cognitive impairment (MCI) refers to the transitional stage between normal aging and dementia. MCI is defined as the presence of subjective cognitive complaints and objective memory impairment, with normal performance in non-anamnestic cognitive domains and overall general functioning, in the absence of dementia[13]. When memory loss is the predominant symptom in MCI, it is termed amnestic MCI. Amnestic MCI is frequently a preliminary state of AD[14]. Additionally, when not memory but other domains show more impairment (non-amnestic MCI), other types of dementia are more likely to develop [15].

PROTEOMICS

Proteomics is the large-scale study of proteins, in particular of their expression levels, post-translational modifications, and interactions, to obtain insight in disease processes, cellular processes and networks at the protein level[16]. Proteomics is frequently used in the search for protein biomarkers, defined as characteristics (in this case proteins) that are objectively measured and evaluated as indicators of normal biological processes, pathological processes, or pharmacological responses to a therapeutic intervention[17]. The field of proteomics can be roughly divided in discovery proteomics and targeted proteomics. Discovery proteomics is the unbiased analysis of proteins or peptides present in complex samples like body fluids and tissues, and can be used in the search for new biomarkers. In targeted proteomics, proteins of interest are selected prior to analysis and information of only these proteins will be obtained. This method is mainly used for the validation of potential biomarkers found by discovery proteomics and is more focused on quantification of proteins and peptides.

Discovery proteomics

For discovery proteomics, mass spectrometry is the most commonly employed technique. It can be applied to map detectable protein expression, as well as to identify the proteins that are differentially expressed[18]. Every mass spectrometry-based proteomics experiment consists of three steps. (1) Isolation of proteins from their biological source and optionally fractionate them. The protein sample is usually enzymatically digested and the resulting peptide sample is further fractionated. (2) The peptides are subjected to mass-spectrometric analysis. (3) The generated datasets are analyzed by dedicated software and peptide identity can be assigned through database searching[19].

Most biological samples used in discovery proteomics studies are very complex. For example, blood contains more than 10,000 proteins with a dynamic range of concentration that spans more than 10 orders of magnitude[20]. Besides the wide dynamic range, the 14 most abundant proteins make up around 95% of the total protein mass. A commonly used strategy to overcome the dynamic range problem is depletion of the most abundant proteins. However, there is also a downside of depletion namely the removal of proteins that complex with these abundant proteins such as albumin and immunoglobulins. Separation or fractionation of proteins, or after digestion peptides, is another way to handle the overwhelming complexity of body fluids and tissues. The classical and now almost old-fashioned approach is to separate proteins by two-dimensional gel electrophoresis (2DE) followed by analysis of the protein(s) in a single gel spot. Nowadays, 2DE is often replaced by non-gel based technologies like liquid chromatography (LC). LC can be used for the separation of both proteins and peptides. Usually, the separation is performed on peptide level, because LC can then directly be (online) coupled to a mass spectrometry (LC-MS) device. The chemical diversity of peptides (charge, isoelectric point, hydrophobicity, size) makes them very well suited to be separated by nearly every liquid-based separation mode. Often used are reversed-phase LC (RPLC, based on differences in polarity), ion exchange chromatography (IEX, based on charge), size exclusion chromatography (SEC) and hydrophilic interaction chromatography (HILIC)[21]. LC columns can consist of a cluster of packed particles (packed columns) or of a single piece of separation media (monolithic columns). The separation capability of an LC column is described by its peak capacity. The peak capacity is defined as the number of peaks that can be separated within a retention window. The peak capacity can be improved by increasing the column length, decreasing the particle size or usage of a more optimal column material, and by optimizing LC characteristics (e.g. temperature, flow rate and gradient time)[21]. These LC methodologies can be used as stand alone; however, very often they are combined (two-dimensional and even multi-dimensional LC) to further raise the possibility to overcome the complexity of digested protein samples and thereby increase the overall resolving power[22]. Well known

combinations are strong-cation exchange (type of IEX) with RPLC, as two separate columns or packed together in one column (multidimensional protein identification technology, MudPIT), RPLC with RPLC, HILIC with RPLC, and SEC with RPLC. RPLC is generally regarded as the back-end separation of choice due to its compatibility with MS.

Mostly MS analysis of peptides or proteins starts with their ionization. Two ionization techniques predominantly used in proteomic research are electrospray ionization (ESI)[23] and matrix assisted laser desorption/ionization (MALDI)[24]. After ionization various types of analyzers can be used to either simply measure the molecular mass or to determine additional structural features including amino acid sequence or the site of attachment and type of posttranslational modifications. In the latter case, after the initial mass determination, specific ions are selected and subjected to fragmentation through collision. These experiments are referred to as tandem mass spectrometry (MS/MS). In table 1.1, the types of mass spectrometers commonly used in proteomics, along with some of their distinctive traits, are summarized[25].

The amount of data generated by (LC-)MS experiments can be too overwhelming for manual interpretation, so automation of data collection and analysis is essential. The raw data needs to be converted to lists of monoisotopic masses, intensities and in case of LC-MS also retention time. This is done by processes like peak detection, removal of non-peptide peaks and (retention time) alignment. The next step is peptide and protein identification by searching the data against an appropriate database. A well-known web-based search engine is Mascot[26]. Statistical analysis can then be performed to look for (a panel of) differentially expressed peptide masses (not identified), peptides or proteins between groups of interest (potential biomarkers).

Table 1.1. Mass analyzers commonly used in proteomics (freely taken from Domon & Aebersold, 2010).

Analyzer	Implementation	Type	Resolving power	Mass accuracy	Limit of detection	Dynamic range
Quadrupole	TQ-QTOF	In-beam	1,000-2,000	Low	Very low	4-5
Ion trap	IT	Trapping	1,000-2,000	Low	Very low	2-3
TOF	Q-TOF	In-beam	>25,000	High	Low	3
OT/ICR	Hybrid	Trapping	>50,000	Very high	Low	3

TOF, time of flight; OT, Orbitrap; ICR, ion cyclotron resonance; TQ, triple quadrupole; IT, iontrap; Q-TOF, quadrupole time of flight.

Targeted proteomics

Traditionally, the enzyme-linked immunosorbent assay (ELISA) has been the major method used for targeted protein quantification and biomarker validation. In case an ELISA assay or high quality antibodies are available, the quantification of proteins with ELISA is relatively straightforward with a good sensitivity and throughput. The limiting factors for the ELISA are the restricted possibility to analyze more than one protein at the same time and the expensive development of new assays. The simultaneous analysis of multiple proteins (dozens or more) can be done by so called multiplex assays, like antibody arrays[27] and bead-based assays (Luminex®). Multiplex assays have a higher throughput, but are usually less sensitive than ELISAs. This is due to the fact that the optimal detection conditions are not similar for all proteins in the assay, and therefore not all proteins will be analyzed under the ideal circumstances[28].

Alternatively, an antibody-free technology for multiplex and quantitative analysis of proteins is multiple reaction monitoring (MRM). MRM is an MS-based technique for the targeted detection and quantification of selected peptides with known fragmentation patterns in a complex sample[29]. The selected peptides can serve as representatives for the candidate protein, when they have an amino acid sequence that is unique for the candidate protein and are easily detected by MS. A triple quadrupole mass spectrometer (TQ) is the most suitable mass spectrometer for quantification with MRM (table 1.1). In the first quadrupole, a specific precursor ion of the peptide is selected based on its mass-to-charge (m/z) ratio. The precursor ion is fragmented in the second quadrupole, which allows for the selection of a specific fragment of the target peptide ion, according to its m/z ratio, in the third quadrupole. For absolute quantification (AQUA) using MRM an internal standard must be spiked into the sample[30]. Often a peptide containing a stable-isotope labeled amino acid is synthesized based on the sequence of the peptide of interest that is being targeted for quantification, and used as an internal standard. MRM coupled with stable isotope dilution MS has shown to be highly reproducible, within and across laboratories, and sensitive to low µg/mL protein concentrations in complex samples without enrichment of the proteins of interest[31].

PROTEIN BIOMARKERS IN DEMENTIA

In dementia and other neurodegenerative disorders, proteomics research is mainly performed in brain tissue and cerebrospinal fluid (CSF). CSF is the only body fluid that directly interchanges with the extracellular fluid of the central nervous system, making it an ideal source for biomarker discovery in neurodegenerative disorders[32]. For AD, the most reproducible findings are decreased levels of Aβ42 peptide and increased levels of total

tau and phosphorylated tau proteins in CSF[33]. Besides these biomarkers for AD, no reliable pathology associated biomarkers have yet been discovered for other types of dementia. However, hundreds of candidate proteins with altered concentrations have been detected that are related to neurodegeneration. They are mainly involved in energy and protein metabolism, signal transduction, transport, and cytoskeleton formation[34]. Some of these proteins identified in CSF and found to be altered in AD in several proteomic analyses are albumin, α-1-antitrypsin, α-1β glycoprotein, amyloid β A4 protein, apolipoprotein A-I, A-II and E, complement C3a and C4a, cystatin C, retinol-binding protein and β-2-microglobulin[35].

Serum biomarker research in AD

Although CSF is a good resource for biomarker research in AD, it is limited by the invasive procedure and the requirement of highly trained personnel to draw CSF. Biomarker research in serum and plasma has therefore become more popular. Several studies have investigated the levels of Aβ peptides in serum and plasma and most groups have found no significant difference between AD and controls[36]. Furthermore, no correlation was found between CSF Aβ levels and plasma Aβ levels in individual patients[37]. Numerous discovery proteomics studies in serum and plasma have been performed over the last decade[38-46]. Proteins reported to be altered in serum or plasma of AD patients as well as in CSF are α-1-antitrypsin, apolipoprotein A-I and E, and complement C3 and C4[45]. Among all other proteins found related to AD, α-2-macroglobulin and complement factor H show the greatest difference in expression between AD and control samples[39].

SCOPE OF THIS THESIS

The overall goal of this thesis is to search for proteins that are differentially expressed in serum of people who later develop AD or with MCI. We applied state of the art proteomics techniques, unbiased as well as targeted, to search for, identify and quantify these proteins. The main part of this thesis is on serum proteomics research related to AD. In **chapter 2**, I describe the results of a search for proteins which could predict the onset of AD. To this end, the serum proteome of persons who developed AD a few years after blood sampling (presymptomatic AD), and matched controls who remained dementia free, was examined using nano LC Orbitrap MS. **Chapter 3** describes an immunohistochemistry study of a protein found to be elevated in serum of presymptomatic AD patients compared to controls and showed a strong expression in AD brain compared to almost no expression in control brain. The goal of this study was to gain more insight in the expression of this protein and its possible role in AD. In **chapter 4**, I have investigated a reported potential

biomarker for AD, clusterin, for its applicability in the prediction of AD. To this end, an MRM assay specific for clusterin was developed to quantify clusterin levels in serum of presymptomatic AD patients and controls.

Chapter 5 describes a proteomics study performed on samples from MCI patients. The serum proteome of MCI patients and controls were compared by means of nano LC Orbitrap MS. Differentially expressed proteins were quantified using MRM assays. **Chapter 6** is of a more technical nature where I compare targeted and label free MS methods for the quantification of proteins.

Finally, the work presented in this thesis is summarized and discussed in **chapter 7**.

REFERENCES

1. Prince M, Jackson J: **World Alzheimer Report 2009.** 2009.

2. Jellinger KA: **Clinicopathological analysis of dementia disorders in the elderly--an update.** *J Alzheimers Dis* 2006, **9:**61-70.

3. Alzheimer A: **Über eine eigenartige Erkrankung der Hirnrinde.** *Allgemeine Zeitschrift für Psychiatrie Psychisch-Gerichtlich Medizine* 1907, **64:**146-148.

4. Alzheimer A, Stelzmann RA, Schnitzlein HN, Murtagh FR: **An English translation of Alzheimer's 1907 paper, "Uber eine eigenartige Erkankung der Hirnrinde".** *Clin Anat* 1995, **8:**429-431.

5. Grundke-Iqbal I, Iqbal K, Quinlan M, Tung YC, Zaidi MS, Wisniewski HM: **Microtubule-associated protein tau. A component of Alzheimer paired helical filaments.** *J Biol Chem* 1986, **261:**6084-6089.

6. Glenner GG, Wong CW: **Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein.** *Biochem Biophys Res Commun* 1984, **120:**885-890.

7. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, Giuffra L, Haynes A, Irving N, James L, et al.: **Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease.** *Nature* 1991, **349:**704-706.

8. Levy-Lahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, Yu CE, Jondro PD, Schmidt SD, Wang K, et al.: **Candidate gene for the chromosome 1 familial Alzheimer's disease locus.** *Science* 1995, **269:**973-977.

9. Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, Chi H, Lin C, Li G, Holman K, et al: **Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease.** *Nature* 1995, **375:**754-760.

10. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA: **Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families.** *Science* 1993, **261:**921-923.

11. Poirier J, Davignon J, Bouthillier D, Kogan S, Bertrand P, Gauthier S: **Apolipoprotein E polymorphism and Alzheimer's disease.** *Lancet* 1993, **342:**697-699.

12. Neugroschl J, Sano M: **Current treatment and recent clinical research in Alzheimer's disease.** *Mt Sinai J Med* 2010, **77:**3-16.

13. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E: **Mild cognitive impairment: clinical characterization and outcome.** *Arch Neurol* 1999, **56:**303-308.

14. Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B: **Current concepts in mild cognitive impairment.** *Arch Neurol* 2001, **58:**1985-1992.

15. Tabert MH, Manly JJ, Liu X, Pelton GH, Rosenblum S, Jacobs M, Zamora D, Goodkind M, Bell K, Stern Y, Devanand DP: **Neuropsychological prediction of conversion to Alzheimer**

disease in patients with mild cognitive impairment. *Arch Gen Psychiatry* 2006, **63:**916-924.

16. Anderson NL, Anderson NG: **Proteome and proteomics: new technologies, new concepts, and new words.** *Electrophoresis* 1998, **19:**1853-1861.

17. Atkinson AJ, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, Oates JA, Peck CC, Schooley RT, Spilker BA, et al: **Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework.** *Clin Pharmacol Ther* 2001, **69:**89-95.

18. Domon B, Aebersold R: **Mass spectrometry and protein analysis.** *Science* 2006, **312:**212-217.

19. Baldwin MA: **Protein identification by mass spectrometry: issues to be considered.** *Mol Cell Proteomics* 2004, **3:**1-9.

20. Anderson NL, Anderson NG: **The human plasma proteome: history, character, and diagnostic prospects.** *Mol Cell Proteomics* 2002, **1:**845-867.

21. Sandra K, Moshir M, D'Hondt F, Verleysen K, Kas K, Sandra P: **Highly efficient peptide separations in proteomics Part 1. Unidimensional high performance liquid chromatography.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2008, **866:**48-63.

22. Sandra K, Moshir M, D'Hondt F, Tuytten R, Verleysen K, Kas K, Francois I, Sandra P: **Highly efficient peptide separations in proteomics. Part 2: bi- and multidimensional liquid-based separation techniques.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2009, **877:**1019-1039.

23. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM: **Electrospray ionization for mass spectrometry of large biomolecules.** *Science* 1989, **246:**64-71.

24. Karas M, Bachmann D, Bahr U, Hillenkamp F: **Matrix-assisted ultraviolet laser desorption of non-volatile compounds.** *International Journal of Mass Spectrometry and Ion Processes* 1987, **78:**53-68.

25. Domon B, Aebersold R: **Options and considerations when selecting a quantitative proteomics strategy.** *Nat Biotechnol* 2010, **28:**710-721.

26. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS: **Probability-based protein identification by searching sequence databases using mass spectrometry data.** *Electrophoresis* 1999, **20:**3551-3567.

27. Barry R, Soloviev M: **Quantitative protein profiling using antibody arrays.** *Proteomics* 2004, **4:**3717-3726.

28. Fu Q, Schoenhoff FS, Savage WJ, Zhang P, Van Eyk JE: **Multiplex assays for biomarker research and clinical application: translational science coming of age.** *Proteomics Clin Appl* 2010, **4:**271-284.

29. Lange V, Picotti P, Domon B, Aebersold R: **Selected reaction monitoring for quantitative proteomics: a tutorial.** *Mol Syst Biol* 2008, **4:**222.

30. Kirkpatrick DS, Gerber SA, Gygi SP: **The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications.** *Methods* 2005, **35:**265-273.

31. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham AJ, Keshishian H, et al: **Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma.** *Nat Biotechnol* 2009, **27:**633-641.

32. Romeo MJ, Espina V, Lowenthal M, Espina BH, Petricoin EF, 3rd, Liotta LA: **CSF proteome: a protein repository for potential biomarker identification.** *Expert Rev Proteomics* 2005, **2:**57-70.

33. Andreasen N, Minthon L, Davidsson P, Vanmechelen E, Vanderstichele H, Winblad B, Blennow K: **Evaluation of CSF-tau and CSF-Abeta42 as diagnostic markers for Alzheimer disease in clinical practice.** *Arch Neurol* 2001, **58:**373-379.

34. Fountoulakis M, Kossida S: **Proteomics-driven progress in neurodegeneration research.** *Electrophoresis* 2006, **27:**1556-1573.

35. Cedazo-Minguez A, Winblad B: **Biomarkers for Alzheimer's disease and other forms of dementia: clinical needs, limitations and future aspects.** *Exp Gerontol* 2010, **45:**5-14.

36. Blennow K, Hampel H, Weiner M, Zetterberg H: **Cerebrospinal fluid and plasma biomarkers in Alzheimer disease.** *Nat Rev Neurol* 2010, **6:**131-144.

37. Mehta PD, Pirttila T, Patrick BA, Barshatzky M, Mehta SP: **Amyloid beta protein 1-40 and 1-42 levels in matched cerebrospinal fluid and plasma from patients with Alzheimer disease.** *Neurosci Lett* 2001, **304:**102-106.

38. German DC, Gurnani P, Nandi A, Garner HR, Fisher W, Diaz-Arrastia R, O'Suilleabhain P, Rosenblatt KP: **Serum biomarkers for Alzheimer's disease: proteomic discovery.** *Biomed Pharmacother* 2007, **61:**383-389.

39. Hye A, Lynham S, Thambisetty M, Causevic M, Campbell J, Byers HL, Hooper C, Rijsdijk F, Tabrizi SJ, Banner S, et al: **Proteome-based plasma biomarkers for Alzheimer's disease.** *Brain* 2006, **129:**3042-3050.

40. Irizarry MC: **Biomarkers of Alzheimer disease in plasma.** *NeuroRx* 2004, **1:**226-234.

41. Liao PC, Yu L, Kuo CC, Lin CJ, Kuo YM: **Proteomics analysis of plasma for potential biomarkers in the diagnosis of Alzheimer's disease.** *Proteomics Clinical Applications* 2007, **1:**506-512.

42. Lopez MF, Mikulskis A, Kuzdzal S, Bennett DA, Kelly J, Golenko E, DiCesare J, Denoyer E, Patton WF, Ediger R, et al: **High-resolution serum proteomic profiling of Alzheimer disease samples reveals disease-specific, carrier-protein-bound mass signatures.** *Clin Chem* 2005, **51:**1946-1954.

43. Lovestone S, Guntert A, Hye A, Lynham S, Thambisetty M, Ward M: **Proteomics of Alzheimer's disease: understanding mechanisms and seeking biomarkers.** *Expert Rev Proteomics* 2007, **4:**227-238.

44. Ray S, Britschgi M, Herbert C, Takeda-Uchimura Y, Boxer A, Blennow K, Friedman LF, Galasko DR, Jutel M, Karydas A, et al: **Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins.** *Nat Med* 2007, **13:**1359-1362.

45. Song F, Poljak A, Smythe GA, Sachdev P: **Plasma biomarkers for mild cognitive impairment and Alzheimer's disease.** *Brain Res Rev* 2009, **61:**69-80.

46. Zhang R, Barker L, Pinchev D, Marshall J, Rasamoelisolo M, Smith C, Kupchak P, Kireeva I, Ingratta L, Jackowski G: **Mining biomarkers in human sera using proteomic tools.** *Proteomics* 2004, **4:**244-256.

Chapter 2

SERUM LEVELS OF PREGNANCY ZONE PROTEIN ARE ELEVATED IN PRESYMPTOMATIC ALZHEIMER'S DISEASE

L. IJsselstijn, L.J.M. Dekker, C. Stingl, M.M. van der Weiden, A. Hofman, J.M. Kros, P.J. Koudstaal, P.A.E. Sillevis Smitt, M.A. Ikram, M.M.B. Breteler, T.M. Luider.
Journal of Proteome Research, 2011, 10(11):4902-4910.

ABSTRACT

We have sought for disease-related proteins which could predict the onset of Alzheimer's disease (AD) in a study population derived from the Rotterdam Scan Study, a population based prospective cohort study, designed to investigate the etiology and natural history of age-related brain changes in the elderly. The serum proteome of 43 persons who developed AD, after an average 4.2 years (±2.6 years SD) after blood sampling, and 43 gender and age matched controls who remained dementia free during follow-up, was investigated by liquid chromatography mass spectrometry. We identified 61 differentially expressed peptides between presymptomatic AD and controls, 9 of which were derived from pregnancy zone protein (PZP). Quantitative measurements using a multiple reaction monitoring assay showed a significant increase in concentration of PZP in presymptomatic AD (34.3 ± 20.6 mg/l) compared with controls (23.6 ± 13.6 mg/l) ($p = 0.006$). The difference in PZP was significant in women. Immunohistochemical validation of the findings on brain tissue sections showed strong PZP expression in senile plaques and in microglial and glial cells in AD with only low expression in some scattered glial cells in controls.

INTRODUCTION

It is estimated that there are currently 35.6 million people with dementia worldwide and that this number will nearly double every 20 years, to 65.7 million in 2030 and 115.4 million in 2050[1]. Dementia mainly affects elderly; after 65 years of age the prevalence doubles with every five-year increase in age. Alzheimer's disease (AD) represents 50 to 75% of all dementias. In people with AD, the increasing impairment of learning and memory eventually leads to the diagnosis. The pathological characteristics of AD (amyloid-β containing senile plaques and neurofibrillary tangles) may occur decades before the onset of the clinical symptoms[2]. At present, five drugs have been approved by the United States Food and Drugs Administration (US FDA) to treat AD[3]. Four of these are acetylcholinesterase inhibitors and the fifth is an NMDA antagonist. Because these drugs can delay symptom progression, but are not curative, it is important to initiate treatment as early as possible. Therefore, early detection in the presymptomatic phase of the disease is of paramount importance. So far, there is no test to detect people with presymptomatic AD.

Because of the large-scale accessibility and applicability of blood sampling, the identification of a (panel of) serum biomarker(s) for presymptomatic AD would be ideal. Several proteomics studies in plasma or serum have identified proteins as potential biomarkers for AD[4-12] and even very early stages of AD[13, 14], but not for the presymptomatic condition. For example, clusterin levels in plasma were associated with the progression and severity of AD[10] but were not helpful in discovering presymptomatic AD[15].

In this study, we have sought to identify serum biomarkers associated with presymptomatic AD. We used blood samples of people who were enrolled in a population-based prospective cohort study designed to detect determinants of physical and psychological health and their predictors in the elderly. As part of this survey, blood samples were taken at regular points in time. In addition, repetitive neuropsychological testing was part of the survey protocol. From this population, the serum proteome of 43 persons who later developed AD and 43 gender and age matched controls was investigated by various mass spectrometric techniques.

MATERIALS AND METHODS

Study population

The study population was derived from the Rotterdam Scan Study (RSS), a population based prospective cohort study initiated in 1995-1996, which included 1077 non-

demented healthy subjects aged between 60 to 90 years. Participants in the RSS were randomly selected from two ongoing population based studies, the Rotterdam Study[16] and the Zoetermeer Study[17]. The Erasmus University Medical Ethics Committee approved the study and written informed consent was obtained from all participants.

The RSS was designed to study the etiology and natural history of age-related brain changes in the elderly. At baseline (1995-1996), information was obtained on socio-demographic characteristics, medical history, current health status, medication use and determinants for chronic diseases. In addition, participants underwent brain imaging, veni-puncture and several physical examinations including a neurological and cognitive screening. Cognitive screening included the Mini Mental State Examination (MMSE) and the Geriatric Mental State (GMS) schedule. Apolipoprotein E (APOE) genotyping was performed on coded genomic DNA, as described in detail elsewhere[18]. Follow-up re-examinations took place in 1997-1999 and 2002-2004, and included the same standardized interviewing, veni-puncture and physical examination.

Of the 1077 subjects 43 developed AD (4%) during the 10-year follow-up period according to the AD assessment protocol described elsewhere[19]. Persons that were screened positive for MMSE (< 26) and GMS (> 0) during follow-up, were assessed with the Cambridge Examination for Mental Disorders of the Elderly interview (CAMDEX). Persons suspected to have AD were further examined by a neurologist and underwent additional neuropsychological testing by a neuropsychologist. Besides the above described tests for AD assessment, the memory function of all subjects was also tested by means of a 15-word verbal learning test, the Stroop test and the letter-digit substitution test. These tests were used to determine and exclude other cognitive stages like mild cognitive impairment (MCI). For each AD case we randomly selected a control, matched on gender, and in strata of age (5 years). The selected controls had no cognitive problems, remained dementia-free and were still alive at the end of follow-up (2005).

The blood samples were taken in the Research Centers of the Rotterdam Study and the Zoetermeer Study between 9.00 AM and 4.00 PM. After coagulation for 30 minutes serum was obtained by centrifugation for 10 minutes at 3,000 rpm (Microcentrifuge 5417R; Eppendorf, Hamburg, Germany) at room temperature. Serum samples were aliquoted into 1 ml portions, fast frozen at -196 °C in liquid nitrogen, transferred to the laboratory and subsequently stored at -80 °C until further use.

To prepare the serum samples for liquid chromatography mass spectrometry (LC-MS), an enzymatic digestion was performed as described previously[15]. Briefly, neat serum was diluted 500 times in 0.1% RapiGest SF (Waters, Milford, MA), reduced using dithiothreitol and alkylated using iodoacetamide. Subsequently, trypsin was added for digestion. After incubation the pH was reduced to terminate the digestion reaction and the serum digest were aliquoted (25 µl) and stored at -20 °C until further use.

Nano LC Orbitrap MS

LC-MS measurements were carried out on a nano LC system (Ultimate 3000; Dionex, Amsterdam, Netherlands) online coupled to a hybrid linear ion trap/Orbitrap MS (LTQ Orbitrap XL; Thermo Fisher Scientific, Bremen, Germany). 1 µl of serum digest was injected onto the nano LC system, which held a C18 trap column (PepMap C18, 300 µm ID × 5 mm, 5 µm particle size and 100 Å pore size; Dionex) and a 50 cm long analytic column (PepMap C18, 75 µm ID × 500 mm, 3 µm particle size and 100 Å pore size; Dionex). A 180-minute gradient with a 300 nl/min flow was run with solvent A (H2O/acetonitrile (ACN) 98/2 (v/v), 0.1% formic acid (FA)) and solvent B (H2O/ACN 20/80 (v/v), 0.1% FA): 0-25% solvent B in 120 min and 25-50% solvent B in the next 60 min. All solvents used were purchased from Biosolve (Valkenswaard, Netherlands). The separation of the peptides was monitored by a UV detector (absorption at 214 nm).

High resolution full scan MS was obtained from the Orbitrap (resolution 30,000; AGC 1,000,000); MS/MS spectra were obtained by CAD fragmentation. MS/MS was performed on the top five masses in the full scan spectra. Dynamic exclusion was used, with a repeat count of one; exclusion duration was set at 3 min and exclusion width at ±5 ppm.

To check the reproducibility of the nano LC system and the performance of the Orbitrap MS, one of the serum digests was re-measured every ten runs as a technical control.

Data analysis was done using the software package Progenesis LC-MS (version 2.5; Nonlinear Dynamics Ltd, New Castle, UK). This package processes raw data files by peak modeling, LC-MS run alignment, peptide detection, and normalization of peak areas to the total ion current[20]. After peptide detection a filter was used, yielding only peptides with charges two to eight and more than two isotope peaks.

For peptide identification, MSMS data was searched against the human SwissProt database using Mascot (version 2.2.06; Matrix Science Inc., London, UK). The following settings were used; enzyme: trypsin, fixed modifications: carbamidomethyl (C), variable modifications: oxidation (M), mass values: monoisotopic, protein mass: unrestricted, peptide mass tolerance: ±10 ppm (# $^{13}C = 2$), fragment mass tolerance: ±0.5 Da, max missed cleavages: 2, instrument type: ESI-TRAP. The identifications were imported in Progenesis and a matrix was generated containing the following information for each peptide; retention time, normalized peak area and when present its sequence.

Quantification of Pregnancy Zone Protein

Fifteen unique peptides of pregnancy zone protein (PZP; Uniprot accession number P20742) were identified. For one of the peptides, VVVQTESGGR (primary amino acid position 204-213) a multiple reaction monitoring (MRM) assay was developed according to the method described previously[15] (for peptide selection, see supplementary data). A synthetic version

of the selected peptide together with its internal standard, VVGVQTESGGR, were obtained from Pepscan Presto (Lelystad, Netherlands) with a peptide purity > 95%. The selected peptide and its internal standard showed comparable intensities and ion efficiency, only a slight shift in retention time was observed (0.8 min). Calibration curves were made for both synthetic peptides in digested serum. The following concentrations of synthetic peptide were measured: 0, 0.16, 0.8, 4, and 20 fmol/µl. Each concentration point for the calibration curve was measured in triplicate. The limit of detection (LOD) and the lower limit of quantification (LLOQ) were determined. The LOD was defined as three times the signal-to-noise ratio of an unspiked serum digest and the LLOQ as the lowest calibration point of the curve that could be measured with a CV < 20%.

Serum digests of case-control samples were spiked with 20 fmol/µl of internal standard. Chromatographic separation of spiked serum digests, was performed on a nano LC system (Ultimate 3000; Dionex, Amsterdam, Netherlands). Two µl of spiked serum digest was loaded onto a C18 trap column (PepMap C18, 300 µm ID × 5 mm, 5 µm particle size and 100 Å pore size; Dionex) and washed for 5 min at a flow rate of 20 µL/min 0.1% TFA in H2O. Next, the trap column was switched in line with the analytic column (PepMap C18, 75 µm ID × 150 mm, 3 µm particle size and 100 Å pore size; Dionex). Peptides were eluted at a flow rate of 300 nl/min with the following gradient: 0-45% solvent B in 30 min, solvent A (H2O/acetonitrile (ACN) 98/2 (v/v), 0.1% formic acid (FA)) and solvent B (H2O/ACN 20/80 (v/v), 0.1% FA). MRM detection was performed by means of a triple quadruple tandem mass spectrometer (4000 QTRAP; AB SCIEX, Concord, Ontario, Canada) in the positive ion mode.

For each transition, the peak area was determined using the program Skyline v0.6[21]. The concentration of PZP in the serum digests was calculated using the ratio of the original peptide to the internal standard.

Immunohistochemical staining

For immunohistochemical staining four brains of subjects with AD and two control brains were taken from the files of the Department of Pathology, Erasmus Medical Center, Rotterdam, The Netherlands. The brains are not from persons who participated in the RSS study and whose serum proteome was examined in this study. The four brains with AD came from male subjects (including one patient with a familial form of AD) while the controls came from females. The brain samples had been fresh-frozen. Immunostaining was performed on frontal cortex and hippocampus in all cases.

Sections of 5 µm from each sample were made and staining was performed following the manufacturer's instructions (alkaline phosphatase technique) using rabbit polyclonal antibody for PZP (GeneTex, Irvine, USA). Briefly, sections were fixed in acetone for 15

minutes and then air-dried. The sections were washed with phosphate-buffer saline (PBS) and incubated with the antibody for 30 minutes. After washing the sections with PBS, the secondary antibody was added and incubated 30 minutes at room temperature. Afterwards, the sections were washed with tap water, counterstained and coverslipped with permanent mounting medium. The sections were examined by the neuropathologist (JMK) for the presence and spatial localization of PZP in the tissue.

Statistical analysis

To test for homogeneity between the presymptomatic AD patients and the cognitive normal controls, a Mann-Whitney test was used with a significant level of 0.05. A non-parametric test was used, since not all variables were normally distributed. To determine whether peptides present in sera from presymptomatic AD patients were significantly higher or lower than that in sera from cognitive normal controls, the normalized peak areas of the nano LC Orbitrap MS measurements were compared using a two-sided t-test with a significance level of 0.05. The same t-test was used in the MRM assay to test for differences in quantitative levels of PZP between presymptomatic AD and controls. To determine the influence of the absence or presence of one or two APOE ε4 alleles on the PZP levels measured, an ANOVA was used followed by a t-test for the separate comparisons. All statistical analyses were performed using the SPSS statistical software package (version 15.0).

RESULTS

Subject characteristics

At the time that the serum samples were collected for this study, no differences between cases and controls were observed in age and gender (table 2.1). Although the Mini Mental State Examination (MMSE) scores were above 25 for both groups, the cases scored significantly lower (26) compared with the controls (27). In addition, the cases were more likely to carry one or two APOE ε4 alleles. The diagnosis of AD was made on average 4.2 ± 2.6 (SD) years after the blood samples were taken.

Table 2.1. Subject characteristics.

	Subjects		
	Cases	**Controls**	**p-value[a]**
Number	43	43	1
Mean age (years; SD)	78 (6.5)	78 (6.8)	0.88
Female/Male	32/11	32/11	1
Mean MMSE[b] score (SD)	26 (2.2)	27 (2.7)	0.01
Mean duration until AD diagnosis (years; SD)	4.2 (2.6)	0 (0)	0
APOE ε4/- (%)	15 (35%)	5 (12%)	0.01
APOE ε4/4 (%)	5 (12%)	0 (0%)	0.02

[a]p-value, a Mann-Whitney test was used to test for homogeneity between cases and controls; [b]MMSE, Mini Mental State Exam, maximum score 30.

Nano LC Orbitrap MS

A technical control serum digest, that was measured every ten samples, was used to determine the coefficient of variance (CV) as an indication of LC reproducibility. The CV was calculated using the retention time of eight peptides belonging to eight different proteins. The peptides were present in all runs and had retention times evenly distributed throughout the run. The control serum digest showed an average % CV of 3.2 ± 1.6 (SD) for the retention time (in minutes), indicating good reproducibility.

The nano LC Orbitrap MS measurements detected a total of 47,715 features. Of these features, 6,070 were identified as peptides. The total number of unique peptides in this sample set was 2,370 belonging to 377 proteins (179 proteins were identified with a minimum of two peptides).

Statistical analysis of the 2,370 identified peptides revealed 61 peptides that discriminated between presymptomatic AD and controls (t-test $p < 0.05$, table 2.2). Eighty-four percent (51/61) of the discriminating peptides were more abundant in presymptomatic AD than in controls. Proteins found with more than two discriminating peptides were pregnancy zone protein (PZP; 9 of 15 peptides), apolipoprotein B-100 (6 of 162 peptides) and complement C2 (3 of 13 peptides). Since PZP was the only protein of which the majority of its identified peptides (9 of 15) differed significantly between presymptomatic AD and controls this protein was analyzed further. Furthermore, all 15 identified PZP peptides were more abundant in presymptomatic AD (figure 2.1) and this was not the case for all identified peptides of apolipoprotein B-100 and complement C2 (not shown).

Table 2.2. Information of identified peptides that were able to discriminate between presymptomatic AD and controls based on nano LC Orbitrap MS measurements.

Peptide sequence	p-Value	More abundant in	Accession number	Protein
ASPAFLASQNTK	0.0047	Presymptomatic AD	P20742	Pregnancy zone protein
ISEITNIVSK	0.0052	Presymptomatic AD	P20742	Pregnancy zone protein
IREEGTDLEVTANR	0.0117	Presymptomatic AD	P20742	Pregnancy zone protein
VLNCDKQEVCEEFSQQLNSNGCITQQVHTK	0.0163	Presymptomatic AD	P20742	Pregnancy zone protein
TLLVEAEGIEQEK	0.0237	Presymptomatic AD	P20742	Pregnancy zone protein
DLFHCVSFTLPR	0.0288	Presymptomatic AD	P20742	Pregnancy zone protein
VQTVPQTCDGHK	0.0422	Presymptomatic AD	P20742	Pregnancy zone protein
SLFTDLVAEK	0.0450	Presymptomatic AD	P20742	Pregnancy zone protein
QGIPFFAQVLLVDGK	0.0452	Presymptomatic AD	P20742	Pregnancy zone protein
LPQQANDYLNSFNWER	0.0184	Presymptomatic AD	P04114	Apolipoprotein B-100
INNQLTLDSNTK	0.0189	Presymptomatic AD	P04114	Apolipoprotein B-100
IADFELPTIIVPEQTIEIPSIK	0.0265	Presymptomatic AD	P04114	Apolipoprotein B-100
EIFNMAR	0.0323	Presymptomatic AD	P04114	Apolipoprotein B-100
CVQSTKPSLMIQK	0.0377	Presymptomatic AD	P04114	Apolipoprotein B-100
SISAALEHK	0.0490	Presymptomatic AD	P04114	Apolipoprotein B-100
EVVTDQFLCSGTQEDESPCKGESGGAVFLER	0.0096	Presymptomatic AD	P06681	Complement C2
GALISDQWVLTAAHCFR	0.0369	Presymptomatic AD	P06681	Complement C2
DFHINLFR	0.0477	Presymptomatic AD	P06681	Complement C2
MYYSAVDPTK	0.0167	Presymptomatic AD	P00450	Ceruloplasmin
SVPPSASHVAPTETFTYEWTVPK	0.0463	Presymptomatic AD	P00450	Ceruloplasmin
LESEETMVLEAHDAQGDVPVTVTVHDFPGK	0.0329	Presymptomatic AD	P01024	Complement C3
LKGPLLNK	0.0396	Presymptomatic AD	P01024	Complement C3
IDTQDIEASHYR	0.0050	Presymptomatic AD	P01031	Complement C5
LSMDIDVSYK	0.0197	Presymptomatic AD	P01031	Complement C5
GGSGGSHGGGSGFGGESGGSYGGGEEASGSGGGYGGGSGK	0.0046	Presymptomatic AD	P35527	Keratin, type I cytoskeletal 9
GGGGSFGYSYGGGSGGGFSASSLGGGFGGGSR	0.0166	Presymptomatic AD	P35527	Keratin, type I cytoskeletal 9
TNAENEFVTIKK	0.0122	Presymptomatic AD	P04264	Keratin, type II cytoskeletal 1

Table 2.2. (continued)

Peptide sequence	p-Value	More abundant in	Accession number	Protein
SLVNLGGSK	0.0419	Presymptomatic AD	P04264	Keratin, type II cytoskeletal 1
DFTCVHQALK	0.0094	Presymptomatic AD	P05155	Plasma protease C1 inhibitor
LEDMEQALSPSVFK	0.0421	Presymptomatic AD	P05155	Plasma protease C1 inhibitor
CCAAADPHECYAK	0.0125	Controls	P02768	Serum albumin
FKDLGEENFK	0.0429	Presymptomatic AD	P02768	Serum albumin
DCKDVDECSLKPSICGTAVCK	0.0290	Presymptomatic AD	P07225	Vitamin K-dependent protein S
NIPGDFECECPEGYR	0.0449	Presymptomatic AD	P07225	Vitamin K-dependent protein S
CLAPLEGAR	0.0468	Presymptomatic AD	P04217	Alpha-1B-glycoprotein
DNSVHWERPQKPK	0.0266	Controls	P01023	Alpha-2-macroglobulin
GDECQLCEVENR	0.0297	Presymptomatic AD	O75882	Attractin
EEIIYECDK	0.0344	Presymptomatic AD	P04003	C4b-binding protein alpha chain
ALLAFQESK	0.0137	Presymptomatic AD	P20851	C4b-binding protein beta chain
SGYLLHGSNEITCNR	0.0231	Presymptomatic AD	P05160	Coagulation factor XIII B chain
GDSGGAFAVQDPNDK	0.0384	Presymptomatic AD	P09871	Complement C1s subcomponent
ECDNPAPQNGGASCPGR	0.0388	Controls	P07357	Complement component C8 alpha chain
LLCNGDNDCGDQSDEANCRR	0.0245	Controls	P07358	Complement component C8 beta chain
KEAGIPEFYDYDVALIK	0.0266	Presymptomatic AD	P00751	Complement factor B
WSAGLTSSQVDLYIPK	0.0304	Presymptomatic AD	P08185	Corticosteroid-binding globulin
LALWEGR	0.0499	Controls	Q9Y4F1	FERM, RhoGEF and pleckstrin domain-containing protein 1
NSLFEYQK	0.0249	Presymptomatic AD	P02671	Fibrinogen alpha chain
QYNVGPSVSK	0.0074	Presymptomatic AD	P02751	Fibronectin
LLVVYPWTQR	0.0270	Controls	P68871	Hemoglobin subunit beta
DIVMTQSPDSLAVSLGER	0.0498	Presymptomatic AD	P01625	Ig kappa chain V-IV region Len
IYGNQDTSSQLKK	0.0065	Controls	P19823	Inter-alpha-trypsin inhibitor heavy chain H2
AIGGGLSSVGGGSSTIK	0.0206	Controls	P02538	Keratin, type II cytoskeletal 6A
YNSQNQSNNQFVLYR	0.0450	Presymptomatic AD	P01042	Kininogen-1
NALTGLPPGLFQASATLDTLVLK	0.0350	Presymptomatic AD	P02750	Leucine-rich alpha-2-glycoprotein

Table 2.2. (continued)

Peptide sequence	p-Value	More abundant in	Accession number	Protein
LGTSKDLQPR	0.0281	Controls	Q9BQI9	Nuclear receptor-interacting protein 2
GFQQLLQELNQPR	0.0450	Presymptomatic AD	P05154	Plasma serine protease inhibitor
CEEDEEFTCR	0.0340	Presymptomatic AD	P00747	Plasminogen
CLKDGAGDVAFVK	0.0355	Controls	P02787	Serotransferrin
VTQVYAENGTVLQGSTVASVYK	0.0473	Presymptomatic AD	P27169	Serum paraoxonase/arylesterase 1
GGTLSTPQTGSENDALYEYLR	0.0257	Controls	P05452	Tetranectin
SYLSMVGSCCTSASPTVCFLK	0.0407	Presymptomatic AD	P02774	Vitamin D-binding protein

For each identified peptide the p-value is given together with the group in which the peptide was more abundant. Per peptide also the protein from which it originated and its accession number is listed.

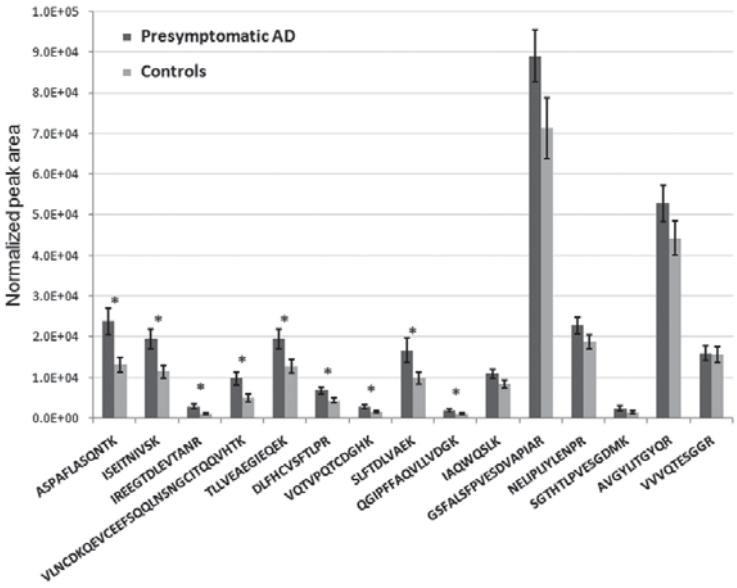


Figure 2.1. Normalized peak areas of identified peptides specific for pregnancy zone protein measured with nano LC Orbitrap MS. Peptides are arranged in order of increasing p-value. Dark bars, presymptomatic AD (n = 43); light bars, controls (n = 43); error bars, standard error (SE); * significant difference between presymptomatic AD and controls ($p < 0.05$).

Quantitative levels of PZP

Calibration curves of the original peptide of PZP (VVVQTESGGR) and its internal standard (VVGVQTESGGR), showed comparable slopes and intercepts. The developed MRM assay had a limit of detection of 0.12 fmol/µl and a lower limit of quantification of 0.16 fmol/µl. The expected concentration of PZP in control serum is around 40 mg/l (in female)[22], corresponding to 0.24 fmol/µl in our serum digests.

The MRM assay showed a significantly higher concentration of PZP in presymptomatic AD (34.3 ± 20.6 mg/l) compared with controls (23.6 ± 13.6 mg/l) ($p = 0.006$). We found that the difference in PZP was significant in women but not in men. In women, PZP levels were 41.7 ± 18.8 mg/l for presymptomatic AD (n = 32) and 26.5 ± 14.0 mg/l for controls (n = 32) ($p < 0.005$) whereas in men PZP levels were 12.8 ± 3.1 mg/l for presymptomatic AD (n = 11) and 15.1 ± 8.3 mg/l for controls (n = 11) ($p = 0.40$) (figure 2.2).

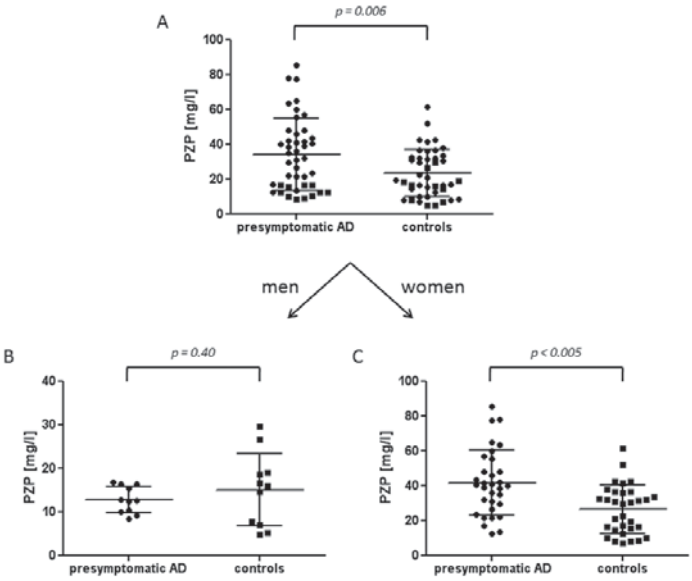


Figure 2.2. PZP concentrations measured by LC-MRM. (A) PZP concentrations in presymptomatic AD and controls. For both groups the mean is indicated together with the standard deviation (error bars). A two-sided t-test shows that there is a significant difference between both groups ($p = 0.006$). The sample set was divided into men (B) and women (C). In men, there is no significant difference ($p = 0.40$). In women, there is a significant difference in PZP levels between presymptomatic AD and controls ($p < 0.005$).

When the presymptomatic AD cases were grouped based on the presence of no, one or two APOE ε4 alleles, no significant difference (ANOVA, $p = 0.22$) was observed between the different groups. A (non-significant) difference in PZP concentration was observed between persons with no or one APOE ε4 alleles present and persons with two APOE ε4 alleles present (t-test, $p = 0.08$ and $p = 0.12$) (figure 2.3).

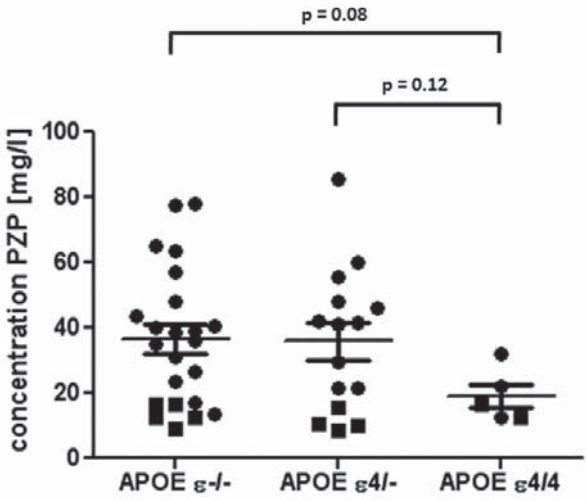


Figure 2.3. Association between PZP concentrations in presymptomatic AD patients and APOE ε4 genotype. An ANOVA shows that there is no significant difference in PZP levels between persons with no, one or two APOE ε4 alleles present (p = 0.22). A non-significant difference is observed between persons with no or one APOE ε4 alleles and persons with two APOE ε4 alleles (t-test, p = 0.08 and p = 0.12). APOE ε-/-, no ε4 alleles present (n = 23); APOE ε4/-, one ε4 allele present (n = 15); APOE ε4/4, two alleles present (n = 5). For all groups the mean is indicated together with the standard deviation (error bars). Squares indicate males while circles indicate females.

Immunohistochemistry

In figure 2.4 immunostaining of PZP is illustrated in cerebral cortex of four AD patients and two control patients. In the AD brains strong immunopositivity was seen in (micro) glial cells and in senile plaques. In the cerebral cortex of the control patients only some scattered glial cells were immunopositive. PZP staining in the cerebral cortex of a familial AD patient (figure 2.5) also showed immunopositivity in senile plaques, microglial and glial cells with additional immunoreactivity around blood vessels.

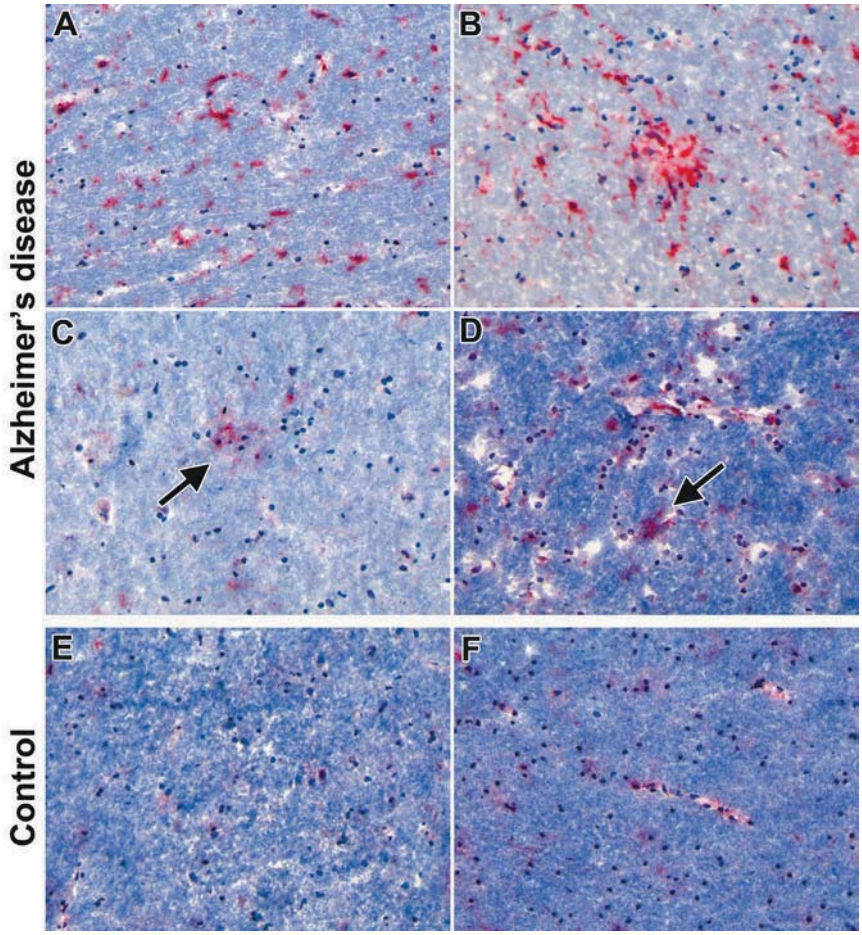


Figure 2.4. Immunohistochemical analysis of PZP expression in AD and normal brain sections. (A-D) Cerebral cortex of patients with AD, stained with the PZP antibody, shows immunopositivity in microglial cells and in senile plaques (arrows). In addition, some adventitial immunopositivity is present (x 200). (E-F) Cerebral cortex of control patients (cause of death cardiovascular arrest), stained with the PZP antibody. There is immunopositivity in some scattered glial cells (x200).

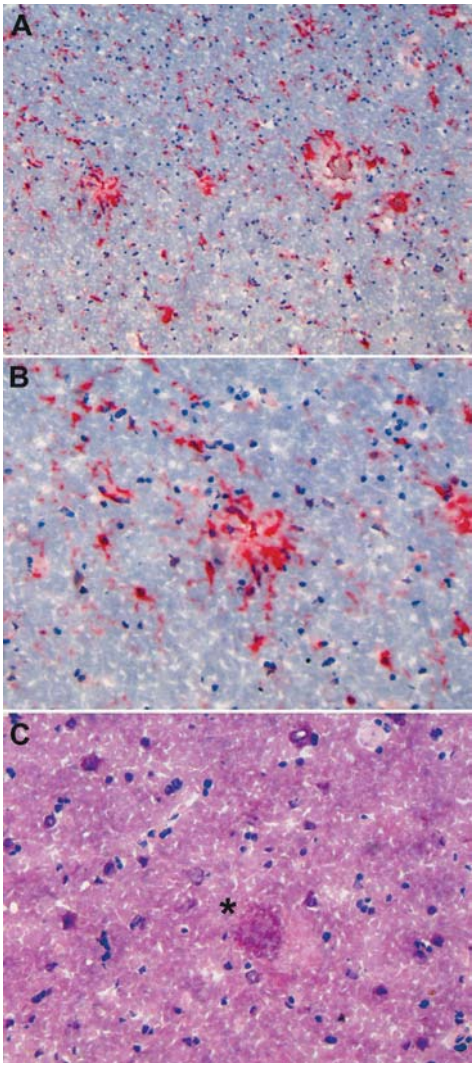


Figure 2.5. Immunohistochemical analysis of PZP expression in familial AD brain sections. (A) Cerebral cortex of a patient with familial AD shows PZP immunopositivity in the senile plaques, microglial and glial cells, and around blood vessels (x 20). (B) Magnification of A: detail of the immunoreactivity of the senile plaque and the microglial and glial cells (x 200). (C) Section adjacent slide to (B) shows the senile plaque (asterix) and microglial and glial cells in H&E staining (x 200).

DISCUSSION

In this population-based study, we found that pregnancy zone protein (PZP) levels were significantly higher in persons who were to develop AD compared with gender and age matched controls who remained dementia free. The difference was significant in women, not in men. Immunohistochemical validation of the findings on brain tissue sections showed strong PZP expression in senile plaques and in microglial and glial cells in AD with only low expression in some scattered glial cells in controls.

Elevated PZP levels

PZP is one of the major pregnancy-associated plasma proteins. In serum of healthy non-pregnant persons PZP concentration is low (women, 10-40 mg/l; men, <10 mg/l)[22]. With advancing gestation, circulating levels of PZP increase up to 1000 mg/l and return to pre-pregnancy levels within a few weeks after delivery[22, 23]. Increased levels of PZP also occur after oestrogen administration in both sexes[24]. Folkersen *et al*. reported a slight, but significant, age-dependent increase in both healthy males and females[22]. Confounders of gender and age were minimized in the present study by selecting matching controls. PZP concentrations in men were approximately 3 to 4 times lower than in women, as described in the literature. The sex difference in serum PZP concentration prompted us to compare the values separately for men and women (figure 2.2) and a clear difference between presymptomatic AD and controls was found only for females. Apart from the generally lower concentrations of PZP in men, a possible explanation may be that the number of males in this study was small. Further, the range of quantification of the MRM assay used in this study did not reach lower than 0.16 fmol/µl corresponding to a concentration of 25 mg/l. Since the majority of concentrations measured in men were below this value, it may well be that a significant difference is also present in males. Furthermore, immunohistochemistry showed the presence of PZP in AD brain, which were all from males. At this time point no female AD brains were available in our archives. Further immunohistochemical studies will be conducted on prospectively collected tissue from other institutes, to determine whether a difference in PZP expression in male and female AD brain is present. To our knowledge, nothing is known in the literature about PZP expression in the brain and the capability of PZP crossing the blood-brain barrier. The only related literature is by Chiabrando *et al*.[25] showing that methylamine-activated PZP was capable of inhibiting nerve growth factor-promoted neurite extension in cell cultures, while normal PZP had little or no effect. The only indication we have about the presence of PZP in the central nervous system, comes from a researcher within our group who has occasionally observed PZP peptides in cerebrospinal fluid (not published).

Since our study has a long and virtually complete follow-up, it was possible to search for a relation between the measured PZP concentrations and the period between blood sampling and the diagnosis of AD. We found no correlation between PZP concentrations and the time to diagnosis (4.2 ± 2.6 years).

Possible link PZP and AD

To the best of our knowledge, a relation between PZP and AD has never been reported in the literature. The primary structure of PZP is remarkably similar to alpha-2-macroglobulin (A2MG). Of all corresponding amino acid residues 71% are identical in both proteins and this increases to 78% when chemically similar residues are included (V = I = L, K = R, T = S, E = D, F = Y, G = A, Q = N)[26]. Both proteins are capable of inhibiting serine proteinases. When a bait region in the protein is cleaved by a proteinase, a conformational change is induced in the protein which traps the proteinase. Since the bait-region sequence appears to be the region with lowest similarity between PZP and A2MG, there are probably different or complementary functions for the proteinase inhibition exerted by both proteins[27]. The close link between PZP and A2MG is interesting since A2MG has been implicated in AD. Hye et al. found elevated levels of A2MG in plasma of AD patients by comparing two-dimensional gel images of cases and controls[7]. A2MG is also found in plaques in the brain of AD patients, where it forms complexes with amyloid-β[28]. In the present study the presence of PZP was immunohistochemically detected in and around blood vessels, in microglial and glial cells, and also, like A2MG, in plaques present in the brains of the AD patients (figure 2.4 and 2.5). Importantly, the used antibody was raised against a unique region in PZP which is not present in A2MG. It has been suggested that A2MG plays a role in amyloid-β clearance via the low-density lipoprotein receptor-related protein (LRP)[29]. LRP is widely studied in relation to AD, since three of the susceptibility genes, amyloid precursor protein, apolipoprotein E and A2MG, encode proteins that are ligands for LRP[30]. LRP is also responsible for the clearance of PZP-proteinase complexes[31]. Although PZP and A2MG are highly similar, a difference in binding properties to LRP is reported[27]. An important role in the different binding properties could be that A2MG exists exclusively in a tetrameric form where PZP exists predominantly as a dimer. The close resemblance in structure and function of these proteins and the much lower concentrations of PZP may well explain that PZP was not previously detected in AD. The fact that apolipoprotein E (APOE) is also a ligand for LRP is very interesting, since we noticed a difference in PZP levels between subjects with two, and with no or only one, APOE ε4 allele (figure 2.3). The APOE ε4 allele is a strong susceptibility factor for development of AD[32]. Risk increases 3-fold in persons with one APOE ε4 allele and about 9- to 12-fold in persons with two APOE ε4 alleles as compared to persons with no APOE ε4 alleles[33]. We observed lower

PZP levels in persons with two APOE ε4 alleles, which was the opposite effect of what one would expect since these persons have an increased risk for developing AD. The small number of persons in the presymptomatic AD group with two APOE ε4 alleles, however, does not allow any conclusion on the relationship between PZP and APOE.

In summary, our search for proteins related to presymptomatic AD resulted in the discovery of PZP. PZP levels were significantly higher in persons who later developed AD compared to gender and age matched controls who remained dementia free. The developed MRM assay confirmed this finding. The difference was significant in women, not in men. The resemblance of PZP to A2MG and the fact that PZP is observed by immunohistochemistry in brain of AD patients is intriguing. Further studies are warranted to unravel the role of PZP in AD.

REFERENCES

1. Prince M, Jackson J: **World Alzheimer Report 2009.** 2009.

2. Braak H, Braak E: **Neuropathological stageing of Alzheimer-related changes.** *Acta Neuropathol* 1991, **82:**239-259.

3. Neugroschl J, Sano M: **Current treatment and recent clinical research in Alzheimer's disease.** *Mt Sinai J Med* 2010, **77:**3-16.

4. Cutler P, Akuffo EL, Bodnar WM, Briggs DM, Davis JB, Debouck CM, Fox SM, Gibson RA, Gormley DA, Holbrook JD, et al: **Proteomic identification and early validation of complement 1 inhibitor and pigment epithelium -derived factor: Two novel biomarkers of Alzheimer's disease in human plasma.** *Proteomics Clinical Applications* 2008, **2:**467-477.

5. German DC, Gurnani P, Nandi A, Garner HR, Fisher W, Diaz-Arrastia R, O'Suilleabhain P, Rosenblatt KP: **Serum biomarkers for Alzheimer's disease: proteomic discovery.** *Biomed Pharmacother* 2007, **61:**383-389.

6. Guntert A, Campbell J, Saleem M, O'Brien DP, Thompson AJ, Byers HL, Ward MA, Lovestone S: **Plasma gelsolin is decreased and correlates with rate of decline in Alzheimer's disease.** *J Alzheimers Dis* 2010, **21:**585-596.

7. Hye A, Lynham S, Thambisetty M, Causevic M, Campbell J, Byers HL, Hooper C, Rijsdijk F, Tabrizi SJ, Banner S, et al: **Proteome-based plasma biomarkers for Alzheimer's disease.** *Brain* 2006, **129:**3042-3050.

8. Liao PC, Yu L, Kuo CC, Lin CJ, Kuo YM: **Proteomics analysis of plasma for potential biomarkers in the diagnosis of Alzheimer's disease.** *Proteomics Clinical Applications* 2007, **1:**506-512.

9. Liu HC, Hu CJ, Chang JG, Sung SM, Lee LS, Yuan RY, Leu SJ: **Proteomic identification of lower apolipoprotein A-I in Alzheimer's disease.** *Dement Geriatr Cogn Disord* 2006, **21:**155-161.

10. Thambisetty M, Simmons A, Velayudhan L, Hye A, Campbell J, Zhang Y, Wahlund LO, Westman E, Kinsey A, Guntert A, et al: **Association of plasma clusterin concentration with severity, pathology, and progression in Alzheimer disease.** *Arch Gen Psychiatry* 2010, **67:**739-748.

11. Zhang R, Barker L, Pinchev D, Marshall J, Rasamoelisolo M, Smith C, Kupchak P, Kireeva I, Ingratta L, Jackowski G: **Mining biomarkers in human sera using proteomic tools.** *Proteomics* 2004, **4:**244-256.

12. Akuffo EL, Davis JB, Fox SM, Gloger IS, Hosford D, Kinsey EE, Jones NA, Nock CM, Roses AD, Saunders AM, et al: **The discovery and early validation of novel plasma biomarkers in mild-to-moderate Alzheimer's disease patients responding to treatment with rosiglitazone.** *Biomarkers* 2008, **13:**618-636.

13. Craig-Schapiro R, Perrin RJ, Roe CM, Xiong C, Carter D, Cairns NJ, Mintun MA, Peskind ER, Li G, Galasko DR, et al: **YKL-40: a novel prognostic fluid biomarker for preclinical Alzheimer's disease.** *Biol Psychiatry* 2010, **68:**903-912.

14. Ghidoni R, Benussi L, Glionna M, Desenzani S, Albertini V, Levy E, Emanuele E, Binetti G: **Plasma cystatin C and risk of developing Alzheimer's disease in subjects with mild cognitive impairment.** *J Alzheimers Dis* 2010, **22:**985-991.

15. IJsselstijn L, Dekker LJ, Koudstaal PJ, Hofman A, Sillevis Smitt P, Breteler MM, Luider T: **Serum Clusterin Levels are not increased in Presymptomatic Alzheimer's Disease.** *J Proteome Res* 2011, **10:**2006-2010.

16. Hofman A, Breteler MM, van Duijn CM, Janssen HL, Krestin GP, Kuipers EJ, Stricker BH, Tiemeier H, Uitterlinden AG, Vingerling JR, Witteman JC: **The Rotterdam Study: 2010 objectives and design update.** *Eur J Epidemiol* 2009, **24:**553-572.

17. Breteler MM: **Vascular involvement in cognitive decline and dementia. Epidemiologic evidence from the Rotterdam Study and the Rotterdam Scan Study.** *Ann N Y Acad Sci* 2000, **903:**457-465.

18. Slooter AJ, Cruts M, Kalmijn S, Hofman A, Breteler MM, Van Broeckhoven C, van Duijn CM: **Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study.** *Arch Neurol* 1998, **55:**964-968.

19. Prins ND, van Dijk EJ, den Heijer T, Vermeer SE, Koudstaal PJ, Oudkerk M, Hofman A, Breteler MM: **Cerebral white matter lesions and the risk of dementia.** *Arch Neurol* 2004, **61:**1531-1534.

20. http://www.nonlinear.com.

21. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ: **Skyline: an open source document editor for creating and analyzing targeted proteomics experiments.** *Bioinformatics* 2010, **26:**966-968.

22. Folkersen J, Teisner B, Grunnet N, Grudzinskas JG, Westergaard JG, Hindersson P: **Circulating levels of pregnancy zone protein: normal range and the influence of age and gender.** *Clin Chim Acta* 1981, **110:**139-145.

23. Ekelund L, Laurell CB: **The pregnancy zone protein response during gestation: a metabolic challenge.** *Scand J Clin Lab Invest* 1994, **54:**623-629.

24. Beckman L, von Schoultz B, Stigbrand T: **"Pregnancy zone" protein in sera from patients with prostatic cancer treated with oestrogens.** *Urol Res* 1973, **1:**67-69.

25. Chiabrando GA, Sanchez MC, Skornicka EL, Koo PH: **Low-density lipoprotein receptor-related protein mediates in PC12 cell cultures the inhibition of nerve growth factor-promoted neurite outgrowth by pregnancy zone protein and alpha2-macroglobulin.** *J Neurosci Res* 2002, **70:**57-64.

26. Devriendt K, Van den Berghe H, Cassiman JJ, Marynen P: **Primary structure of pregnancy zone protein. Molecular cloning of a full-length PZP cDNA clone by the polymerase chain reaction.** *Biochim Biophys Acta* 1991, **1088:**95-103.

27. Chiabrando GA, Vides MA, Sanchez MC: **Differential binding properties of human pregnancy zone protein- and alpha2-macroglobulin-proteinase complexes to low-density lipoprotein receptor-related protein.** *Arch Biochem Biophys* 2002, **398:**73-78.

28. Kovacs DM: **alpha2-macroglobulin in late-onset Alzheimer's disease.** *Exp Gerontol* 2000, **35:**473-479.

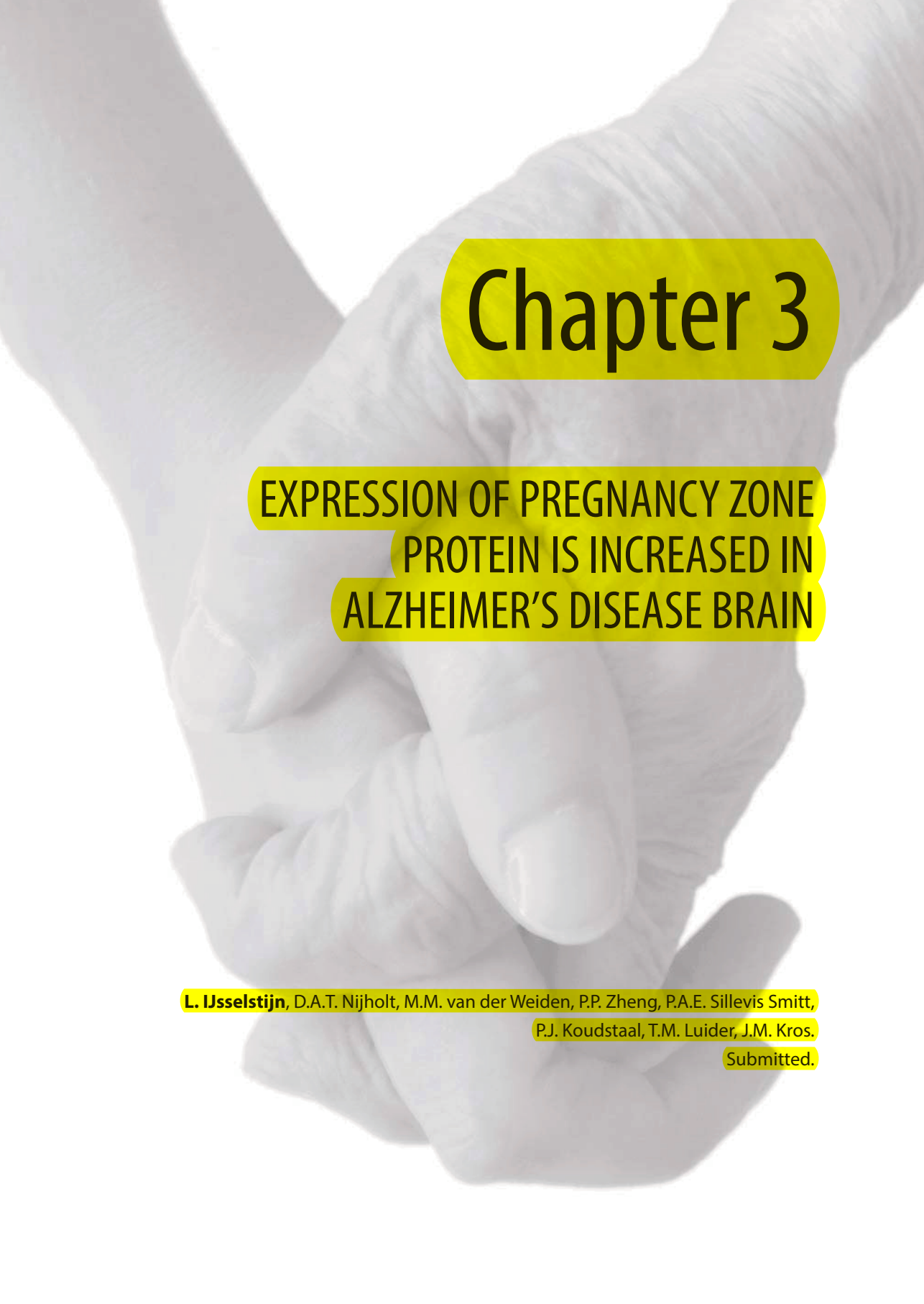
29. Van Uden E, Sagara Y, Van Uden J, Orlando R, Mallory M, Rockenstein E, Masliah E: **A protective role of the low density lipoprotein receptor-related protein against amyloid beta-protein toxicity.** *J Biol Chem* 2000, **275:**30525-30530.

30. Van Uden E, Kang DE, Koo EH, Masliah E: **LDL receptor-related protein (LRP) in Alzheimer's disease: towards a unified theory of pathogenesis.** *Microsc Res Tech* 2000, **50:**268-272.

31. Sanchez MC, Chiabrando GA, Vides MA: **Pregnancy zone protein-tissue-type plasminogen activator complexes bind to low-density lipoprotein receptor-related protein (LRP).** *Arch Biochem Biophys* 2001, **389:**218-222.

32. Raber J, Huang YD, Ashford JW: **ApoE genotype accounts for the vast majority of AD risk and AD pathology.** *Neurobiology of Aging* 2004, **25:**641-650.

33. Saunders AM, Strittmatter WJ, Schmechel D, Georgehyslop PHS, Pericakvance MA, Joo SH, Rosi BL, Gusella JF, Crappermaclachlan DR, Alberts MJ, et al: **Association of Apolipoprotein-E Allele Epsilon-4 with Late-Onset Familial and Sporadic Alzheimers-Disease.** *Neurology* 1993, **43:**1467-1472.

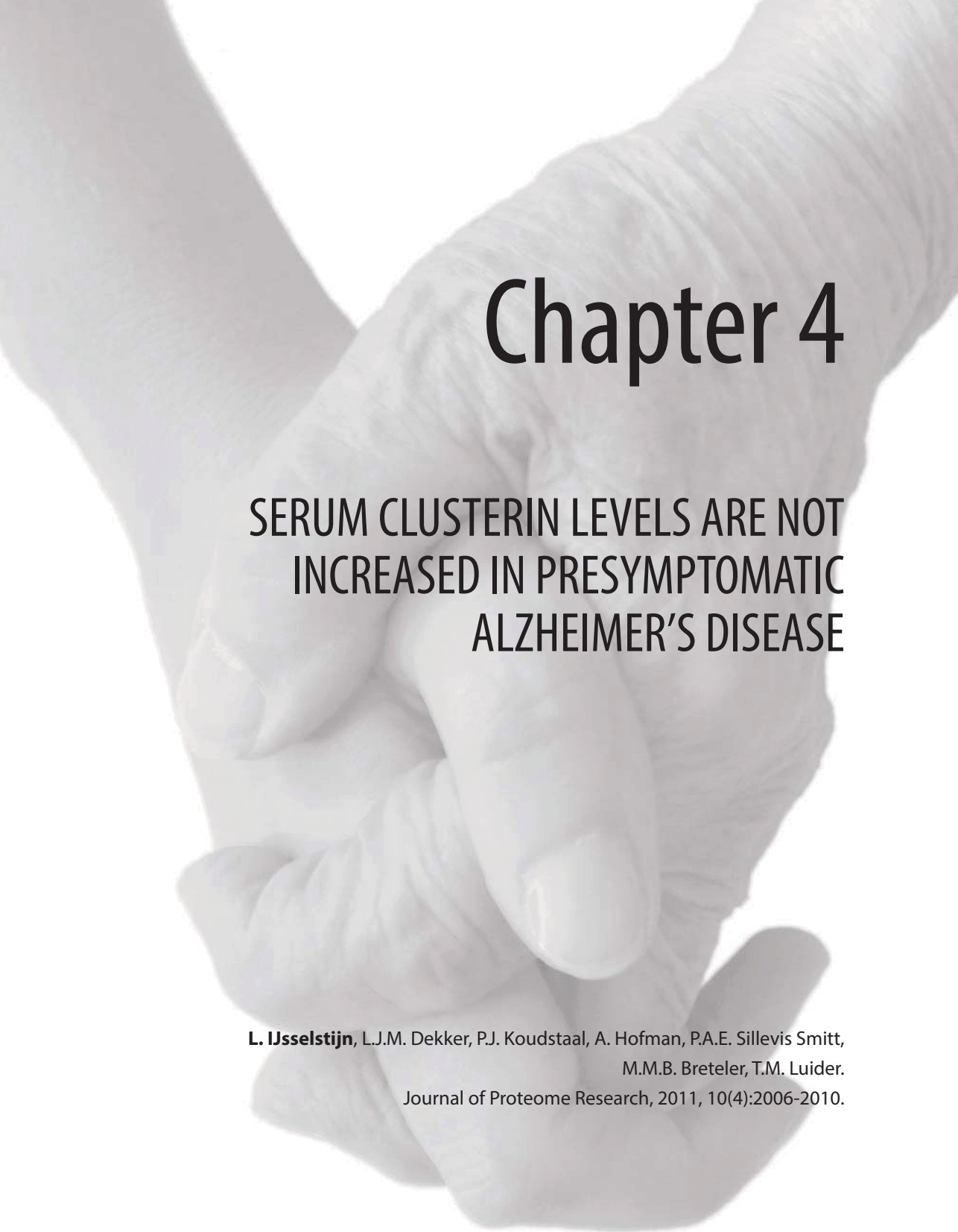


Chapter 3

EXPRESSION OF PREGNANCY ZONE PROTEIN IS INCREASED IN ALZHEIMER'S DISEASE BRAIN

L. IJsselstijn, D.A.T. Nijholt, M.M. van der Weiden, P.P. Zheng, P.A.E. Sillevis Smitt, P.J. Koudstaal, T.M. Luider, J.M. Kros.

Submitted.



Chapter 4

SERUM CLUSTERIN LEVELS ARE NOT INCREASED IN PRESYMPTOMATIC ALZHEIMER'S DISEASE

L. IJsselstijn, L.J.M. Dekker, P.J. Koudstaal, A. Hofman, P.A.E. Sillevis Smitt, M.M.B. Breteler, T.M. Luider.

Journal of Proteome Research, 2011, 10(4):2006-2010.

ABSTRACT

Increased plasma levels of clusterin have recently been found to be associated with severity and progression in Alzheimer's disease (AD). We have investigated clusterin levels in serum of elderly people with presymptomatic AD from a population based prospective cohort study. During 10 years follow-up, 43 participants were diagnosed with AD after on average 4.2 years (± 2.6 years SD) after the initial blood sampling. At the time of blood sampling these participants showed normal cognitive function. For each presymptomatic AD case, a control was matched on gender and age. Furthermore, the selected controls had to remain dementia-free and still be alive at the end of follow-up. Quantitative serum clusterin levels were measured with a newly developed multiple reaction monitoring (MRM) assay. Results of the assay showed no significant difference in clusterin levels between presymptomatic AD and controls (p-value 0.54). In conclusion, serum clusterin is not an early, presymptomatic biomarker for AD.

INTRODUCTION

In 1990, May *et al*.[1] first described increased expression of clusterin (or apolipoprotein J) in Alzheimer's disease (AD) hippocampus. Subsequently, several studies have demonstrated a role for clusterin in the aggregation of amyloid-β (Aβ) peptides[2-4]. It can either prevent Aβ oligomerization or enhance the formation of fibrillar structures, depending on the ratio of clusterin and Aβ-peptides[5]. Also, clusterin may participate in the amyloid-β clearance from the brain across the blood-brain barrier[6, 7].

In 2009, two genome-wide association studies independently identified an association of late-onset AD with a single-nucleotide polymorphism within an intron of the clusterin gene on chromosome 8[8, 9]. Following this genetic discovery, Thambisetty *et al*.[10] reported an association of plasma clusterin levels with atrophy of the entorhinal cortex, a region known to show early pathological changes in AD. They also observed elevated plasma clusterin levels in AD patients with rapid clinical progression.

Because treatment of AD is still aimed at delaying symptoms rather than curing the disease, early detection of the disease and initiation of treatment is of great importance. We have therefore investigated serum clusterin concentrations as a potential biomarker in presymptomatic AD patients from a population based prospective cohort study, the Rotterdam Scan Study (RSS).

MATERIALS AND METHODS

Subjects and samples

The study population was derived from the Rotterdam Scan Study (RSS), a population based prospective cohort study initiated in 1995-1996, which included 1077 non-demented healthy subjects aged between 60 to 90 years. Participants in the RSS were randomly selected from 2 ongoing population based studies, the Rotterdam Study and the Zoetermeer Study[11]. The Erasmus University Medical Ethics Committee approved the study, and written informed consent was obtained from all participants.

The RSS was designed to study the etiology and natural history of age-related brain changes in the elderly. At baseline (1995-1996), information was obtained on socio-demographic characteristics, medical history, current health status, medication use, and determinants for chronic diseases. In addition, participants underwent brain imaging, veni-puncture and several physical examinations, including a neurologic and neuropsychologic screening. Participants have been followed since. In person re-examinations took place in 1997-1999

and 2002-2004, and involved the same standardized interviewing, veni-puncture and physical examination. For the current study we considered a 10-year follow up period.
Among the 1077 subjects 43 developed AD (4%), the AD assessment protocol for these patients has been described elsewhere[12]. Per AD case we randomly selected a control matched on gender, sampled in 5-year age categories. The selected controls remained dementia-free and were still alive at the end of follow-up (2005).
At baseline, blood samples were drawn by veni-puncture from non-fasting subjects at the Research Centres of the Rotterdam Study and the Zoetermeer Study between 9.00 AM and 4.00 PM. After coagulation for 30 minutes serum was obtained by centrifugation for 10 minutes at 3,000 rpm (Microcentrifuge 5417R; Eppendorf, Hamburg, Germany) at room temperature. Serum samples were aliquoted in 1 ml portions, fast frozen at -196 °C in liquid nitrogen, transferred to the laboratory and subsequently stored at -80 °C until further use.

Sample preparation

Serum samples were thawn on ice, 20 µl was taken and diluted 50 times in 50 mM ammonium bicarbonate. Of this dilution 10 µl was further diluted by adding 90 µl of 0.1% RapiGest SF (Waters, Milford, MA). Samples were reduced and alkylated with 1,4-dithiothreitol (DTT) and iodoacetamide. One µl of 0.5 M DTT was added, and incubated at 60 °C for 30 minutes. Next, 5 µl of 0.3 M iodoacetamide was added, and incubated in the dark at room temperature for 30 minutes. For the enzymatic digestion 10 µl of 100 µg/ml gold grade trypsin (Promega, Madison, WI) in 3 mM Tris-HCl was added and samples were incubated overnight at 37 °C. To inactivate trypsin and degrade RapiGest SF, trifluoric acid (TFA; Biosolve, Valkenswaard, Netherlands) was added to a final concentration of 0.5%. Subsequently samples were incubated for 30 minutes. Serum digest were aliquoted and stored at -20 °C until further use.

Peptide selection clusterin

Since different peptides from a single protein can vary in detectability in the mass spectrometer, we chose peptides for our MRM assay based on experimental data from the bottom-up proteomics study described in chapter 2. In this study twelve unique peptides of clusterin (Uniprot accession number P10909) were identified (see table 4.1). These twelve peptides were subjected to criteria described by Han and Higgs[13] including no internal trypsin cleavage sites, no potential modification sites, no ragged end and between 7 and 30 amino acids residues. Of two peptides that met the criteria, EIQNAVNGVK (primary amino acid position 45-54) and ASSIIDELFQDR (183-194), synthetic versions together with an internal standard were obtained from Pepscan Presto (Lelystad, Netherlands) with a

peptide purity > 95%. As internal standards peptides with the same amino acid residue sequences were taken with an additional glycine inserted in the middle of the sequence (EIQGNAVNGVK, ASSGIIDELFQDR). The assumption was made that incorporation of an extra glycine would not alter the properties of the peptide significantly. Amino acid analysis of both internal standards showed that the peptide content for EIQGNAVNGVK and ASSGIIDELFQDR were 70.6% and 68.6%, respectively.

Calibration curves were made for all synthetic peptides in digested serum. The following concentrations of synthetic peptide were measured 0, 0.16, 0.8, 4, 20 and 100 fmol/µl. Each concentration point for the calibration curve was measured in triplicate.

Serum digests of case-control samples were spiked with 20 fmol/µl of internal standards, which were corrected for their peptide content.

Table 4.1. Information of identified clusterin peptides.

Peptide sequence	Start - end	Missed cleavages	m/z measured	Mass calculated	Score	No. of spectra
R.**KYNELLK**.S	340 - 346	1	454.266	906.517	44	70
K.**FMETVAEK**.A	430 - 437	0	477.734	953.453	32	6
K.**EIQNAVNGVK**.Q	45 - 54	0	536.292	1070.572	49	89
K.**TLLSNLEEAK**.K	69 - 78	0	559.308	1116.603	64	47
R.**KTLLSNLEEAK**.K	68 - 78	1	415.906	1244.698	65	77
R.**KTLLSNLEEAK**.K	68 - 78	1	623.355	1244.698	49	60
K.**TLLSNLEEAKK**.K	69 - 79	1	623.356	1244.698	36	1
R.**KTLLSNLEEAKK**.K	68 - 79	2	458.605	1372.793	59	89
R.**ASSIIDELFQDR**.F	183 - 194	0	465.237	1392.689	61	86
R.**ASSIIDELFQDR**.F	183 - 194	0	697.352	1392.689	83	78
R.**EILSVDCSTNNPSQAK**.L	307 - 322	0	588.279	1761.820	56	90
R.**EILSVDCSTNNPSQAK**.L	307 - 322	0	881.919	1761.820	85	103
R.**QQTHMLDVMQDHFSR**.A	168 - 182	0	468.968	1871.841	31	6
R.**QQTHMLDVMQDHFSR**.A	168 - 182	0	624.956	1871.841	37	3
K.**LFDSDPITVTVPVEVSR**.K	409 - 425	0	625.332	1872.983	49	69
R.**VTTVASHTSDSDVPSGVTEVVVK**.L	386 - 408	0	772.065	2313.170	68	117

For each identified peptide the sequence, its position in the protein, number of missed cleavages, measured m/z-value and calculated mass is listed. Some peptides are listed twice, this is because the peptide was detected in two different charge states, resulting in two m/z-values. The score given is a probability based score, a value above 25 indicates identity.

Nano LC-MRM MS

Chromatographic separation of spiked serum digests, was performed on a nano LC system (Ultimate 3000; Dionex, Amsterdam, Netherlands). Two µl of spiked serum digest was loaded onto a C18 trap column (PepMap C18, 300 µm ID × 5 mm, 5 µm particle size and 100 Å pore size; Dionex) and washed for 5 min at a flow rate of 20 µL/min 0.1% TFA in H_2O. Next, the trap column was switched in line with the analytic column (PepMap C18, 75 µm ID × 150 mm, 3 µm particle size and 100 Å pore size; Dionex). Peptides were eluted at a flow rate of 300 nl/min with the following gradient: 0-45% solvent B in 30 min, solvent A (H_2O/acetonitrile (ACN) 98/2 (v/v), 0.1% formic acid (FA)) and solvent B (H_2O/ACN 20/80 (v/v), 0.1% FA). All solvents used were purchased at Biosolve. The separation of the peptides was monitored by a UV detector (absorption at 214 nm).
MRM detection was performed by means of a triple quadruple tandem mass spectrometer (4000 QTRAP; AB SCIEX, Concord, Ontario, Canada) in the positive ion mode. MRM parameters were optimized using the synthetic peptides, the optimized parameters are listed in table 4.2. For all peptides the entrance potential was set to 9 volts and the dwell time to 100 msec. The mass spectrometer (MS) was controlled using Analyst 1.5.1 (AB SCIEX). As technical control one spiked serum digest was re-measured every ten runs.

Table 4.2. Parameters for MRM assay.

Peptide	Charge	Q1	Q3	Collision energy	Declustering potential
EIQNAVNGVK	2	536.29	516.31 (y5)	25.5	80
EIQ<u>G</u>NAVNGVK	2	564.80	516.31 (y5)	28.5	90
ASSIIDELFQDR	2	697.35	922.43 (y7)	33	95
ASS<u>G</u>IIDELFQDR	2	725.86	922.43 (y7)	35	95

For each peptide, the precursor ion charge state is given together with the selected transitions (Q1 and Q3). Between brackets the numbered fragmented y-ion is listed. In the last columns the optimized collision energy and declustering potential (both in volts) are given.

Data analysis

For each transition, the peak area was determined using the MRM data analysis program Skyline v0.6[14]. By setting the internal standard concentration constant, the concentration of clusterin in the serum digests was calculated using the ratio of the internal standard

to the original peptide. A two-sided t-test was used to test for differences in the peptide concentrations between cases and controls.

RESULTS

Subject characteristics

By design, cases and controls were similar with respect to age and gender and non-demented at the time that the serum samples were collected for this study (table 4.3). The average Mini Mental State Examination (MMSE) score was however significantly lower in cases (26) than in controls (27). In addition, cases were more likely to carry one or two APOE ε4 alleles. The cases were diagnosed with AD after an average of 4.2 years (± 2.6 years SD) following blood sampling.

Table 4.3. Subject characteristics.

	Subjects		
	Cases	**Controls**	**p-value[a]**
Number	43	43	1
Mean age (years; SD)	78 (6.5)	78 (6.8)	0.88
Female/Male	32/11	32/11	1
Mean MMSE[b] score (SD)	26 (2.2)	27 (2.7)	0.01
Mean duration until AD diagnosis (years; SD)	4.2 (2.6)	0 (0)	0
APOE ε4/- (%)	15 (35%)	5 (12%)	0.01
APOE ε4/4 (%)	5 (12%)	0 (0%)	0.02

[a]p-value, a Mann-Whitney test was used to test for homogeneity between cases and controls; [b]MMSE, Mini Mental State Exam, maximum score 30.

MRM assay

Enhanced product ion scans were obtained from the four synthetic peptides, two original clusterin peptides and their two internal standards. On the basis of these experiments, the strongest transition per peptide that could be explained based on the structure was selected for MRM measurements. Once the transition was chosen, the collision energy and declustering potential were optimized. Extracted ion chromatograms resulted in

corresponding retention times and intensities for the original peptides and its internal standards.

Calibration curves were made for all synthetic peptides in digested serum. For both original peptides, the calibration curve of the original peptide was plotted together with the curve of the internal standard (figure 4.1). To exclude significant interference of the additional glycine the curves would ideally have the same slope and intercept. The slopes and intercepts for peptide EIQNAVNGVK and its internal standard EIQGNAVNGVK, were not significantly different with a p-value of 0.826 and 0.409, respectively. Peptide ASSIIDELFQDR and its internal standard ASSGIIDELFQDR, were significantly different with a p-value of 0.054 for the slopes and 0.024 for the intercepts.

For both internal standards, the limit of detection (LOD) and the lower limit of quantification (LLOQ) were determined. The LOD was defined as three times the signal-to-noise ratio of an unspiked serum digest and the LLOQ as the lowest calibration point of the curve that could be measured with a CV (coefficient of variance) less than 20%. For the MRM assay with internal standard EIQGNAVNGVK, the LOD was 0.5 fmol/µl and the LLOQ 0.8 fmol/µl. For the MRM assay with internal standard ASSGIIDELFQDR, the LOD and LLOQ were higher with values of 2.5 fmol/µl and 4 fmol/µl, respectively. Based on literature[15, 16] the expected concentration of clusterin in control serum is 100 ± 50 mg/l which corresponds to 1.9 ± 0.9 fmol/µl in our serum digests.

Since the calibration curve of the internal standard ASSGIIDELFQDR is significantly different from its original peptide and both the LOD and LLOQ of the MRM assay using this peptide is above the expected concentration, only the MRM assay with internal standard EIQGNAVNGVK was used to determine clusterin concentrations in cases and controls.

Clusterin levels in presymptomatic AD

There was no significant difference in clusterin concentration between presymptomatic AD (62.7 ± 13.9 mg/l) and controls (61.0 ± 11.4 mg/l) (figure 4.2A). When the presymptomatic AD cases were arranged on decreasing time to diagnosis, no (positive) trend in the concentration of clusterin could be observed (Pearson correlation, p-value 0.37). Also no correlation was found between the presence of one or two APOE ε4 alleles and the concentration of clusterin (figure 4.2B). The clusterin concentration measured with the MRM assay are in line with concentrations given in literature[15, 16].

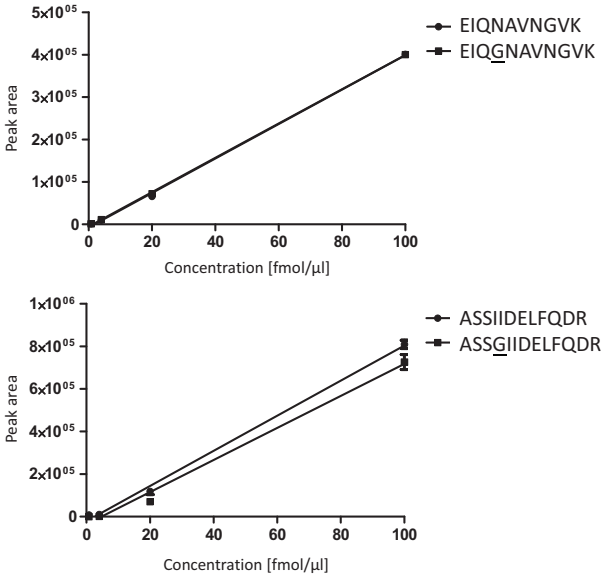


Figure 4.1. MRM calibration curves for the synthetic clusterin peptides together with their internal standards. All synthetic peptides were spiked into a serum digest at final concentrations of 0.16, 0.8, 4, 20 and 100 fmol/µl (three replicates per concentration point). Regression factors (R^2) for EIQNAVNGVK and EIQGNAVNGVK are both > 0.99. For ASSIIDELFQDR and ASSGIIDELFQDR the R^2 are > 0.99 and 0.98, respectively.

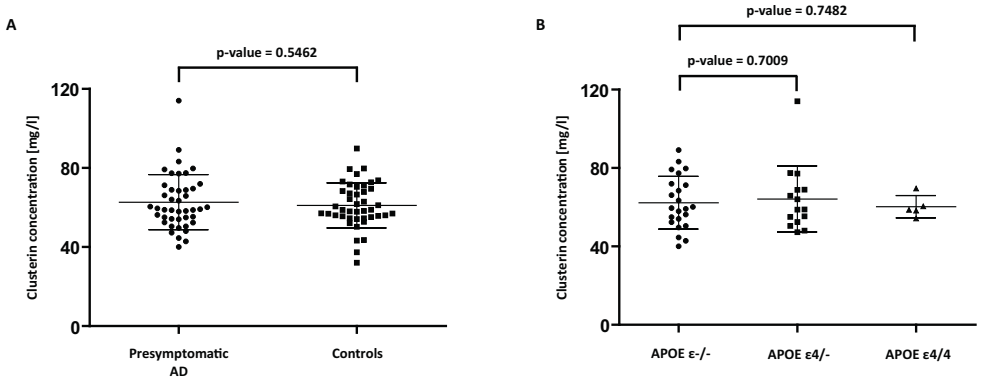


Figure 4.2. (A) Clusterin concentrations measured in presymptomatic AD and controls. For both groups the mean is indicated together with the standard deviation (error bars). A two-sided t-test shows that there is no significant difference between both groups (p-value 0.5462). (B) Clusterin concentrations within the presymptomatic AD group divided based on the presence of APOE ε4 alleles. APOE ε-/-, no ε4 alleles present; APOE ε4/-, one ε4 allele present; APOE ε4/4, two alleles present. For all groups the mean is indicated together with the standard deviation (error bars). A two-sided t-test shows that there is no significant difference in clusterin levels between persons with no APOE ε4 alleles present and persons with one or two APOE ε4 alleles present (p-value 0.7009 and p-value 0.7482).

DISCUSSION

In this population based cohort study, we found no difference in clusterin levels between persons with presymptomatic AD and gender and aged matched controls who remained dementia free for at least 10 years. The presence of one or two APOE ε4 alleles did not influence the clusterin concentrations.

An important strength of our study is the long and virtually complete follow-up, and the rigorous assessment of AD. Furthermore, we used an MRM assay for the measurement of clusterin. MRM has shown to be a highly reproducible and sensitive technique[17]. The choice of the standard is of great importance as it needs to be stable, show linearity and have a good sensitivity. We investigated an alternative for heavy isotope labeled peptides and were able to demonstrate that glycine inserted peptide standards has comparable chemical behaviour in retention time, intensity and ion efficiency. Of the two peptide standards selected for the MRM assay to measure clusterin concentrations in serum, only one proved to be more suitable (EIQGNAVNGVK) in respect to sensitivity and identical ion efficiency (figure 4.1). Therefore, we measured clusterin concentrations with an assay based on EIQNAVNGVK. A weakness of our study is the relatively small sample size, and the selective Caucasian ethnicity of our population. This means that the results of our study should be viewed with caution, particularly with respect to generalizability.

Thambisetty *et al.*[10] reported, besides the association of plasma clusterin levels in the progression and severity of AD, also an association between plasma clusterin concentrations and the amount of amyloid deposition in the brain assessed 10 years later. Based on these results they implicated a role for clusterin in the very early process of AD. However, their subjects showed no cognitive decline according to MMSE results and were not diagnosed with AD at the time of amyloid assessment (supplementary data Thambisetty *et al*.[10]). Our result of normal serum clusterin levels in AD patients approximately 4 years before diagnosis, does not support their hypothesis.

In summary, we have developed an MRM assay to measure clusterin levels in serum and applied this assay to serum samples from 43 persons who later developed AD and 43 gender and aged matched controls who remained dementia free. The results showed no significant difference in clusterin levels between the two groups. These findings suggest that clusterin is not a promising biomarker for the very early detection of AD.

REFERENCES

1. May PC, Lampert-Etchells M, Johnson SA, Poirier J, Masters JN, Finch CE: **Dynamics of gene expression for a hippocampal glycoprotein elevated in Alzheimer's disease and in response to experimental lesions in rat.** *Neuron* 1990, **5:**831-839.

2. Ghiso J, Matsubara E, Koudinov A, Choi-Miura NH, Tomita M, Wisniewski T, Frangione B: **The cerebrospinal-fluid soluble form of Alzheimer's amyloid beta is complexed to SP-40,40 (apolipoprotein J), an inhibitor of the complement membrane-attack complex.** *Biochem J* 1993, **293 (Pt 1):**27-30.

3. Matsubara E, Soto C, Governale S, Frangione B, Ghiso J: **Apolipoprotein J and Alzheimer's amyloid beta solubility.** *Biochem J* 1996, **316 (Pt 2):**671-679.

4. Oda T, Wals P, Osterburg HH, Johnson SA, Pasinetti GM, Morgan TE, Rozovsky I, Stine WB, Snyder SW, Holzman TF, et al.: **Clusterin (apoJ) alters the aggregation of amyloid beta-peptide (A beta 1-42) and forms slowly sedimenting A beta complexes that cause oxidative stress.** *Exp Neurol* 1995, **136:**22-31.

5. Yerbury JJ, Poon S, Meehan S, Thompson B, Kumita JR, Dobson CM, Wilson MR: **The extracellular chaperone clusterin influences amyloid formation and toxicity by interacting with prefibrillar structures.** *FASEB J* 2007, **21:**2312-2322.

6. Zlokovic BV: **Cerebrovascular transport of Alzheimer's amyloid beta and apolipoproteins J and E: possible anti-amyloidogenic role of the blood-brain barrier.** *Life Sci* 1996, **59:**1483-1497.

7. DeMattos RB, Cirrito JR, Parsadanian M, May PC, O'Dell MA, Taylor JW, Harmony JA, Aronow BJ, Bales KR, Paul SM, Holtzman DM: **ApoE and clusterin cooperatively suppress Abeta levels and deposition: evidence that ApoE regulates extracellular Abeta metabolism in vivo.** *Neuron* 2004, **41:**193-202.

8. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, Pahwa JS, Moskvina V, Dowzell K, Williams A, et al: **Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease.** *Nat Genet* 2009, **41:**1088-1093.

9. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, Combarros O, Zelenika D, Bullido MJ, Tavernier B, et al: **Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease.** *Nat Genet* 2009, **41:**1094-1099.

10. Thambisetty M, Simmons A, Velayudhan L, Hye A, Campbell J, Zhang Y, Wahlund LO, Westman E, Kinsey A, Guntert A, et al: **Association of plasma clusterin concentration with severity, pathology, and progression in Alzheimer disease.** *Arch Gen Psychiatry* 2010, **67:**739-748.

11. Breteler MM: **Vascular involvement in cognitive decline and dementia. Epidemiologic evidence from the Rotterdam Study and the Rotterdam Scan Study.** *Ann N Y Acad Sci* 2000, **903:**457-465.

12. Prins ND, van Dijk EJ, den Heijer T, Vermeer SE, Koudstaal PJ, Oudkerk M, Hofman A, Breteler MM: **Cerebral white matter lesions and the risk of dementia.** *Arch Neurol* 2004, **61:**1531-1534.

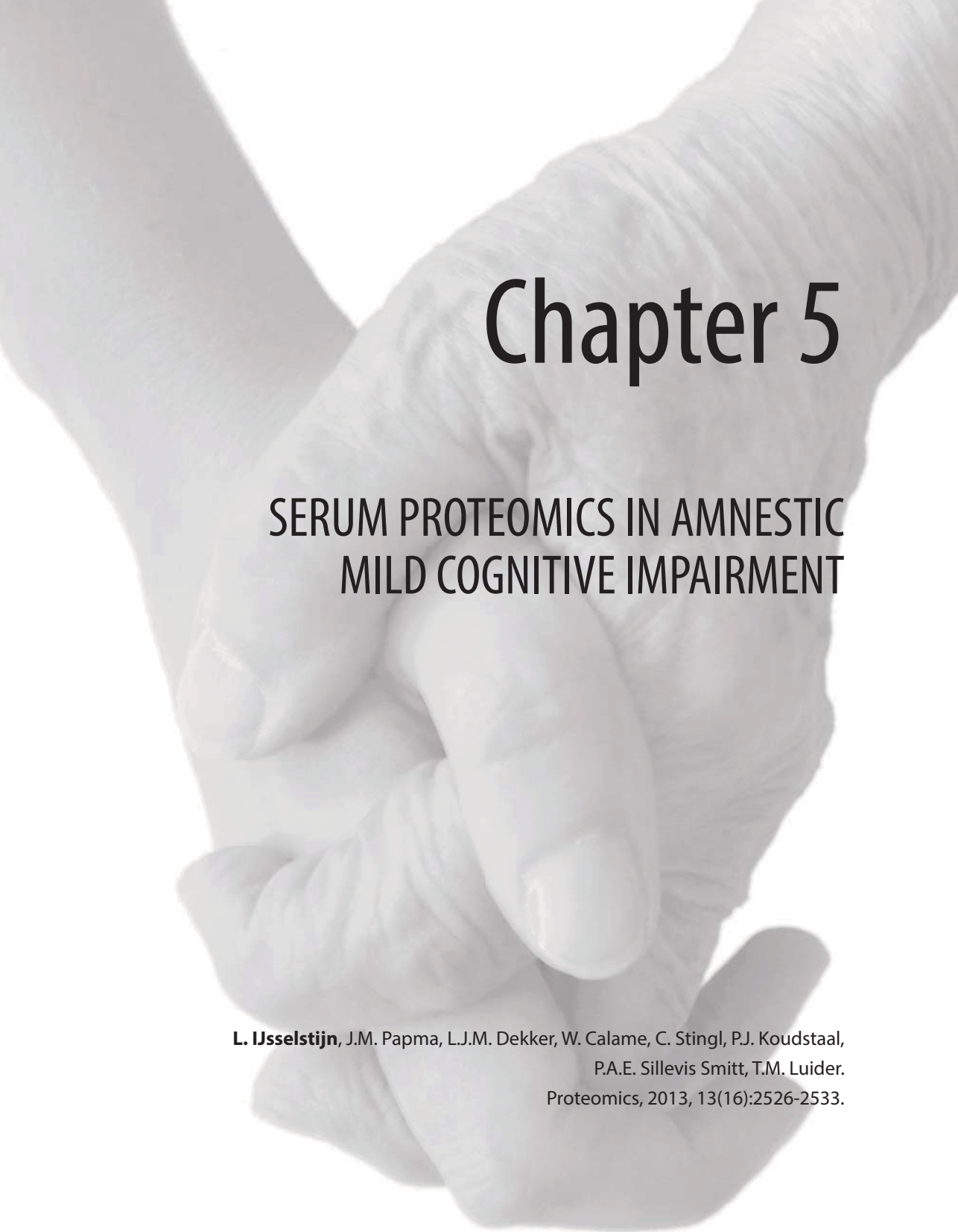
13. Han B, Higgs RE: **Proteomics: from hypothesis to quantitative assay on a single platform. Guidelines for developing MRM assays using ion trap mass spectrometers.** *Brief Funct Genomic Proteomic* 2008, **7:**340-354.

14. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ: **Skyline: an open source document editor for creating and analyzing targeted proteomics experiments.** *Bioinformatics* 2010, **26:**966-968.

15. Choi NH, Tobe T, Hara K, Yoshida H, Tomita M: **Sandwich ELISA assay for quantitative measurement of SP-40,40 in seminal plasma and serum.** *J Immunol Methods* 1990, **131:**159-163.

16. Morrissey C, Lakins J, Moquin A, Hussain M, Tenniswood M: **An antigen capture assay for the measurement of serum clusterin concentrations.** *J Biochem Biophys Methods* 2001, **48:**13-21.

17. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham AJ, Keshishian H, et al: **Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma.** *Nat Biotechnol* 2009, **27:**633-641.



Chapter 5

SERUM PROTEOMICS IN AMNESTIC MILD COGNITIVE IMPAIRMENT

L. IJsselstijn, J.M. Papma, L.J.M. Dekker, W. Calame, C. Stingl, P.J. Koudstaal, P.A.E. Sillevis Smitt, T.M. Luider.

Proteomics, 2013, 13(16):2526-2533.

ABSTRACT

We have explored proteins related to mild cognitive impairment (MCI). The serum proteome of 35 amnestic MCI patients and 35 cognitively healthy persons was investigated by liquid chromatography mass spectrometry. We identified 108 differentially expressed peptides between MCI patients and controls, belonging to 39 proteins. Eight proteins were selected for further investigation by quantitative protein measurements using a multiple reaction monitoring (MRM) assay; apolipoprotein E (APOE), carboxypeptidase N subunit 2 (CPN2), complement factor B (CFAB), galectin-3 binding protein (LG3BP), lumican (LUM), serum amyloid A-4 protein (SAA4), serum amyloid P-component (SAMP), and sex hormone binding globulin (SHGB). Results of the quantitative protein measurements showed significantly decreased levels of CPN2, CFAB, LG3BP, SAA4 and SAMP in serum from amnestic MCI patients compared with cognitive healthy controls (two-sided t-test; $p<0.05$). APOE and LUM showed no significant difference in protein levels, SHGB could not be quantified since the MRM assay did not reach the required sensitivity. A model based on the three most significantly decreased proteins (CFAB, LG3BP and SAA4) showed a sensitivity and specificity of 73% and 66%, respectively, for the initial sample set. A small external validation set yielded 77% sensitivity and 75% specificity.

INTRODUCTION

Dementia is a clinical syndrome of cognitive impairment that interferes with social or occupational functioning. Affected areas of cognition may be memory, attention, language and executive functioning. The most common causes of dementia at older age are Alzheimer's disease (AD) and cerebral vascular disease. As our population ages, dementia becomes a major health issue. Research regarding the early stages of dementia is of value, in particular for the development of novel strategies to prevent or delay dementia. Amnestic mild cognitive impairment (MCI) is considered to be a pre-stage of Alzheimer's disease, and can be defined as the presence of a memory complaint, objective memory impairment abnormal for age with relatively preserved general cognition, and essentially intact activities of daily living (no dementia)[1-3]. In the past decades, several techniques have been used to identify early alterations in this at-risk group[4], and most research has focused on clinical biomarkers for MCI in cerebrospinal fluid (CSF) or structural and functional neuroimaging[5]. Even though blood samples are more easily to obtain than CSF samples or MRI images, biomarkers for MCI derived from serum or plasma are scarce. Most serum or plasma biomarker research in the field of dementia and cognitive decline has been performed in patients with AD in the dementia stage. MCI patients are often included in these studies as a control group for AD cases[6]. A previous study by Blasko *et al.*, showed that plasma Aβ-42 levels are high in MCI and drop as the cognitive functions of patients further decline in AD[7]. Therefore, it is of interest to examine the serum or plasma proteome of MCI patients in comparison with healthy persons.

In this study, we used an unbiased proteomics approach to search for and identify serum biomarkers associated with MCI. We investigated the serum proteome of 35 amnestic MCI patients and 35 cognitively healthy persons using liquid chromatography mass spectrometry. Subsequently, we used multiple reaction monitoring to quantify the detected differences on the same samples and on extra independent samples.

MATERIALS AND METHODS

Subjects and samples

We recruited 35 MCI patients aged 65 years or older, from outpatient clinics of the departments of Geriatrics and Neurology of the Erasmus MC, Rotterdam, and surrounding clinics. All patients met the criteria for amnestic MCI described by Petersen *et al.*[8]. Briefly, these criteria are 1) presence of cognitive complaints from patient or caregiver, 2) decline in one or more cognitive domains as measured by neuropsychological assessment, 3)

absence of dementia according to the DSM-IV or NINCS ADRDA criteria for dementia[9], and 4) preserved overall general functioning. Patients were not included if they had a history of neurological or psychiatric disorders which could negatively affect cognition (e.g. stroke or depression). All MCI patients underwent MRI imaging of the brain, routine physical and neurological examination, extensive neuropsychological testing, and veni-puncture. We used the results of the MRI imaging, neurological examination and neuropsychological tests to assure study eligibility of the MCI patients. In addition, we recruited 35 control persons from the Erasmus Stroke Study, who were friends or spouses, no family members, of patients treated for transient ischaemic attack or stroke at the Erasmus MC[10]. They had neither cognitive complaints nor cognitive impairment, and were matched with MCI patients within a five-year age range. The Erasmus MC Medical Ethics Committee approved the study and written informed consent was obtained from all participants.

The collected serum samples were aliquoted in 1 mL portions, fast frozen in liquid nitrogen and subsequently stored at -80 °C until further use. To prepare the serum samples for liquid chromatography mass spectrometry (LC-MS), an enzymatic digestion was performed as described previously[11]. Briefly, serum was diluted 500 times in 0.1% RapiGest SF (Waters, Milford, MA), reduced using dithiothreitol and alkylated using iodoacetamide. Subsequently, trypsin was added for digestion. After incubation the pH was reduced to terminate the digestion reaction and the serum digests were aliquoted (25 µL) and stored at −20 °C until further use.

Nano LC Orbitrap MS

LC-MS measurements were carried out on a nano LC system (Ultimate 3000; Dionex, Amsterdam, Netherlands) online coupled to a hybrid linear ion trap/Orbitrap MS (LTQ Orbitrap XL; Thermo Fisher Scientific, Bremen, Germany). Four microliters of serum digest was loaded onto a C18 trap column (PepMap C18, 300 µm ID × 5 mm, 5 µm particle size and 100 Å pore size; Dionex) and desalted for 10 min at a flow rate of 20 µL/min 0.1% TFA in H_2O. Next, the trap column was switched inline to an analytic column (PepMap C18, 75 µm ID × 500 mm, 3 µm particle size and 100 Å pore size; Dionex). A 180-minute gradient with a 300 nl/min flow was run with solvent A (H_2O/acetonitrile (ACN) 98/2 (v/v), 0.1% formic acid (FA)) and solvent B (H_2O/ACN 20/80 (v/v), 0.1% FA): 0-25% solvent B in 120 min and 25-50% solvent B in the next 60 min. All solvents used were purchased at Biosolve (Valkenswaard, Netherlands). The separation of the peptides was monitored by a UV detector (absorption at 214 nm).

High resolution full scan MS spectra were obtained from the Orbitrap (resolution 30,000; AGC 1,000,000), MS/MS spectra were obtained by CAD fragmentation. MS/MS was performed on the top five masses in the full scan spectra. Dynamic exclusion was used,

with a repeat count of one; exclusion duration was set at 3 min and exclusion width at ± 5 ppm. As a technical control a pool of serum digests was re-measured every ten runs.
Data analysis/management was done using the software package Progenesis LC-MS (version 3.0; Nonlinear Dynamics Ltd, New Castle, UK). After peptide detection a filter was used, yielding only features charged with two to eight protons and containing more than two isotope peaks. Of the whole experiment a matrix was generated, which includes all features with their corresponding (normalized) intensities and abundances.
For peptide identification, MS/MS data was searched against the human SwissProt database (version July 23th 2009) using Mascot (version 2.3.01; Matrix Science Inc., London, UK). The following settings were used; enzyme: trypsin, fixed modifications: carbamidomethylation of cysteine (+57.021 u), variable modifications: oxidation of methionine (+15.995 u), peptide mass tolerance: ± 10 ppm (# ^{13}C = 2), fragment mass tolerance: ± 0.5 Da, max missed cleavages: 2. Peptide identifications with a score above 25 were imported in Progenesis and a matrix was generated containing the following information for each peptide feature; retention time, normalized peak area and when present its sequence.

Capillary LC MRM MS

For the proteins of interest, stable isotope labeled peptides were ordered (Pepscan Presto, Lelystad, Netherlands) to use as internal standards. The C-terminal amino acid (arginine or lysine) of the isotope labeled peptides contained the isotopes ^{13}C and ^{15}N in place of ^{12}C and ^{14}N. We prepared calibration series of the internal standards ranging from 0.2 to 100 fmol/µL in serum digest matrix. Each concentration of the calibration series was measured in triplicate. Serum digests of case-control samples were spiked with 20 fmol/µL of each standard. For each standard, the limit of detection (LOD) and the lower limit of quantification (LLOQ) were determined. The LOD was defined as three times the signal-to-noise ratio of an unspiked serum digest and the LLOQ as the lowest calibration point of the curve that could be measured with a CV (coefficient of variance) less than 20%.
Chromatographic separation of spiked serum digests was performed on a capillary LC system (Ultimate 3000; Dionex, Amsterdam, Netherlands). Fifteen microliters of spiked serum digest was loaded onto a C18 trap column (PepMap100 C18, 500 µm ID × 5mm, 5 µm particle size and 100 Å pore size; Dionex) and washed for 5 min at a flow rate of 25 µL/min 0.1% TFA in H_2O. Next, the trap column was switched in line with the analytic column (PepMap C18, 300 µm ID × 150 mm, 3 µm particle size and 100 Å pore size; Dionex). Peptides were eluted at a flow rate of 6 µL/min with the following gradient: 4-25% solvent B in 20 min and 25-50% solvent B in the next 10 min, solvent A (H_2O, 0.1% formic acid (FA)) and solvent B (H_2O/acetonitrile 20/80 (v/v), 0.08% FA). All solvents used were purchased

at Biosolve. The separation of the peptides was monitored by a UV detector (absorption at 214 nm).
MRM detection was performed by means of a triple quadruple tandem mass spectrometer (4000 QTRAP; AB SCIEX, Concord, Ontario, Canada) in the positive ion mode. A turbo-V source was used with ion spray voltage of 4500 V at 200 °C. A scheduled MRM method was used with the MRM detection window set to 240 sec and the target scan time to 3 sec. The dwell times of the various peptides varied between 71 and 500 ms. The MS system was controlled using Analyst 1.5.1 (AB SCIEX). As technical control one spiked serum digest was repeatedly measured every 10 runs.
For each transition, the peak area was determined using the program Skyline v0.6[12]. The concentrations of the endogenous peptides in the serum digests were calculated using the ratios of the original peptide to its internal standard.

Statistical analysis

The present study was conducted as pilot study using a data set which was available at our department. Recognizing this fact a proportional power analysis[13] using an α-level of 0.05 and a power (1-β) level of 0.80 revealed that for two-sided testing at least 90 MCI patients and 90 controls were needed to show an increase from 0.50 to 0.70 in either sensitivity or specificity qualifications. It is emphasized that the outcome in the current study was unknown, moreover, to our knowledge, no information on our approach from the international literature is available. Since the actual number in both groups, being 35 in both arms, was substantially lower than required for the above mentioned increase randomized boot strap approach was applied to minimize variation issues in the outcome. Moreover, this approach enabled us to detect potential impact by specific persons on the outcome in sensitivity and specificity determinations. Accepting these facts the extra 13 MCI persons and 4 controls was only used to establish to which extent the results of these small groups would strengthen or contradict the results obtained in the 35 patients and 35 controls.
To test for homogeneity in the obtained data between the MCI patients and controls, an ANOVA or chi square test was used with a significance level of $p<0.05$. To determine whether peptides present in sera from MCI patients were significantly higher or lower than in sera from controls, the normalized peak areas of the nano LC Orbitrap MS measurements were compared using a two-sided t-test with a significance level of $p<0.05$. For the proteins selected as significantly different by nano LC Orbitrap MS measurements, an independent and quantitative mass spectrometry analysis was performed. Differences in protein concentrations between MCI patients and controls were tested for significance by a two-sided t-test and ANCOVA, corrected for gender and age (significance level of

$p<0.05$). The statistical tests were performed using the SPSS statistical software package (version PASW 17.0.2).

The concentrations of the quantified proteins were also used to build models to differentiate between MCI and controls. The 35 observations of MCI patients and the 35 observations of the controls were randomly split into two groups: group I and II. This resulted in groups of either 18 observations for MCI and 17 for control, or vice versa. This process was randomly (using Latin square design) repeated at least twenty times, yielding different combinations of patients and controls, to gain insight in the robustness and the sensitivity and specificity of the various models consisting of varying combinations of the peptides. Model building was done via a combination of canonical linear discriminant analysis[14] and stepwise logistic regression analysis[15] in order to yield that model with the highest sensitivity and specificity in group I which was validated using the data in group II. Canonical linear discriminant analysis was applied to quantify the difference in the incidence of a "0" (control) versus that of a "1" (MCI). It provides information on the difference in group means on canonical variables indicating the relative difference using these peptides between cases and controls. The logistic regression model was applied to determine which peptide could be taken out without a major drop in either sensitivity or specificity characteristics as observed via canonical linear discriminant analysis. In other words: could we minimize the number of peptides without losing sensitivity and specificity. This resulted in various models with different combinations of the proteins evaluated. In the logistic regression analysis control data were assigned with "0" and MCI data with "1". The outcome of this process was statistically analyzed using the parameter free Wilcoxon matched-pairs signed-rank test. This allowed identification of the optimal combination of the seven identified proteins. Since after the start of the present study new patients (and controls) became available the obtained models were checked for sensitivity and specificity in a small group of 13 patients and 4 controls as well (external validation) having not been included in the original analysis.

The analyses were done using STATA (version 10; StataCorp, Texas, US).

RESULTS

Subject characteristics

The MCI patients and controls showed no statistically significant differences regarding age, gender and MMSE score (table 5.1). Although not significant, MCI patients were more likely to carry one or two APOE ε4 alleles. Neuropsychological assessment confirmed that MCI patients consisted of the amnestic MCI subtype[2], either single or multiple-domain (data not shown).

Table 5.1. Subject characteristics.

	MCI (n = 35)	Controls (n = 35)	p-Value†
Mean age	74 (4.6)	73 (4.6)	0.535
Male/Female	27/8	21/14	0.122
MMSE (max. 30)	27 (2.0)	29 (1.2) [a]	0.514
APOE ε4/- / APOE ε4/4	16 / 3	1 / 0 [a]	0.074

Values are unadjusted means (standard deviation) or proportions. [a]MMSE and APOE data were only present for 6/35 controls. † By means of one way ANOVA or chi square test. MMSE = Mini Mental State Examination.

Nano LC Oribtrap MS

The nano LC Orbitrap MS measurements resulted in a total of 31,122 detected features, whereby, 3,356 of these features could be assigned to a peptide. The total number of unique peptides identified in this sample set was 1,981 belonging to 326 proteins.

Statistical analysis of the 1,981 identified peptides revealed 108 peptides that could discriminate between MCI and controls (t-test $p<0.05$). Proteins that met the following criteria were selected for further quantitative analysis: 1) the protein must be identified by at least two peptides, 2) of the identified peptides of a protein more than 15% of the peptides needed to show a significant difference between MCI and controls (no cut-off value for the fold change was used), and 3) in general, all peptides of the protein should either all show an increase or a decrease. Proteins fulfilling these criteria were apolipoprotein E (APOE), carboxypeptidase N subunit 2 (CPN2), complement factor B (CFAB), galectin-3 binding protein (LG3BP), lumican (LUM), serum amyloid A-4 protein (SAA4), serum amyloid P-component (SAMP), and sex hormone binding globulin (SHGB) (table 5.2). LUM and SHGB were increased in MCI compared to control serum samples, the other six proteins were decreased (table 5.2).

Table 5.2. Proteins selected for further quantitative analysis.

Protein	Accession number	Significant peptides (total peptides)	Increased or decreased in MCI	Peptides selected for MRM assay[a]
Apolipoprotein E (APOE)	P02649	3 (17)	↓	**AATVGSLAGQPLQER** **LGPLVEQGR**
Carboxypeptidase N subunit 2 (CPN2)	P22792	1 (6)	↓	AGGSWDLAVQER **GQVVPALNEK**
Complement factor B (CFAB)	P00751	8 (24)	↓	**ISVIRPSK** **QLNEINYEDHK**
Galectin-3 binding protein (LG3BP)	Q08380	1 (3)	↓	**LADGGATNQGR** YSSDYFQAPSDYR
Lumican (LUM)	P51884	2 (8)	↑	LPSGLPVSLLTLYLDNNK **NNQIDHIDEK**
Serum amyloid A-4 protein (SAA4)	P35542	1 (2)	↓	**GPGGVWAAK**
Serum amyloid P-component (SAMP)	P02743	3 (6)	↓	**AYSLFSYNTQGR** **VGEYSLYIGR**
Sex hormone-binding globulin (SHBG)	P04278	1 (3)	↑	IALGGLLFPASNLR

[a]These peptides were selected and tested for the development of an MRM assay for quantitative protein measurements. Peptides in bold were used in the final MRM assay.

Protein quantification using MRM

The peptides given in table 5.2 were synthesized with stable isotope labeled arginine or lysine at the C-terminal end of the peptide and were used as internal standards to develop an MRM assay for protein quantification. The limit of detection and lower limit of quantification were determined for each internal standard. Four internal standards had limits above the expected protein concentration and were therefore not suitable for quantitative measurements. SHBG concentrations could not be determined as the only peptide selected for this protein had detection limits above the expected serum concentration. Results of the protein levels for APOE, CPN2, CFAB, LG3BP, LUM, SAA4 and SAMP are listed in table 5.3. Serum concentrations of CPN2, CFAB, LG3BP, SAA4 and SAMP were significantly lower in MCI compared to controls. When corrected for gender and age using ANCOVA, the serum concentrations of CPN2, CFAB, LG3BP, SAA4 and SAMP remained significantly lower in MCI patients.

Table 5.3. Quantitative levels of APOE, CPN2, CFAB, LG3BP, LUM, SAA4 and SAMP in MCI and controls.

Protein	MCI [mg/L]	Controls [mg/L]	p-Value[a]
Apolipoprotein E (APOE)	81.4 (29.4)	100.8 (51.9)	0.06
Carboxypeptidase N subunit 2 (CPN2)	68.1 (8.9)	73.4 (9.2)	0.02
Complement factor B (CFAB)	304.5 (58.5)	353.2 (60.1)	0.001
Galectin-3 binding protein (LG3BP)	18.4 (7.4)	24.0 (8.8)	0.006
Lumican (LUM)	48.6 (6.6)	50.5 (6.9)	0.23
Serum amyloid A-4 protein (SAA4)	59.2 (13.7)	71.5 (23.3)	0.009
Serum amyloid P-component (SAMP)	66.8 (17.2)	76.8 (22.0)	0.04

Values are unadjusted means (standard deviation). [a]A two-sided t-test was used to test for differences in quantitative protein levels between MCI and controls.

Internal validation

Based on the serum levels of the seven proteins in table 5.3, various models were built that could differentiate between MCI patients and controls. Internal validation of these models was performed by randomly splitting the groups of MCI patients and controls multiple times and repeatedly determining the sensitivity and specificity. The various models and their average sensitivities and specificities are listed in table 5.4.

Table 5.4. Average sensitivity and specificity for the various models after randomly splitting the groups of MCI patients and controls multiple times (n).

Model	Proteins used	n	Average sensitivity	Average specificity
1	APOE, CPN2, CFAB, LG3BP, LUM, SAA4, SAMP	40	74.8 %	66.7 %
2	CFAB, LG3BP, SAA4	40	73.1 %	65.7 %
3	CPN2, LUM, SAMP	40	68.0 %	64.0 %
4	CFAB	36	74.3 %	63.9 %
5	LG3BP	24	73.0 %	56.9 %
6	CPN2, CFAB, LG3BP, LUM, SAA4, SAMP	20	73.3 %	66.0 %

Model 3 showed an average sensitivity which is significantly lower ($p<0.05$) than all the other models and for model 5 a significantly ($p<0.05$) lower specificity is observed. All other models listed have similar sensitivity and specificity values, which renders model 2, using the proteins: CFAB, LG3BP and SAA4, the most practical combination to distinguish MCI from control by applying the lowest number of proteins to be determined.

The algorithm of this model is:

Incidence = (-0.00188 x CFAB) + (-0.00933 x LG3BP) + (-0.00371 x SAA4) + 1.56104,
F: 5.14, p<0.005

with incidence being defined as "0" for control and "1" for MCI.

External validation

After the start of this study, we continued inclusion of MCI patients and controls for a prospective MRI study providing us with a potential validation set. Unfortunately, we could only include 13 MCI patients and 4 controls. Quantitative protein measurements of the seven interesting proteins were performed on sera from these subjects. The results of the t-tests, although not reliable due to the small number of controls, only showed a significant difference in protein concentration for LUM. All proteins showed the same change when comparing MCI and controls, with exception of SAA4 and SAMP. Concentrations of these proteins showed an increase in MCI in this validation set, while in the set described in the manuscript they showed a decrease. Notwithstanding the low numbers of cases and controls in this set model 2 yielded 77.0 % sensitivity and 75.0 % specificity.

DISCUSSION

In this study, we found significantly lower concentrations of complement factor B (CFAB), carboxypeptidase N subunit 2 (CPN2), galectin-3 binding protein (LG3BP), serum amyloid A-4 protein (SAA4) and serum amyloid P-component (SAMP) in persons with amnestic MCI compared with controls without cognitive impairment. These proteins (as well as APOE, LUM and SHBG) were not found in a previous study in which the same proteomics approach was taken in the search for differentiating proteins in presymptomatic Alzheimer's disease (AD)[16]. A model based on CFAB, LG3BP and SAA4 (the three most significantly decreased proteins) showed a sensitivity and specificity of 73% and 66%, respectively. Further inclusion of MCI patients and controls resulted in only a small external validation set (n=13 and n=4 for MCI patients and controls, respectively). Despite of the small number of subjects in the validation set, the model based on CFAB, LG3BP and SAA4 yielded 77% sensitivity and 75% specificity.

CFAB showed the largest and most significant (p=0.001) difference in concentrations between MCI patients and controls and was of great influence on the sensitivity and specificity in the different models. CFAB is part of the alternative complement pathway which acts as a rapid and efficient immune surveillance system in the absence of specific antibodies. CFAB binds to hydrolyzed molecules of complement C3 and is cleaved by complement factor D. This cleavage generating C3 convertase that activates complement

by cleaving C3 into its active fragments C3a and C3b[17]. Complement proteins are mainly produced by the liver, but can also be synthesized locally in the brain. There is evidence that the complement system is involved in neurodegenerative diseases like AD, Parkinson's disease, Pick's disease and prion disease[18]. In these diseases, complement proteins interact or are activated by a number of proteins of different etiology that form highly fibrillar deposits with specific motifs, like amyloid-β fibrils in AD[19]. Whether the neurotoxicity in these diseases is a result of the aggregates itself or of a reaction activated by these aggregates remains to be determined.

Another protein found differentially expressed in this study, CPN2, is also linked to the complement system. CPN2 is part of the zinc metalloprotease carboxypeptidase N (CPN), which consists of two enzymatically active small subunits (CPN1) and two large subunits (CPN2) that protect the protein from degradation. CPN cleaves the basic amino acids, lysine and arginine, from the carboxy-terminus of biologically active peptides and proteins. Substrates for CPN are complement anaphylatoxins (C3a, C4a and C5a), creatine kinase and kinins[20]. To our knowledge nothing is known in the literature on CPN2 or CPN in relation to MCI or neurodegeneration. However, kinins (especially bradykinin) are known to be involved in modulating cerebrovascular tone, eliciting vasodilatation and increasing the permeability of the blood-brain barrier (BBB). There is evidence that the permeability of the BBB is increased in AD[21] and epidemiological studies have shown that many risk factors for cognitive decline and dementia relate to cerebrovascular disease (e.g., hypertension, atherosclerosis)[22]. Since the vascular condition of the subjects could possibly be linked to the CPN levels, a closer look was taken at the medication of the subjects. Medication to control blood pressure was taken by two subjects in the MCI group and by two subjects in the control group. This medication use did not have an influence on the CPN2 levels.

LG3BP, also known as 90K or mac-2 binding protein, is a secreted glycoprotein. LG3BP, originally identified by its ability to bind to galectin-3, binds proteins and molecules mediating in cell-matrix and cell-cell adhesions such as collagens, integrins and fibronectin[23]. Although its function in not yet entirely defined, LG3BP has been investigated in relation to the enhancement of natural killer cell activation, the production of various cytokines, interleukins and growth factors, and its role in tumor growth suppression[24-26]. Increased levels for LG3BP have been observed in cancer tissue and sera, and have been demonstrated to be associated with poor progression[27]. To our knowledge LG3BP has never been associated with MCI, dementia or other neurodegenerative diseases.

Serum amyloid A proteins (SAAs) are apolipoproteins of high density lipoprotein (HDL) and can be divided in two groups. The first group comprises the acute phase SAAs which associate with HDL during inflammation. The second group consists of constitutive SAAs which exist as minor apolipoproteins on HDL but constitute more than 90% of the total

SAA during homeostasis[28]. SAA4, the SAA we found in lower concentrations in MCI patients compared to controls is part of the latter. For SAA4 no relation with MCI or dementia was ever proposed in the literature. However, acute phase SAAs were detected in brain of AD patients as well as in brain of patients with conditions with a known inflammatory component, but not in brains of patients with other neurological diseases[29].
SAMP is a glycoprotein of the pentraxin family which is always present in amyloid deposits *in vivo where it* binds strictly calcium-dependent, to ligands present on all amyloid fibrils[30]. Amyloidoses, the disorders in which amyloid deposits are found, are of diverse etiology and ultimately lead to progressive destruction of tissue. Amyloidoses can be localized like Aβ deposits in AD, or systemic like in amyloid light-chain amyloidosis (AL), the most common form of amyloidosis, in which the amyloid deposits are derived from immunoglobulin light chains[31]. In the second most common type of amyloidosis worldwide, the deposits are formed by the acute phase SAAs described previously (AA amyloidosis). Tennent *et al.* showed that SAMP prevents proteolysis of the amyloid fibrils in AD, AL and AA and may thereby contribute to their persistence in vivo[32]. SAMP is not an enzyme inhibitor and is protective only when bound to the fibrils. Levels of SAMP serum and/or CSF levels have been investigated as a biomarker for AD with contradictory results[33-38]. Verwey *et al.* studied next to AD patients and controls also MCI patients and found lower SAMP levels in CSF of MCI patients that progressed to dementia compared to nonprogressors[38].

Of the five proteins that were significantly different between persons with amnestic MCI and controls with normal cognition, only SAMP had a direct link to dementia. The other four proteins are all involved in complement or immune reactions. Confirmation of our findings in a larger external validation set is needed to unravel the role of CFAB, CPN2, LG3BP, SAA4 and SAMP in amnestic MCI.

REFERENCES

1. Burns A, Zaudig M: **Mild cognitive impairment in older people.** *Lancet* 2002, **360:**1963-1965.

2. Petersen RC: **Mild cognitive impairment as a diagnostic entity.** *J Intern Med* 2004, **256:**183-194.

3. Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B: **Current concepts in mild cognitive impairment.** *Arch Neurol* 2001, **58:**1985-1992.

4. Ewers M, Walsh C, Trojanowski JQ, Shaw LM, Petersen RC, Jack CR, Jr., Feldman HH, Bokde AL, Alexander GE, Scheltens P, et al: **Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance.** *Neurobiol Aging* 2010.

5. Frank RA, Galasko D, Hampel H, Hardy J, de Leon MJ, Mehta PD, Rogers J, Siemers E, Trojanowski JQ: **Biological markers for therapeutic trials in Alzheimer's disease. Proceedings of the biological markers working group; NIA initiative on neuroimaging in Alzheimer's disease.** *Neurobiol Aging* 2003, **24:**521-536.

6. Song F, Poljak A, Smythe GA, Sachdev P: **Plasma biomarkers for mild cognitive impairment and Alzheimer's disease.** *Brain Res Rev* 2009, **61:**69-80.

7. Blasko I, Jellinger K, Kemmler G, Krampla W, Jungwirth S, Wichart I, Tragl KH, Fischer P: **Conversion from cognitive health to mild cognitive impairment and Alzheimer's disease: prediction by plasma amyloid beta 42, medial temporal lobe atrophy and homocysteine.** *Neurobiol Aging* 2008, **29:**1-11.

8. Artero S, Petersen R, Touchon J, Ritchie K: **Revised criteria for mild cognitive impairment: validation within a longitudinal population study.** *Dement Geriatr Cogn Disord* 2006, **22:**465-470.

9. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM: **Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease.** *Neurology* 1984, **34:**939-944.

10. van den Herik EG, Cheung EY, de Lau LM, den Hertog HM, Leebeek FW, Dippel DW, Koudstaal PJ, de Maat MP: **gamma'/total fibrinogen ratio is associated with short-term outcome in ischaemic stroke.** *Thromb Haemost* 2011, **105:**430-434.

11. IJsselstijn L, Dekker LJ, Koudstaal PJ, Hofman A, Sillevis Smitt PA, Breteler MM, Luider TM: **Serum clusterin levels are not increased in presymptomatic Alzheimer's disease.** *J Proteome Res* 2011, **10:**2006-2010.

12. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ: **Skyline: an open source document editor for creating and analyzing targeted proteomics experiments.** *Bioinformatics* 2010, **26:**966-968.

13. Cohen J: *Statistical Power Analysis for the Behavioral Sciences.* 2nd edn. Mahwah, New Jersey: Lawrence Erlbaum Associates; 1988.

14. Afifi A, Clark V, May S: *Computer-Aided Multivariate Analysis* 4th edn. Boca Raton, Florida: Chapman & Hall; 2004.

15. Kleinbaum DG, Kupper LL, Muller KE, Nizam A: *Applied regression analysis and other multivariable methods.* 3rd edn. Pacific grove, California: Duxbury Press; 1998.

16. IJsselstijn L, Dekker LJ, Stingl C, van der Weiden MM, Hofman A, Kros JM, Koudstaal PJ, Sillevis Smitt PA, Ikram MA, Breteler MM, Luider TM: **Serum Levels of Pregnancy Zone Protein are Elevated in Presymptomatic Alzheimer's Disease.** *J Proteome Res* 2011, **10:**4902-4910.

17. Ricklin D, Hajishengallis G, Yang K, Lambris JD: **Complement: a key system for immune surveillance and homeostasis.** *Nat Immunol* 2010, **11:**785-797.

18. Veerhuis R, Nielsen HM, Tenner AJ: **Complement in the brain.** *Mol Immunol* 2011, **48:**1592-1603.

19. Velazquez P, Cribbs DH, Poulos TL, Tenner AJ: **Aspartate residue 7 in amyloid beta-protein is critical for classical complement pathway activation: implications for Alzheimer's disease pathogenesis.** *Nat Med* 1997, **3:**77-79.

20. Matthews KW, Mueller-Ortiz SL, Wetsel RA: **Carboxypeptidase N: a pleiotropic regulator of inflammation.** *Mol Immunol* 2004, **40:**785-793.

21. Farrall AJ, Wardlaw JM: **Blood-brain barrier: ageing and microvascular disease--systematic review and meta-analysis.** *Neurobiol Aging* 2009, **30:**337-352.

22. Breteler MM: **Vascular involvement in cognitive decline and dementia. Epidemiologic evidence from the Rotterdam Study and the Rotterdam Scan Study.** *Ann N Y Acad Sci* 2000, **903:**457-465.

23. Inohara H, Akahani S, Koths K, Raz A: **Interactions between galectin-3 and Mac-2-binding protein mediate cell-cell adhesion.** *Cancer Res* 1996, **56:**4530-4534.

24. Ullrich A, Sures I, D'Egidio M, Jallal B, Powell TJ, Herbst R, Dreps A, Azam M, Rubinstein M, Natoli C, et al.: **The secreted tumor-associated antigen 90K is a potent immune stimulator.** *J Biol Chem* 1994, **269:**18401-18407.

25. Fukaya Y, Shimada H, Wang LC, Zandi E, DeClerck YA: **Identification of galectin-3-binding protein as a factor secreted by tumor cells that stimulates interleukin-6 expression in the bone marrow stroma.** *J Biol Chem* 2008, **283:**18573-18581.

26. Kalayci O, Birben E, Tinari N, Oguma T, Iacobelli S, Lilly CM: **Role of 90K protein in asthma and TH2-type cytokine expression.** *Ann Allergy Asthma Immunol* 2004, **93:**485-492.

27. Grassadonia A, Tinari N, Iurisci I, Piccolo E, Cumashi A, Innominato P, D'Egidio M, Natoli C, Piantelli M, Iacobelli S: **90K (Mac-2 BP) and galectins in tumor progression and metastasis.** *Glycoconj J* 2004, **19:**551-556.

28. de Beer MC, Yuan T, Kindy MS, Asztalos BF, Roheim PS, de Beer FC: **Characterization of constitutive human serum amyloid A protein (SAA4) as an apolipoprotein.** *J Lipid Res* 1995, **36:**526-534.

29. Liang JS, Sloane JA, Wells JM, Abraham CR, Fine RE, Sipe JD: **Evidence for local production of acute phase response apolipoprotein serum amyloid A in Alzheimer's disease brain.** *Neurosci Lett* 1997, **225:**73-76.

30. Pepys MB, Dyck RF, de Beer FC, Skinner M, Cohen AS: **Binding of serum amyloid P-component (SAP) by amyloid fibrils.** *Clin Exp Immunol* 1979, **38:**284-293.

31. Picken MM: **Amyloidosis-where are we now and where are we heading?** *Arch Pathol Lab Med* 2010, **134:**545-551.

32. Tennent GA, Lovat LB, Pepys MB: **Serum amyloid P component prevents proteolysis of the amyloid fibrils of Alzheimer disease and systemic amyloidosis.** *Proc Natl Acad Sci U S A* 1995, **92:**4299-4303.

33. Hawkins PN, Rossor MN, Gallimore JR, Miller B, Moore EG, Pepys MB: **Concentration of serum amyloid P component in the CSF as a possible marker of cerebral amyloid deposits in Alzheimer's disease.** *Biochem Biophys Res Commun* 1994, **201:**722-726.

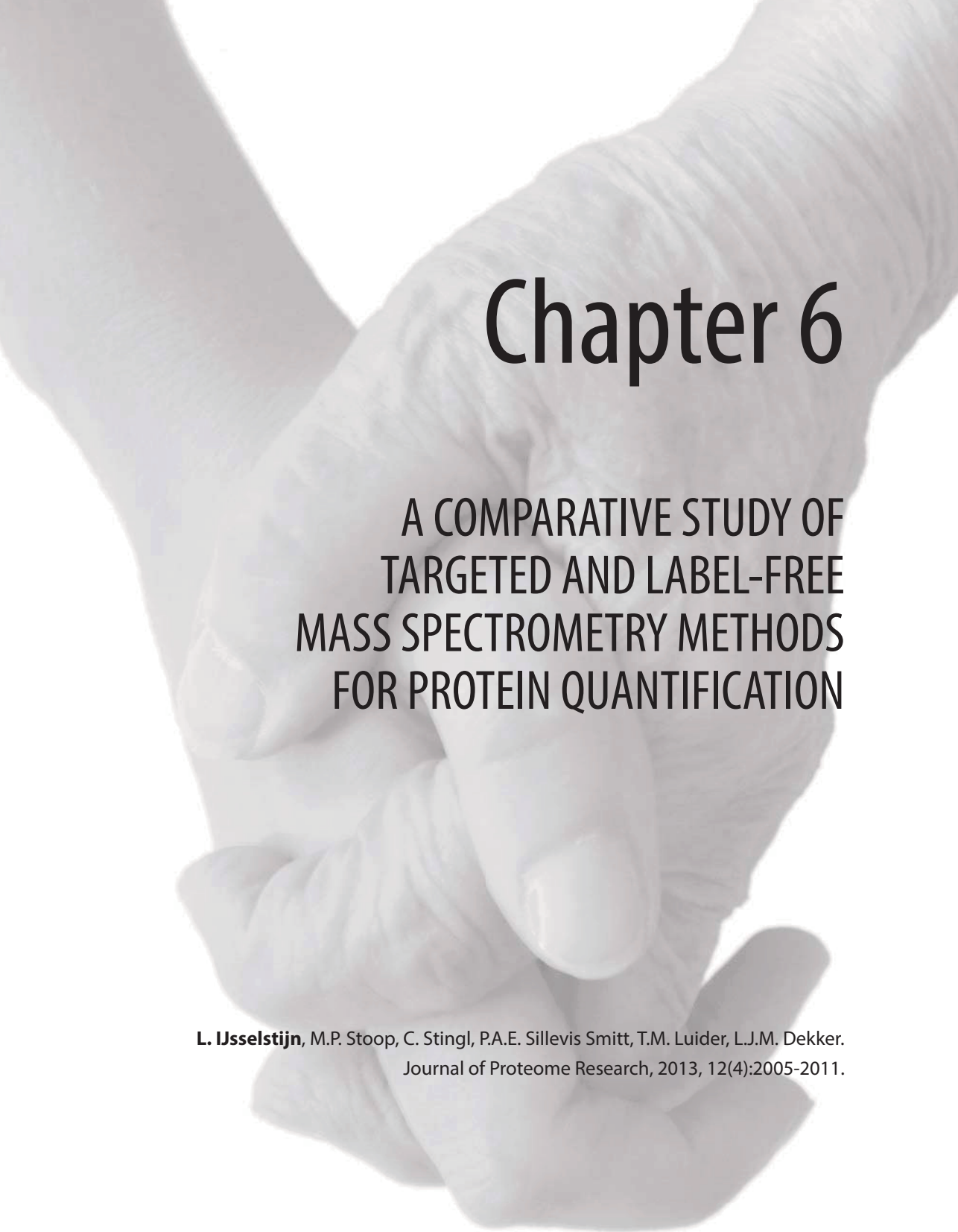
34. Kimura M, Asada T, Uno M, Machida N, Kasuya K, Taniguchi Y, Fujita T, Nishiyama E, Iwamoto N, Arai H: **Assessment of cerebrospinal fluid levels of serum amyloid P component in patients with Alzheimer's disease.** *Neurosci Lett* 1999, **273:**137-139.

35. Mulder C, Schoonenboom SN, Wahlund LO, Scheltens P, van Kamp GJ, Veerhuis R, Hack CE, Blomberg M, Schutgens RB, Eikelenboom P: **CSF markers related to pathogenetic mechanisms in Alzheimer's disease.** *J Neural Transm* 2002, **109:**1491-1498.

36. Nishiyama E, Iwamoto N, Kimura M, Arai H: **Serum amyloid P component level in Alzheimer's disease.** *Dementia* 1996, **7:**256-259.

37. Nybo M, Olsen H, Jeune B, Andersen-Ranberg K, Holm Nielsen E, Svehag SE: **Increased plasma concentration of serum amyloid P component in centenarians with impaired cognitive performance.** *Dement Geriatr Cogn Disord* 1998, **9:**126-129.

38. Verwey NA, Schuitemaker A, van der Flier WM, Mulder SD, Mulder C, Hack CE, Scheltens P, Blankenstein MA, Veerhuis R: **Serum amyloid p component as a biomarker in mild cognitive impairment and Alzheimer's disease.** *Dement Geriatr Cogn Disord* 2008, **26:**522-527.



Chapter 6

A COMPARATIVE STUDY OF TARGETED AND LABEL-FREE MASS SPECTROMETRY METHODS FOR PROTEIN QUANTIFICATION

L. IJsselstijn, M.P. Stoop, C. Stingl, P.A.E. Sillevis Smitt, T.M. Luider, L.J.M. Dekker.
Journal of Proteome Research, 2013, 12(4):2005-2011.

ABSTRACT

We compared data acquired on an LTQ-Orbitrap MS used in a typical shotgun proteomics setting (optimized for protein identification) with data from a quadrupole ion trap MS operated in the MRM mode. Six relative abundant proteins were quantified in identical sets of serum and CSF samples by the following methods; a qual/quant method with and without use of internal standards and a quantitative method (MRM with use of internal standards). Comparison of these methods with an antibody-based method in CSF samples showed good linearity for both methods (R^2 of 0.961 and 0.971 for the qual/quant method with use of internal standards and the quantitative method, respectively). Besides its better linearity, the quantitative method was also more reproducible with lower CVs for all samples. Next to these comparisons we also explored why a qual/quant approach had typically a lower reproducibility compared to MRM analyses. We observed that modified peptides, or peptides with a cysteine or a methionine, yielded a significant increase in CV. Furthermore, a positive correlation was found between the length of the peptide and the CV. We conclude that qual/quant is an alternative for the quantification of abundant proteins and that the use of internal standards in qual/quant could be advantageous. Furthermore, the ongoing development in MS techniques increases the possibilities of qual/quant in protein quantification.

INTRODUCTION

Mass spectrometry (MS) is a well-established tool in protein quantification and is increasingly used for the verification of biomarkers. multiple reaction monitoring (MRM) has been a reference quantitative technique to analyze small molecules for more than 30 years[1], but the past decade it has emerged as a tool in quantitative proteomics[2, 3]. MRM can be used for relative quantification, in which a reference sample is labeled either metabolically (SILAC)[4], chemically (ICAT[5], mTRAQ[6]) or enzymatically with stable isotopes (O^{18})[7]. By addition of internal standards, like synthetic stable-isotope labeled homologs[8] or glycine inserted peptides[9], the absolute amount of the corresponding endogenous peptide can be determined. MRM coupled with stable isotope dilution MS has shown to be highly reproducible, within and across laboratories, and sensitive to low µg/mL protein concentrations in complex samples without enrichment of the proteins of interest[10]. At these concentration levels, immunoassays can be replaced by MRM assays in the verification of larger numbers of candidate biomarkers in a single experiment. In the biomarker discovery phase, often qualitative MS measurements are performed aiming at identifying large numbers of proteins for instance in shotgun proteomic experiments[11]. Due to higher accuracy and faster scan speeds of modern equipment, workflows have been developed which combine qualitative measurements with high-resolution, accurate mass measurements for quantitative purposes. These approaches that combine qualitative and quantitative aspects are called qual/quant. Qual/quant measurements have a considerable advantage over a targeted approach, because prior to the MS measurements no target proteins or peptides have to be selected. All observed peptides by MS are in principle available for quantification and can under circumstance improve the quality of quantification.

In this study, we investigated whether qual/quant measurements performed on equipment routinely used for protein identification, are of sufficient sensitivity and reproducibility to use for quantification of proteins in complex samples. Secondly, we examined why a qual/quant approach has typically a lower reproducibility and quantitative performance compared to MRM analyses. To answer these questions, six proteins (ranging from mg/mL to low µg/mL concentrations) were quantified in identical sets of serum and cerebrospinal fluid (CSF) digests, by qual/quant measurements using an Orbitrap-platform and by quantitative measurements using a Qtrap-platform. We used the two instruments operating in a general way which has been described in previous chapters (Orbitrap-platform for biomarker discovery and Qtrap-platform for biomarker verification/validation)[12-14]. Stable-isotope labeled peptides were used as internal standards in the

quantitative measurements. In addition, these standards were also spiked in the samples for the qual/quant analysis, in order to determine if this improves quantification.

MATERIALS AND METHODS

Sample background

The serum samples used in this study were a selection of control samples from a previous biomarker study[14]. The CSF samples originated from a previous proteomics and metabolomics analysis of normal CSF samples[15]. Both studies were approved by the Erasmus University Medical Ethics Committee, and written informed consents were obtained from all participants.

Protein selection

The six proteins selected for a quantitative analyses are proteins which were detected in previous proteomics studies in serum[14] as well as in CSF[16]. The concentrations of the proteins are in the mg/mL (albumin) to low µg/mL (galectin-3 binding protein) range. For each protein two peptides were selected for quantification. These peptides had to meet the following criteria; no internal trypsin cleavage sites, no potential modification sites (cysteine and methionine), no ragged ends, and a maximum of 30 amino acid residues. For each selected peptide, a stable-isotope labelled peptide standard was synthesized (Pepscan Presto, Lelystad, the Netherlands), in which the arginine (R) or lysine (K) at the C-terminus was replaced with the heavy form of the amino acid. The sequences of the selected peptides and the corresponding standards can be found in table 6.1.

Table 6.1. Transitions used in MRM assay of the selected peptides and their standards.

Protein	Peptide	Q1 [m/z]	Q3 [m/z]
Albumin	SLHTLFGDK	509.27	579.31 (y5), 680.36 (y6), 817.42 (y7)
	SLHTLFGD**K**	513.27	587.33 (y5), 688.38 (y6), 825.43 (y7)
	FQNALLVR	480.78	500.36 (y4), 571.39 (y5), 685.44 (y6)
	FQNALLV**R**	485.79	510.36 (y4), 581.40 (y5), 695.44 (y6)
Complement C3	IWDVVEK	444.74	474.29 (y4), 589.32 (y5), 775.40 (y6)
	IWDVVE**K**	448.75	482.31 (y4), 597.33 (y5), 783.41 (y6)
	VVLVAVDK	421.77	432.25 (y4), 531.31 (y5), 644.40 (y6)
	VVLVAVD**K**	425.78	440.26 (y4), 539.33 (y5), 652.41 (y6)
Vitamin D binding protein	VLEPTLK	400.25	458.30 (y4), 587.34 (y5), 700.42 (y6)
	VLEPTL**K**	404.26	466.31 (y4), 595.35 (y5), 708.44 (y6)
	HLSLLTTLSNR	627.86	691.37 (y6), 917.54 (y8), 1004.57 (y9)
	HLSLLTTLSN**R**	632.87	701.38 (y6), 927.55 (y8), 1014.58 (y9)
Complement factor B	QLNEINYEDHK	701.83	805.35 (y6), 918.43 (y7), 1161.52 (y9)
	QLNEINYEDH**K**	705.84	813.36 (y6), 926.45 (y7), 1169.53 (y9)
	ISVIRPSK	450.29	600.38 (y5), 699.45 (y6), 786.48 (y7)
	ISVIRPS**K**	454.29	608.40 (y5), 707.47 (y6), 794.50 (y7)
Galectin-3 binding protein	LADGGATNQGR	530.26	575.29 (y5), 760.37 (y8), 875.40 (y9)
	LADGGATNQG**R**	535.27	585.30 (y5), 770.38 (y8), 885.41 (y9)
	YSSDYFQAPSDYR	799.84	338.18 (y2), 637.29 (y5), 708.33 (y6)
	YSSDYFQAPSDY**R**	804.85	348.19 (y2), 647.30 (y5), 718.34 (y6)
Apolipoprotein E	LGPLVEQGR	484.78	360.20 (y3), 588.31 (y5), 701.39 (y6)
	LGPLVEQG**R**	489.78	370.21 (y3), 598.32 (y5), 711.40 (y6)
	AATVGSLAGQPLQER	749.40	642.36 (y5), 827.44 (y7), 898.47 (y8)
	AATVGSLAGQPLQE**R**	754.41	652.37 (y5), 837.45 (y7), 908.48 (y8)

For each peptide the transitions (Q1 and Q3) are given (all doubly charged). Peptides with a bold K (lysine) or R (arginine) and the C-terminus are stable-isotope labeled. Stable-isotope labeled amino acids K and R contained ^{13}C and ^{15}N in place of ^{12}C and ^{14}N.

Albumin concentration

Albumin concentrations were determined by the department of Clinical Chemistry, Erasmus Medical Center with an antibody-based method on the Immage 800 (Beckman Coulter, Brea, CA). Shortly, the samples were diluted, an albumin-specific antibody was added and the resulting immunoprecipitation was measured by nephelometry.

Preparation spiked digests

To prepare the spiked serum digests, serum samples (n=20) were thawed on ice, 20 µL was taken and diluted 50 times in 50 mM ammonium bicarbonate. From this dilution 10 µL was further diluted by adding 90 µL of 0.1% RapiGest SF (Waters, Milford, MA). During this second dilution step, synthetic stable-isotope labelled peptide standards (Pepscan Presto, Lelystad, Netherlands) were added to yield final concentrations after digestion of 100 fmol/µL for the albumin peptides and 25 fmol/µL for all other peptides investigated. Subsequently, samples were reduced and alkylated with 1,4-dithiothreitol (DTT) and iodoacetamide. For the enzymatic digestion 10 µL of 100 µg/mL gold grade trypsin (Promega, Madison, WI) in 3 mM Tris-HCl was added and samples were incubated overnight at 37 °C. To inactivate trypsin and degrade RapiGest SF, trifluoroacetic acid (TFA; Biosolve, Valkenswaard, Netherlands) was added to a final concentration of 0.5%. Subsequently, samples were incubated for 30 minutes. The spiked serum digests were aliquoted to perform LC Orbitrap MS and LC MRM MS measurements on identical samples. To prepare the spiked CSF digests, CSF samples (n=21) were thawed on ice and 20 µL was taken. Twenty microliters of 0.2 % RapiGest SF was added together with the synthetic peptide standards to yield final concentrations after digestion of 100 fmol/µL for the albumin peptides and 25 fmol/µL for all other peptides. After reduction and alkylation, 4 µL of trypsin was added for the enzyme digestion and the samples were incubated overnight at 37 °C and treated as described above for serum. The spiked CSF digests were aliquoted to perform LC Orbitrap MS and LC MRM MS measurements on identical samples.

LC Orbitrap MS

LC-MS measurements were carried out on a nano LC system (Ultimate 3000; Dionex, Amsterdam, Netherlands) online coupled to a hybrid linear ion trap/Orbitrap MS (LTQ Orbitrap XL; Thermo Fisher Scientific, Bremen, Germany). Four microliters of spiked serum digest or 5 µL of spiked CSF digest was injected onto the nano LC system, which held a C18 trap column (PepMap C18, 300 µm ID × 5 mm, 5 µm particle size and 100 Å pore size; Dionex) and a 25 cm long analytic column (PepMap C18, 75 µm ID × 500 mm, 3 µm particle size and 100 Å pore size; Dionex). A 90-minute gradient with a 300 nL/min flow was run with solvent A (H_2O/acetonitrile (ACN) 98/2 (v/v), 0.1% formic acid (FA)) and solvent B (H_2O/ACN 20/80 (v/v), 0.1% FA): 0-25% solvent B in 60 min and 25-50% solvent B in the next 30 min. All solvents used were purchased from Biosolve (Valkenswaard, Netherlands). The separation of the peptides was monitored by a UV detector (absorption at 214 nm).
High resolution full scan MS was obtained from the Orbitrap (resolution 30,000; AGC 1,000,000), MS/MS spectra were obtained by CAD fragmentation. MS/MS was performed

on the top five masses in the full scan spectra. Dynamic exclusion was used, with a repeat count of one; exclusion duration was set at 3 min and exclusion width at ± 5 ppm.
To check the performance of the system, one of the spiked serum or CSF digests was remeasured every 5 runs as a technical control.

LC MRM MS

Chromatographic separation of spiked digests, was performed on a nano LC system (Ultimate 3000; Dionex, Amsterdam, Netherlands). Two microliters of spiked serum or CSF digest was loaded onto a C18 trap column (PepMap100 C18, 300 µm ID × 5mm, 5 µm particle size and 100 Å pore size; Dionex) and washed for 5 min at a flow rate of 20 µL/min 0.1% TFA in H_2O. Next, the trap column was switched in line with the analytic column (PepMap C18, 75 µm ID × 150 mm, 3 µm particle size and 100 Å pore size; Dionex). Peptides were eluted at a flow rate of 300 nL/min with the following gradient: 0-45% solvent B in 30 min, solvent A (H_2O, 0.1% formic acid (FA)) and solvent B (H_2O/acetonitrile 20/80 (v/v), 0.08% FA). All solvents used were purchased at Biosolve. The separation of the peptides was monitored by a UV detector (absorption at 214 nm).
MRM detection was performed by means of a quadrupole ion trap tandem mass spectrometer (4000 QTRAP; AB SCIEX, Concord, Ontario, Canada) in the positive ion mode. A scheduled MRM method was used with the MRM detection window set to 240 sec and the target scan time to 3 sec. The transitions used are listed in table 1. The mass spectrometer was controlled using Analyst 1.5.1 (AB SCIEX).
To check the performance of the system, one of the spiked serum and CSF digests was remeasured every 5 runs as a technical control.
Calibration curves for the standards were made in digested serum and CSF. The following concentrations of standards were measured 0, 0.5, 2, 5, 20, 50, and 200 fmol/µL. Each concentration point for the calibration curve was measured in triplicate.

Data analysis

The program Skyline v1.1[17] was used to analyse both LC Orbitrap MS and LC MRM MS data measured with internal standards. For quantification of Orbitrap data, we used the sum of the integrated areas of the extracted ion chromatograms (XIC) of the monoisotopic precursor and first ^{13}C isotope precursor. For the MRM data, the peak areas for all measured transitions were determined. The concentration of the peptides was calculated using the ratio of the original peptide to its standard. The program Progenesis LC-MS (version 3.0; Nonlinear Dynamics Ltd, Newcastle-upon-Tyne, UK) was used additionally for analysis of LC Orbitrap MS data. After peptide detection a filter was used, yielding only features charged with two to eight protons and containing more than two isotope peaks.

Of the whole experiment a matrix was generated, which includes all features with their corresponding intensities and (normalized) abundances. For peptide identification, MS/MS data was searched against the human SwissProt database (version July 23th 2009) using Mascot (version 2.3.01; Matrix Science Inc., London, UK). The following settings were used; enzyme: trypsin, fixed modifications: carbamidomethylation of cysteine (+57.021 u), variable modifications: oxidation of methionine (+15.995 u), peptide mass tolerance: ± 10 ppm (# $^{13}C = 2$), fragment mass tolerance: ± 0.5 Da, maximally allowed missed cleavages: 2.

Statistical analysis

To test for differences in CV of the six proteins between serum and CSF, a paired t-test was used with a significance level of 0.05. For each peptide of the six proteins identified by the qual/quant method, the CV was calculated. To determine whether there was a significant difference in CV between identified peptides with and without a modification, a potential modification site present (cysteine or methionine), or a missed cleavage, a two-sided t-test with a significance level of 0.05 was used. To test for the presence of a significant correlation between the CV and the number of amino acids, the intensity or the number of co-eluting peptides within 0.1 minutes time window, the Pearson's correlation was used with a significance level of 0.05. The statistical tests were performed using the SPSS statistical software package (version PASW 17.0.2).

RESULTS

Linearity

The albumin concentrations in cerebrospinal fluid (CSF) measured by qual/quant and quantitative methods were compared with the albumin concentrations determined by an antibody-based method (figure 6.1). The regression coefficient (R^2) was taken as a measure for the linearity of the methods. The quantitative method (figure 6.1A) showed the best linearity (highest R^2) and the qual/quant method based on the average intensity of all identified peptides of albumin the worst linearity (figure 6.1D). The use of an internal standard in the qual/quant method improved this linearity (figure 6.1B). The albumin concentrations measured with both the quantitative and the qual/quant methods differed systematically from the antibody-based method with factor of 1.7.

The same methods were applied to determine the concentrations of albumin in serum. The results of these measurements were similar to the results obtained in CSF.

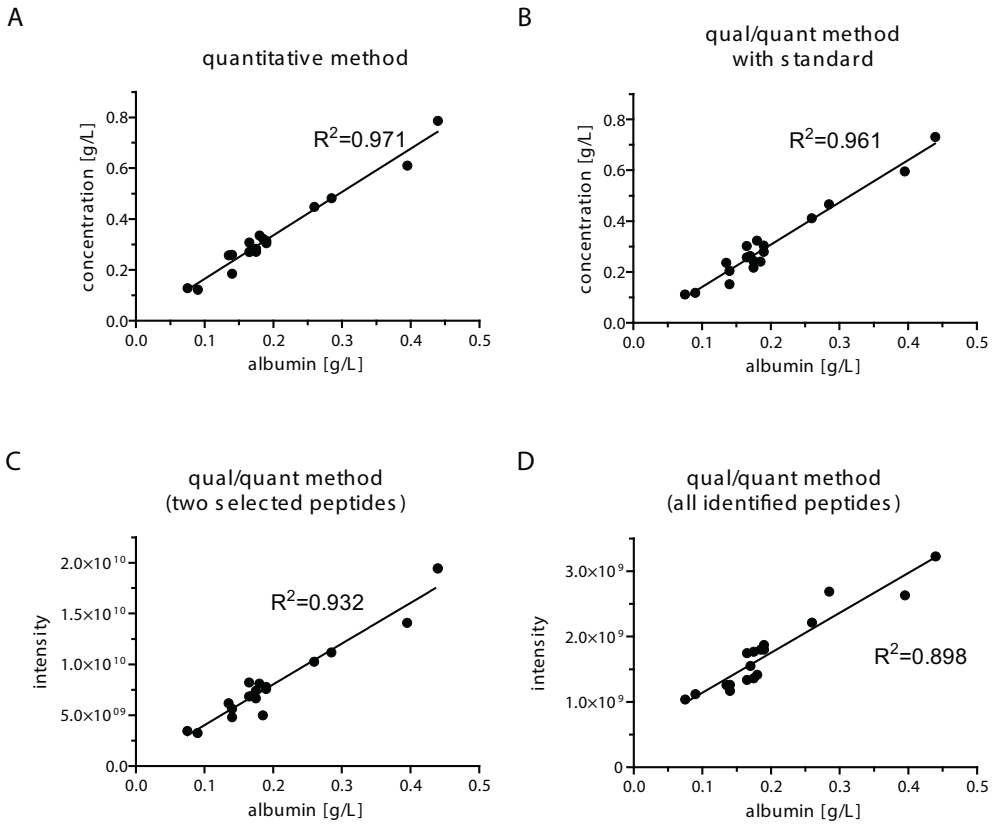


Figure 6.1. Albumin concentrations in CSF determined via various methods plotted against the concentration measured with the anti-body based method on the x-axis; (A) Quantitative method (Qtrap-platform) with internal standards. (B) Qual/quant method (Orbitrap-platform) with internal standards. (C) Qual/quant method (Orbitrap-platform), average intensity displayed of two peptides used in the methods with internal standard. (D) Qual/quant method (Orbitrap-platform), average intensity displayed of all identified peptides for albumin.

Reproducibility

The technical reproducibility of the different methods was determined by calculating the CV for each of the six proteins quantified in one serum and one CSF sample that was repeatedly measured during the sequence (table 6.2). The quantitative method showed the lowest CV for all six proteins in serum and CSF expect for complement factor B measured in CSF. The CV calculated for the quantitative method increased when proteins became less abundant. The addition of internal standards to the sample using the qual/quant method resulted in lower CVs for the less abundant proteins, when only the peptides corresponding to the standards were analyzed. The CVs for the qual/quant method using all identified peptides per protein (last column of table 6.2) showed lower CVs than the

qual/quant methods based on only two peptides, with or without the addition of internal standards. The various methods did not show significant differences in CVs between serum and paired CSF sample, except for the CVs calculated for the qual/quant method where only two selected peptides for quantitation were included (p-value 0.029).

Table 6.2. CVs for the level of six different proteins measured in serum and CSF by quantitative and qual/quant methods.

Proteins in serum	Quantitative	Qual/quant with standard	Qual/quant (2 peptides)	Qual/quant (all peptides)
Albumin	0.98	22.24	27.15	6.18
Complement C3	3.22	10.82	22.55	8.95
Vitamin D binding protein	5.50	21.27	26.15	6.61
Complement factor B	5.55	4.18	30.00	9.07
Apolipoprotein E	5.97	10.69	43.22	10.30
Galectin-3 binding protein	12.17	29.07	24.11	34.65
Proteins in CSF	**Quantitative**	**Qual/quant with standard**	**Qual/quant (2 peptides)**	**Qual/quant (all peptides)**
Albumin	1.49	17.21	12.07	6.01
Complement C3	3.22	8.67	8.03	4.47
Vitamin D binding protein	3.84	7.97	24.16	3.92
Complement factor B	14.63	9.52	8.79	4.30
Apolipoprotein E	2.32	11.06	11.72	3.99
Galectin-3 binding protein	2.32	9.44	22.73	4.88
p-value†	0.712	0.188	**0.029***	0.130

†p-Value of paired t-test to check for differences in CVs between serum and CSF (* significant difference $p < 0.05$).

CVs for internal standards in just water were calculated for the quantitative method and the qual/quant method (table 6.3). The CVs for the quantitative method varied between 2.08 and 5.82 for eleven of the twelve standards. Peptide HLSLLTTLSNR had a divergent CV of 17.47. For the qual/quant methods the CVs ranged from 6.96 to 17.30.

Table 6.3. CVs for the internal standards measured in water by the quantitative method (Qtrap-platform) and by the qual/quant method (Orbitrap-platform).

Standards in water		Quantitative	Qual/quant
Albumin	SLHTLFGD**K**	5.82	10.33
	FQNALLV**R**	2.53	11.64
Apolipoprotein E	LGPLVEQG**R**	2.84	11.05
	AATVGSLAGQPLQE**R**	2.87	8.97
Complement C3	IWDVVE**K**	3.21	13.01
	VVLVAVD**K**	4.62	6.96
Complement factor B	QLNEINYEDH**K**	4.62	17.30
	ISVIRPS**K**	2.08	13.07
Galectin-3 binding protein	LADGGATNQG**R**	3.90	7.51
	YSSDYFQAPSDY**R**	3.61	9.49
Vitamin D binding protein	VLEPTL**K**	2.10	9.88
	HLSLLTTLSN**R**	17.47	13.31

Influence on reproducibility in the qual/quant method

In the CSF samples, 242 peptides were identified belonging to the six proteins with an average CV for the normalized intensity of 12.9. The CV's of these peptides were used to determine which factors were of influence on the reproducibility of the qual/quant measurements. We examined the influence of the presence of modifications, the presence of a cysteine or the presence of a methionine and missed cleavages, the length of the peptide (number of amino acids), the number of co-eluting peptides within 0.1 minutes and the intensity of the peptide (table 6.4). When the peptide had one or more modifications, or when a cysteine or a methionine were present, a significant increase in the CV was observed ($p=0.005$, $p=0.005$ and $p<0.001$, respectively). A positive correlation was found between the length of the peptide and the CV (Pearson correlation coefficient 0.238, $p<0.001$).

To confirm the influence of these factors in practice, we used data from two other studies. In these studies both qual/quant (Orbitrap-platform) and quantitative (Qtrap-platform) measurements have been performed. In one study apolipoprotein E levels were determined in 70 serum digests (see chapter 5) and in the other study haptoglobin levels were determined in 34 CSF digests (unpublished data). In both studies we determined the correlation (R^2) between the intensity measured of each individual peptide using the Orbitrap-platform and the concentration determined using the Qtrap-platform. In figure 6.2, the R^2 for the individual peptides are plotted for both the filtered and the unfiltered

situation. We observed that the average correlation improved when the peptide list was filtered using the criteria described above and when peptides were removed containing missing values in the qual/quant approach, since these were the peptides with the lowest R^2 values.

Table 6.4. Influence of various factors on the CVs of identified peptides using the Orbitrap-platform.

Factor	CV_{pres}	CV_{abs}	p-value†
Modification	18.2	12.0	0.005
Cysteine	16.5	11.5	0.005
Methionine	18.6	11.4	< 0.001
Missed cleavage	14.1	12.2	0.25
Factor	**Correlation coefficient**		**p-value‡**
Number of amino acids	0.238		< 0.001
Number of co-eluting peptides (within 0.1 minute)	-0.054		0.05
Intensity of the peptide	-0.043		0.51

†A t-test was used to test for differences in CVs of peptides with the presence (CV_{pres}) or absence (CV_{abs}) of the factor (average CVs are listed). ‡For these factors, the presence of a correlation was determined using Pearson's correlation (correlation coefficients are listed).

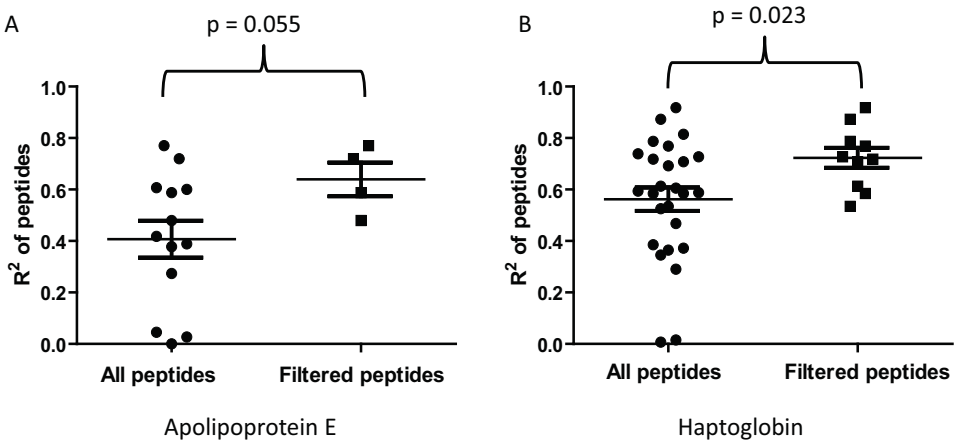


Figure 6.2. Plot of the correlation coefficients of the concentration determined on the Qtrap-platform and the intensity measured on the Orbitrap-platform for each individual peptide. Plot A shows the values for apolipoprotein E in serum and plot B for Haptoglobin in CSF. In the plots the R^2 for all the peptides are shown and the peptides that have been filtered based on the criteria proposed in this manuscript.

DISCUSSION

In this study, we investigated the differences in quantitative performance between a purely quantitative method (Qtrap-platform) and a qual/quant method with and without the use of internal standards and partially cross-validated these methods with an antibody-based technique that is routinely used in clinical chemistry laboratories. The Qtrap-platform is a dedicated system that is optimized for quantification. On the other hand, the settings for the Orbitrap MS were intentionally not optimized for quantitative measurements because this instrument is intended for identification. Yet, we were interested in its quantitative performance next to its (optimized) qualitative performance. Identical samples were used throughout all the measurements to ensure a fair comparison between all quantification methods.

The linearity of the measured albumin concentrations in CSF compared to the antibody-based technique was best for the quantitative method, followed by the qual/quant method with use of internal standards and without use of internal standards. Although the linearity of the qual/quant method without prior knowledge of the identified peptides was the least accurate method, there was still a regression coefficient (R^2) of 0.898 in CSF when compared to the concentrations determined using an antibody-based method. We therefore conclude that qual/quant is a good method to give a first and rather accurate quantitative impression. The factor 1.7 observed between the concentrations measured by MS methods and the antibody-based method can be explained by the amino acid content of the used internal standards, which is in general 60-80% of the weighted mass of the standard. If one corrects for this percentage, the values correlate within acceptable ranges.

When comparing the three qual/quant analysis (with internal standards and without internal standards taken two peptides or all identified peptides into account), the analysis using the internal standards showed a better reproducibility, but the best reproducibility for the qual/quant methods were observed for the analysis performed without internal standards and when all peptides were taken into account, so measurements without prior knowledge of the identified peptides. This was to be expected since the intensity of the protein is dominated by its most intense peptides which are not necessarily the best signature peptides for the protein to be used in a quantitative method.

The CVs were only significantly lower in CSF than in serum when measured by the qual/quant method without internal standards taken only two peptides into account. All other methods also showed lower CVs for CSF but these differences were not significant. To determine the lowest possible CVs, we calculated the CV for the pure internal standards without biological sample added. For the quantitative method, the CVs in water for the

highest abundant serum / CSF proteins are in the same range as the CVs for these proteins in serum or CSF. For the qual/quant method on the other hand the CVs in water are clearly lower indicating a larger influence of the other compounds present in the biological sample on the CV.

We examined why peptide measurements by qual/quant approaches typically have a lower reproducibility than peptide measurements using a targeted quantitative approach. To this end we examined peptide properties that influence digestion and chromatographic separation, as well as the intensity of the measured peptides for their influence on the CV. Factors that showed a significant influence on the CV were the presence of modifications, the presence of a cysteine and the presence of a methionine. The presence of missed cleavages did not have a significant influence per se, but the closely related length of the peptide was of significant influence on the CV. Related to the chromatographic separation the influence of co-eluting peptides was examined, which was found not to be significant. The CV was surprisingly also not influenced by the intensity of the peptides. This could be due to the fact that it was based on a very high abundant protein (albumin). With exception of the last two factors mentioned, the described factors are also the factors taken into account when peptides for MRM assays are selected[18]. Applying these criteria to two independent data sets resulted in an improved average correlation between MRM determined protein concentration and average peptide intensity in the qual/quant measurements. In addition, we observed that the peptides with the lowest R^2 were removed from the dataset. We therefore conclude that peptides fulfilling the general criteria for peptide selection for MRM are also the most reproducible peptides in qual/quant measurements and that the quantitative performance of qual/quant measurements is significantly improved when these criteria in combination with the removal of peptides with missing values is applied to a dataset.

Recently, new instrumentation has become available which is even better suited for qual/quant approaches. Both the new types of Q-TOF instruments and the Q-Orbitrap (Q-Exactive) have improved quantitative capabilities[19]. The advantages of these instruments are that the resolution for both MS and MS/MS is improved which generates theoretically an additional selectivity. New mass spectrometry methods like all ion fragmentation and MS^E on a high resolution instrument are used on these instruments to perform fragmentation on all peptides present in an MS scan[20, 21]. The high resolution of all ion fragment spectra can be used to quantify peptides based on specific fragments. For the quantitation of small molecules a shift is already observed from MRM to qual/quant using high resolution MS[22, 23]. At the moment however quadrupole ion trap instruments remain the mass spectrometer of choice for quantitation of peptides in complex samples. The data presented in this manuscript confirms that the most reproducible quantitative results are obtained using a Qtrap-platform. Also in terms of sample throughput this

platform outperformed the Orbitrap-platform using the qual/quan method. However, we showed that the quantitative capabilities of a qual/quan approach can considerably be improved by applying previously proposed filtering criteria for peptides and the addition of internal standards. The quantitative performance for the proteins for which internal standards have been spiked in the samples is improved and absolute quantitation for spiked peptides/proteins is possible. The quantitative results obtained from a limited number of proteins can be used for normalization of the samples, to decrease the run-to-run variability in long samples sequences. Additionally, internal standards can be used to monitor the quality of the data in time. Mass spectrometry response, retention time reproducibility and other chromatographic parameters can even be fully automatically determined for each individual sample using the signals of the internal standards. Besides the quantitative data of proteins of interest also a large shotgun proteomics data set remains which can easily be used for reanalyzes of other proteins.

We have shown that qual/quant in combination with modern LC and high resolution measurements could offer an alternative for the quantification of abundant proteins in complex samples. In the near future, we foresee an improvement in specifications of qual/quant methods and qual/quant becoming an equivalent method to pure quantitative methods. The largest challenge in this technology is still to reach the ng/ml range.

REFERENCES

1. Baty JD, Robinson PR: **Single and multiple ion recording techniques for the analysis of diphenylhydantoin and its major metabolite in plasma.** *Biomed Mass Spectrom* 1977, **4:**36-41.

2. Gallien S, Duriez E, Domon B: **Selected reaction monitoring applied to proteomics.** *J Mass Spectrom* 2011, **46:**298-312.

3. Anderson L, Hunter CL: **Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins.** *Mol Cell Proteomics* 2006, **5:**573-588.

4. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M: **Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.** *Mol Cell Proteomics* 2002, **1:**376-386.

5. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R: **Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.** *Nat Biotechnol* 1999, **17:**994-999.

6. DeSouza LV, Taylor AM, Li W, Minkoff MS, Romaschin AD, Colgan TJ, Siu KW: **Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues.** *J Proteome Res* 2008, **7:**3525-3534.

7. Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C: **Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus.** *Anal Chem* 2001, **73:**2836-2842.

8. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP: **Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS.** *Proc Natl Acad Sci U S A* 2003, **100:**6940-6945.

9. IJsselstijn L, Dekker LJ, Koudstaal PJ, Hofman A, Sillevis Smitt P, Breteler MM, Luider T: **Serum Clusterin Levels are not increased in Presymptomatic Alzheimer's Disease.** *J Proteome Res* 2011.

10. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham AJ, Keshishian H, et al: **Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma.** *Nat Biotechnol* 2009, **27:**633-641.

11. Domon B, Aebersold R: **Options and considerations when selecting a quantitative proteomics strategy.** *Nat Biotechnol* 2010, **28:**710-721.

12. de Groot CJ, Guzel C, Steegers-Theunissen RP, de Maat M, Derkx P, Roes EM, Heeren RM, Luider TM, Steegers EA: **Specific peptides identified by mass spectrometry in placental tissue from pregnancies complicated by early onset preeclampsia attained by laser capture dissection.** *Proteomics Clin Appl* 2007, **1:**325-335.

13. Guzel C, Ursem NT, Dekker LJ, Derkx P, Joore J, van Dijk E, Ligtvoet G, Steegers EA, Luider TM: **Multiple reaction monitoring assay for pre-eclampsia related calcyclin peptides in formalin fixed paraffin embedded placenta.** *J Proteome Res* 2011, **10:**3274-3282.

14. IJsselstijn L, Dekker LJ, Stingl C, van der Weiden MM, Hofman A, Kros JM, Koudstaal PJ, Sillevis Smitt PA, Ikram MA, Breteler MM, Luider TM: **Serum Levels of Pregnancy Zone Protein are Elevated in Presymptomatic Alzheimer's Disease.** *J Proteome Res* 2011.

15. Stoop MP, Coulier L, Rosenling T, Shi S, Smolinska AM, Buydens L, Ampt K, Stingl C, Dane A, Muilwijk B, et al: **Quantitative proteomics and metabolomics analysis of normal human cerebrospinal fluid samples.** *Mol Cell Proteomics* 2010, **9:**2063-2075.

16. Stoop MP, Singh V, Dekker LJ, Titulaer MK, Stingl C, Burgers PC, Sillevis Smitt PA, Hintzen RQ, Luider TM: **Proteomics comparison of cerebrospinal fluid of relapsing remitting and primary progressive multiple sclerosis.** *PLoS One* 2010, **5:**e12442.

17. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ: **Skyline: an open source document editor for creating and analyzing targeted proteomics experiments.** *Bioinformatics* 2010, **26:**966-968.

18. Han B, Higgs RE: **Proteomics: from hypothesis to quantitative assay on a single platform. Guidelines for developing MRM assays using ion trap mass spectrometers.** *Brief Funct Genomic Proteomic* 2008, **7:**340-354.

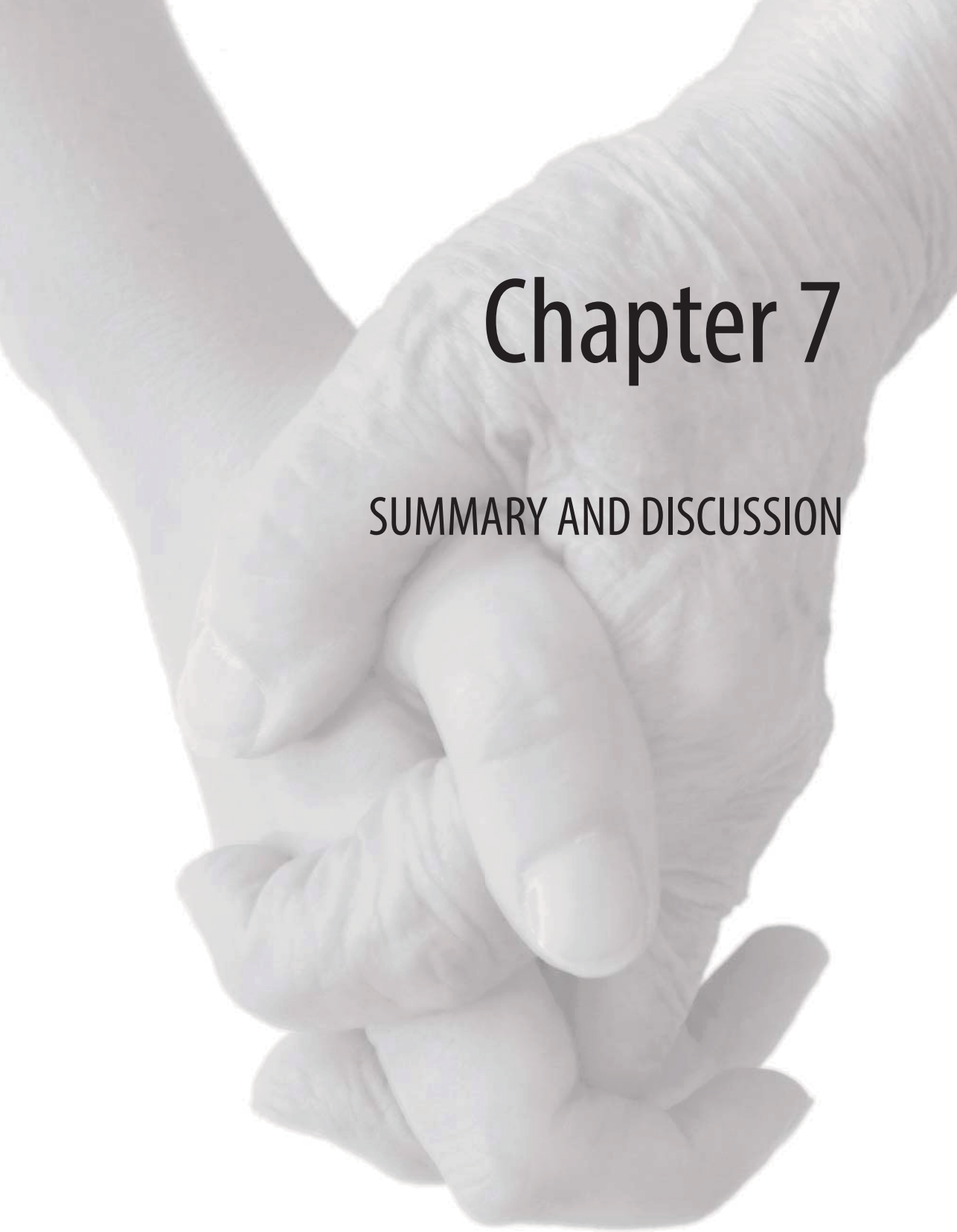
19. Gallien S, Duriez E, Crone C, Kellmann M, Moehring T, Domon B: **Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer.** *Mol Cell Proteomics* 2012, **11:**1709-1723.

20. Geiger T, Cox J, Mann M: **Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation.** *Mol Cell Proteomics* 2010, **9:**2252-2261.

21. Levin Y, Hradetzky E, Bahn S: **Quantification of proteins using data-independent analysis (MSE) in simple and complex samples: a systematic evaluation.** *Proteomics* 2011, **11:**3273-3287.

22. Ramanathan R, Jemal M, Ramagiri S, Xia YQ, Humpreys WG, Olah T, Korfmacher WA: **It is time for a paradigm shift in drug discovery bioanalysis: from SRM to HRMS.** *J Mass Spectrom* 2011, **46:**595-601.

23. Kaufmann A, Butcher P, Maden K, Walker S, Widmer M: **Quantitative and confirmative performance of liquid chromatography coupled to high-resolution mass spectrometry compared to tandem mass spectrometry.** *Rapid Commun Mass Spectrom* 2011, **25:**979-992.



Chapter 7

SUMMARY AND DISCUSSION

Dementia is not a single disease of the brain, but rather a syndrome in which memory, thinking, orientation, language and judgment are affected. Dementia mainly affects the elderly and it is estimated that by the year 2050 there will be 115.4 million people with dementia around the world. Alzheimer's disease (AD) is with 50-70% the most common type of dementia. Mild cognitive impairment (MCI) refers to the transitional stage between normal aging and dementia and is frequently a preliminary state of AD. As described in **chapter 1**, protein biomarker research in dementia is mainly performed on brain tissue and cerebrospinal fluid (CSF), and is often focused on end-stage dementia. The overall goal of this thesis is to search for proteins that are differentially expressed in serum of people who later develop AD or people with MCI, using various proteomics techniques.

Chapter 2 presents the search for proteins which are able to predict the onset of AD using serum samples from the Rotterdam Scan Study (RSS), a population based prospective cohort study. From the RSS, 43 persons who developed AD, after an average 4.2 years (±2.6 years SD) after blood sampling, and 43 gender and age matched controls who remained dementia free during follow-up were selected of which the serum proteome was examined using mass spectrometric techniques. Persons who were to develop AD (presymptomatic AD) were found to have significantly higher pregnancy zone protein (PZP) levels compared to controls. The difference was significant in women, not in men. It could well be that a significant difference is also present in males, but this was not detected in this study due to the small number of males included and the fact the majority of the measured PZP levels in males, which are generally lower than in females, were below the quantification limit of the developed multiple reaction monitoring (MRM) assay. No correlation was found between the measured PZP levels and the period between blood sampling and the diagnosis of AD. The primary structure of PZP is similar to alpha-2-macroglobulin (A2MG). A2MG was found to be elevated in plasma of AD patients and to be present in plaques in the brain of AD patients, where it forms complexes with amyloid-β peptides. PZP and A2MG are both ligands for the low-density lipoprotein receptor-related protein (LRP) receptor, which is also true for apolipoprotein E (APOE). Immunohistochemical validation on a limited set of brain tissue sections showed PZP expression in senile plaques and a subset of (micro)glial cells, suggesting a possible role for PZP in AD.

The finding of PZP expression in the brain of AD patients as described in chapter 2, was further examined and presented in **chapter 3.** The expression pattern of PZP in AD (n=20) and control (n=7) superior frontal gyrus (SFG) was examined via immunohistochemistry. The specificity of the PZP antibody was confirmed using mass spectrometry and cross reactivity with A2MG was excluded. PZP immunoreactivity was observed in microglia in both AD and control cases, and occasionally immunoreactive neurons were observed in

AD SFG. A statistically significant increase of both the amount of immunoreactive cells and their staining intensity was apparent in AD compared to control cases in the SFG cortex. In the neighboring white matter the same trend was observed though not significant. Furthermore, PZP immunoreactive cells were found in close association with senile plaques in AD SFG. These cells were presumed to be microglia, based on their location and our observation that PZP is predominantly expressed in microglia. The expression of PZP in microglia, in addition to its similarity with A2MG, suggests that PZP, via the LRP receptor, could be involved in the clearance or deposition of amyloid β. The immunohistochemical data combined with the previous findings described in chapter 2 are suggestive of an early involvement of PZP in AD.

Clusterin has been of renewed interest in AD research since genome-wide association studies identified polymorphisms in the clusterin gene to be associated with AD. Thambisetty *et al*. found increased plasma levels of clusterin to be associated with severity and progression of AD[1]. **Chapter 4** presents the investigation of the applicability of clusterin levels in the prediction of AD. The same serum samples from the RSS were used as in the search for differentially expressed proteins in presymptomatic AD (described in chapter 2) and the clusterin levels were measured with a newly developed MRM assay. No difference in clusterin levels between persons with presymptomatic AD and gender and aged matched controls was found. Thambisetty *et al.* reported, besides the association of plasma clusterin levels in the progression and severity of AD, also an association between plasma clusterin concentrations and the amount of amyloid deposition in the brain assessed 10 years later. Based on these results they implicated a role for clusterin in the very early process of AD. Our result of unchanged serum clusterin levels in presymptomatic AD patients did not support their hypothesis. Another report from the RSS where clusterin levels were measured by a multiplex immunoassay in a larger set of samples, did show association with the presence and progression of AD, but also showed no association with the development or prediction of AD[2].

Chapter 5 presents a study where the serum proteome of 35 amnestic MCI patients and 35 cognitively healthy persons was investigated by liquid chromatography mass spectrometry. This resulted in the selection of seven proteins for further quantitative analysis using MRM. The quantitative protein measurements showed significantly decreased levels in serum from amnestic MCI patients compared with cognitive healthy controls of the following five proteins: carboxypeptidase N subunit 2 (CPN2), complement factor B (CFAB), galactin-3 binding protein (LG3BP), serum amyloid A-4 protein (SAA4) and serum amyloid P-component (SAMP). Based on the serum levels of the three most significantly decreased proteins CFAB, LG3BP and SAA4, a model was built that could

differentiate between MCI patients and controls with a sensitivity of 73% and a specificity of 66%. When this model was applied to a small validation set of 13 MCI patients and 4 controls, it showed a sensitivity and specificity of 77% and 75%, respectively. Of the five proteins that were significantly different between persons with amnestic MCI and controls with normal cognition, only SAMP had a direct link to dementia. SAMP is a glycoprotein which is present in amyloid deposits and prevents their proteolysis. The other four proteins are all involved in complement or immune reactions. Confirmation of our findings in a larger external validation set is needed and to unravel the role of CFAB, CPN2, LG3BP, SAA4 and SAMP in amnestic MCI more research is required.

Chapter 6 describes a study in which the quantitative performance of two different mass spectrometric techniques was examined. The purely quantitative MRM method was compared with a method in which qualitative and quantitative aspects are combined (qual/quant). The qual/quant measurements were performed on an Orbitrap MS which is routinely used for protein identification (see chapter 2 and 5) with and without the use of internal standards. The linearity and reproducibility of the methods were determined in serum and CSF sample sets by measuring the quantitative levels of six relatively abundant proteins. The linearity of the measured albumin concentrations in CSF compared to an antibody-based technique was best for the quantitative method, followed by the qual/quant method with use of internal standards and without use of internal standards. Although the linearity measurements showed that the qual/quant method was the least accurate method, there was still a regression coefficient (R^2) of 0.898 making it a good method to give a first and rather accurate quantitative impression. In addition to its better linearity, the quantitative method also showed a better reproducibility with lower CVs for both serum and CSF samples. Next to these comparisons the question why a qual/quant approach has typically a lower reproducibility compared to MRM analyses was addressed. Factors that showed a significant influence were the presence of modifications and the presence of a cysteine or a methionine in the peptides. The presence of missed cleavages did not have a significant influence per se, but the closely related length of the peptide was of significant influence. Filtering out peptides based on the above described criteria resulted in a better correlation between the qual/quant and the quantitative data as demonstrated on data of apolipoprotein E measured in chapter 5. These criteria were also applied to the data of PZP (chapter 2). In this case however, no increase in correlation was observed. This could be due to the fact that PZP is a low abundant protein and that the signature peptide chosen for the MRM measurements was not the most appropriate. In conclusion, qual/quant could offer an alternative for the quantification of abundant proteins in complex samples. Further improvement of methods and equipment is still needed to reach the ng/ml range.

Future perspectives

The ultimate goal of research in dementia is to unravel the disease mechanism and find a treatment to help dementia patients. The increase in serum PZP in presymptomatic AD patients and the expression of PZP in AD brains are interesting findings but the contribution of these findings to the ultimate goal is still unknown. To understand the role of PZP in AD, information is needed about the interaction of PZP with other proteins. This can be investigated by identifying proteins in senile plaques that interact with PZP. To this end, a method combining laser capture microdissection, immunoprecipitation and mass spectrometry could be used[3]. Mass spectrometry could also be deployed in the determination of the activity of PZP. Chymotrypsin is among others a substrate for PZP[4]. By the binding of chymotrypsin to PZP, chymotrypsin will become inactive and therefore unable to cleave proteins. Monitoring of the cleavage products using MRM could be a method to indirectly determine the activity of PZP. Next to information about the interaction of PZP with other proteins, further research is also needed regarding the serum levels of PZP. The increase of PZP in presymptomatic AD patients needs to be confirmed in a larger independent sample set which preferably also includes sera from other groups like AD patients and MCI patients. Such studies will show the applicability of PZP as a biomarker for presymptomatic AD.

In general the search for biomarkers is a challenging one. Many differentially expressed proteins are found, but only few meet the criteria of a biomarker. In complex diseases like dementia one should probably not focus on a single biomarker but on a combination of multiple biomarkers. This could be a combination of multiple differentially expressed proteins, making models as described in chapter 6 for amnestic MCI, or a combination of biomarkers from different research fields like proteomics, genomics and imaging. The motive for the clusterin study described in chapter 5 was a study in which genomics and proteomics data were combined.

Implementation of protein biomarkers in a clinical setting demands a technique that is able to measure protein levels with a high sensitivity and high sample throughput. There are approximately 200 clinical assays for proteins in plasma and serum approved by the United States Food and Drugs Administration (FDA) which are almost all immunoassays[5]. Although MRM has been a reference technique to analyze small molecules for many years, mass spectrometry is only occasionally used for the analysis of proteins. To create a field called Clinical Chemistry Proteomics, quantitative mass spectrometric techniques need to be further optimized[6]. The biggest challenge is dealing with the complexity of the samples. Solutions could lie in better controlled pre-analytical procedures like enzymatic digestion, protein precipitation, and the use of antibodies to capture the analyte. These procedures however introduce a degree of variability. The development of the mass

spectrometers themselves is also still ongoing resulting in more sensitive and reproducible methods. Clinical Chemistry Proteomics is an emerging field, but in the near future mass spectrometers will remain in the background.

REFERENCES

1. Thambisetty M, Simmons A, Velayudhan L, Hye A, Campbell J, Zhang Y, Wahlund LO, Westman E, Kinsey A, Guntert A, et al: **Association of plasma clusterin concentration with severity, pathology, and progression in Alzheimer disease.** *Arch Gen Psychiatry* 2010, **67:**739-748.

2. Schrijvers EM, Koudstaal PJ, Hofman A, Breteler MM: **Plasma clusterin and the risk of Alzheimer disease.** *Jama* 2011, **305:**1322-1326.

3. Stingl C, van Vilsteren FG, Guzel C, Ten Kate FJ, Visser M, Krishnadath KK, Bergman JJ, Luider TM: **Reproducibility of protein identification of selected cell types in Barrett's esophagus analyzed by combining laser-capture microdissection and mass spectrometry.** *J Proteome Res* 2011, **10:**288-298.

4. Christensen U, Simonsen M, Harrit N, Sottrup-Jensen L: **Pregnancy zone protein, a proteinase-binding macroglobulin. Interactions with proteinases and methylamine.** *Biochemistry* 1989, **28:**9324-9331.

5. Anderson NL: **The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum.** *Clin Chem* 2010, **56:**177-185.

6. Lehmann S, Hoofnagle A, Hochstrasser D, Brede C, Glueckmann M, Cocho JA, Ceglarek U, Lenz C, Vialaret J, Scherl A, Hirtz C: **Quantitative Clinical Chemistry Proteomics (qCCP) using mass spectrometry: general characteristics and application.** *Clin Chem Lab Med* 2013, **51:**919-935.

Appendices

SAMENVATTING
DANKWOORD
LIST OF PUBLICATIONS
PHD PORTFOLIO
CURRICULUM VITAE

Samenvatting

Dementie is een syndroom waarin het geheugen, denkvermogen, oriëntatie, taalgebruik en beoordelingsvermogen zijn aangetast. Dementie treft over het algemeen ouderen en geschat wordt dat er in het jaar 2050 115,4 miljoen mensen wereldwijd lijden aan dementie. De ziekte van Alzheimer (AD) is met 50-70% de meest voorkomende vorm van dementie. Mild cognitive impairment (MCI) refereert naar het overgangsstadium van normale veroudering naar dementie en is vaak het voorstadium van AD. Zoals beschreven in **hoofdstuk 1**, wordt biomarkeronderzoek met name uitgevoerd in hersenweefsel en hersenvocht (CSF) en ligt de focus vaak op het eindstadium van dementie. Het doel van dit proefschrift was het zoeken naar eiwitten die verschillend tot expressie komen in serum van personen die later AD ontwikkelden of van personen met MCI hierbij gebruik makend van verschillende proteomics technieken.

Hoofdstuk 2 presenteert de zoektocht naar eiwitten die het begin van AD kunnen voorspellen in serum monsters van de Rotterdam Scan Study (RSS), een prospectieve populatie cohort studie. Van de RSS studie, 43 personen die AD ontwikkelden na gemiddeld 4,2 jaar (±2.6 jaar SD) na de bloedafname en 43 geslacht en leeftijd gematchte controles werden geselecteerd waarvan het serum proteome werd onderzocht met behulp van massaspectrometrische technieken. Personen die later AD ontwikkelden (presymptomatische AD) bleken een significant hoger niveau van het pregnancy zone protein (PZP) te hebben vergeleken met de controles. Het verschil was significant voor vrouwen, niet voor mannen. Het is goed mogelijk dat een significant verschil tevens bij mannen aanwezig is maar dat dit niet gedetecteerd is als gevolg van het lage aantal mannen in de studie en het feit dat het overgrote deel van de gemeten PZP niveaus bij mannen, welke over het algemeen lager zijn dan bij vrouwen, onder de detectiegrens van het ontwikkelde multiple reaction monitoring (MRM) assay lagen. Er werd geen correlatie gevonden tussen de gemeten PZP niveaus en de tijd verstreken tussen de bloedafname en het moment van diagnose van AD. De primaire structuur van PZP is gelijk aan die van alfa-2-microglobuline (A2MG). A2MG werd in verhoogde mate gevonden in het plasma van AD patiënten en is tevens aanwezig in de plaques in de hersenen van AD patiënten waar het complexen vormt met amyloïd-β peptiden. PZP en A2MG zijn beide liganden voor de low-density lipoprotein receptor-related protein (LRP) receptor, net als apolipoproteïne E (APOE). Immuunhistochemische validatie op een beperkte set hersenweefsel monsters laat de expressie van PZP zien in plaques en in een gedeelte van de (micro)gliale cellen wat wijst op een mogelijke rol van PZP in AD.

De ontdekking van PZP expressie in de hersenen van AD patiënten, zoals beschreven in hoofdstuk 2, werd verder onderzocht en gepresenteerd in **hoofdstuk 3**. Het expressiepatroon van PZP in AD (n=20) en controle (n=7) superior frontal gyrus (SFG) werd onderzocht met behulp van immuunhistochemie. De specificiteit van het PZP-antilichaam werd middels massaspectrometrie bevestigd en kruisreactiviteit met A2MG werd uitgesloten. PZP immuunreactiviteit werd gezien in microglia in zowel AD als in controle weefsel en af en toe werden immuunreactieve neuronen gezien in AD SFG. Een statistisch significante stijging van zowel de hoeveelheid van de immuunreactieve cellen als de intensiteit van de kleuring was zichtbaar in AD vergeleken met de controles in de SFG cortex. In de aangrenzende witte stof werd eenzelfde trend gezien echter was deze niet significant. Tevens werden PZP immuunreactieve cellen gevonden in nauwe associatie met plaques in AD SFG. Deze cellen werden verondersteld microglia te zijn, gebaseerd op hun ligging en onze observatie dat PZP voornamelijk tot expressie komt in microglia. De expressie van PZP in microglia, naast de gelijkheid met A2MG, suggereert dat PZP, via de LRP receptor, een rol zou kunnen spelen bij de opruiming of afzetting van amyloïd-β. De immuunhistochemie data in samenhang met de eerdere bevindingen beschreven in hoofdstuk 2 wijzen mogelijk op een vroege betrokkenheid van PZP in AD.

Clusterine is van hernieuwde interesse in AD onderzoek sinds genome-wide association studies aantoonden dat polymorfismen in het clusterine gen geassocieerd zijn met AD. Thambisetty *et al*. vonden een associatie tussen verhoogde plasma clusterine niveaus en de hevigheid en progressie van AD. **Hoofdstuk 4** presenteert het onderzoek naar de toepasbaarheid van clusterine niveaus in de voorspelling van AD. Dezelfde set serum monsters van de RSS werd gebruikt als in de zoektocht naar differentieel tot expressie komende eiwitten in presymptomatische AD (beschreven in hoofdstuk 2) en de clusterine niveaus werden gemeten middels een nieuw ontwikkeld MRM assay. Er werd geen verschil gevonden in clusterine niveaus tussen personen met presymptomatische AD en geslacht en leeftijd gematchte controles. Thambisetty *et al*. beschreven, naast de associatie tussen verhoogde plasma clusterine niveaus en de hevigheid en progressie van AD, ook een associatie tussen plasma clusterine niveaus en de hoeveelheid van amyloid afzetting in de hersenen na tien jaar. Ons resultaat van onveranderde serum clusterine niveaus in presymptomatische AD ondersteunt deze veronderstelling niet. Een ander verslag van de RSS, waarin clusterine niveaus werden gemeten middels multiplex immuunassays in een grotere set monsters, liet een associatie zien met de aanwezigheid en progressie van AD, maar eveneens geen associatie met de ontwikkeling of voorspelling van AD.

Hoofdstuk 5 presenteert een onderzoek waarbij het serum proteome van 35 amnestische MCI patiënten en 35 cognitief gezonde personen onderzocht werd met behulp van

vloeistofchromatografie gekoppeld aan massaspectrometrie. Dit onderzoek resulteerde in een selectie van zeven eiwitten voor verdere kwantitatieve analyse middels MRM. De kwantitatieve eiwitmetingen lieten een significant lager niveau in serum zien in amnestische MCI patiënten vergeleken met cognitief gezonde personen van de volgende vijf eiwitten: carboxypeptidase N subunit 2 (CPN2), complement factor B (CFAB), galactin-3 binding protein (LG3BP), serum amyloid A-4 protein (SAA4) and serum amyloid P-component (SAMP). Op basis van de serum niveaus van de drie meest significant lagere eiwitten CFAB, LG3BP en SAA4, werd een model gebouwd dat onderscheid kon maken tussen MCI patiënten en controles met een sensitiviteit van 73% en een specificiteit van 66%. Wanneer dit model werd toegepast op een kleine validatie set van 13 MCI patiënten en 4 controles, werd er respectievelijk een sensitiviteit en specificiteit van 77% en 75% gezien. Van de vijf eiwitten die significant verschillend werden gevonden tussen personen met amnestische MCI en controles met normale cognitie, heeft alleen SAMP een directe relatie met dementie. SAMP is namelijk een glycoproteïne wat aanwezig is in amyloïde afzettingen en de afbraak hiervan voorkomt. De overige vier eiwitten zijn allen betrokken bij complement of immuun reacties. Bevestiging van onze bevindingen in een grotere externe validatieset is nodig en voor het ontrafelen van de rol van CFAB, CPN2, LG3BP, SAA4 en SAMP in amnestische MCI is verder onderzoek vereist.

Hoofdstuk 6 beschrijft een studie waarbij de kwantitatieve prestaties van twee verschillende massaspectrometrietechnieken werden onderzocht. De uitsluitend kwantitatieve MRM methode werd vergeleken met een methode waarin kwalitatieve en kwantitatieve aspecten zijn gecombineerd (qual/quant). De qual/quant metingen werden uitgevoerd met behulp van een Orbitrap MS, welke routinematig wordt gebruikt voor eiwit identificatie (zie hoofdstuk 2 en 5), met en zonder gebruik van interne standaarden. De lineariteit en reproduceerbaarheid van de methoden werden bepaald middels het meten van de kwantitatieve eiwit niveaus van zes relatief hoog abundant eiwitten in serum en CSF monsters. De lineariteit van de gemeten albumine niveaus in CSF vergeleken met een antilichaam gebaseerde methode was het beste voor de kwantitatieve methode, gevolgd door de qual/quant methode met gebruik van interne standaarden en zonder gebruik van interne standaarden. Alhoewel de lineariteit metingen lieten zien dat de qual/quant methode de minst accurate methode was, was de regressiecoëfficiënt (R^2) wel 0.898 waardoor het een goede methode is voor een eerste en redelijk accurate kwantitatieve impressie. Behalve de betere lineariteit liet de kwantitatieve methode ook een betere reproduceerbaarheid zien met lagere CVs in zowel de serum als CSF monsters. Naast deze vergelijkingen werd er ook gekeken naar het feit dat de qual/quant methode typisch resulteert in een lagere reproduceerbaarheid vergeleken met MRM analyses. Factoren die een significante invloed lieten zien waren de aanwezigheid van modificaties

en de aanwezigheid van een cysteïne of methionine in de peptiden. De aanwezigheid van missed cleavages was op zichzelf niet van significante invloed, maar de nauw gerelateerde lengte van de peptiden was dit wel. Het weglaten van peptiden gebaseerd op bovenstaande criteria, resulteerde in een betere correlatie tussen de qual/quant en de kwalitatieve data zoals gedemonstreerd werd op data van apolipoproteïne E gemeten in hoofdstuk 5. Deze criteria werden tevens toegepast op de data van PZP (hoofdstuk 2), echter werd hier geen toename in correlatie gezien. Dit kan het gevolg zijn van het feit dat PZP een laag abundant eiwit is en dat het peptide gekozen voor de MRM metingen niet de meest geschikte bleek. Concluderend, qual/quant biedt een alternatief voor het kwantificeren van abundant eiwitten in complexe monsters. Verdere verbetering van methoden en apparatuur is nog altijd nodig om het ng/ml niveau te bereiken.

In **hoofdstuk 7** worden de belangrijkste bevinden van mijn proefschrift beschreven en doe ik suggesties voor toekomstig onderzoek betreffende proteomics in dementie.

Dankwoord

Met het schrijven van dit dankwoord komt er een einde aan een belangrijke periode in mijn leven. Graag wil ik iedereen die direct of indirect heeft bijgedragen aan deze mooie, leerzame, maar soms ook lastige tijd, bedanken en een aantal personen in het bijzonder.

Allereerst mijn co-promotor, Theo Luider. Beste Theo, ik wil je bedanken voor alle kansen en mogelijkheden die je me hebt geboden tijdens mijn onderzoek. Je gave om bij tegenvallende resultaten toch positief te blijven, zal me altijd bijblijven.

Mijn promotor, Peter Sillevis Smitt. Beste Peter, dank dat ik binnen de neurologie en onder jouw hoede mijn onderzoek mocht uitvoeren. Je altijd opbouwende kritiek en je heldere kijk op de zaken heb ik zeer gewaardeerd.

De leescommissie, Peter Koudstaal, Max Kros en Arfan Ikram, bedankt voor de tijd en moeite die jullie hebben gestoken in het beoordelen van mijn proefschrift. Tevens dank voor de fijne samenwerking.

Uiteraard gaat mijn dank ook uit naar alle medeauteurs voor de prettige samenwerking en het komen tot mooie publicaties. Diana, fijn dat je het stokje hebt overgenomen van hoofdstuk 3, het is een prachtig hoofdstuk geworden.

Graag wil ik alle (oud-)collega's van de proteomics groep bedanken voor alle medewerking en gezellige borrels, congressen en kletspraatjes. Linda en Marieke, bedankt voor het delen van al het (promotie) lief en leed.

Lieve paranimfen, heel fijn dat jullie vandaag naast mij staan. Lennard, ik heb ontzettend veel van je geleerd, bedankt! Dominique, het is ons toch maar mooi gelukt en erg leuk dat we voor altijd collega's zullen zijn.

Lieve familie en vrienden, hoewel jullie vaak geen idee hadden wat ik aan het doen was, waren jullie altijd geïnteresseerd en bereid mijn verhalen aan te horen. Dank jullie wel hiervoor en voor alle gezellige momenten waarop ik even niet aan mijn promotie hoefde te denken. Lieve oma, dementie heeft ook u te pakken. Doordat we samen op de omslag staan, voelt het toch alsof u er bij bent. Lieve pap en mam, ik weet dat jullie trots op me zijn.

Tot slot de personen die me het meest dierbaar zijn. Lieve Michel, dank je wel voor al je liefde, vertrouwen en brede schouders. En natuurlijk voor onze prachtige dochter. Lieve Féline, mama is de mooiste titel die er bestaat! Dikke kus!

List of publications

L. IJsselstijn, J.M. Papma, L.J.M. Dekker, W. Calame, C. Stingl, P.J. Koudstaal, P.A.E. Sillevis Smitt, T.M. Luider. *Serum proteomics in amnestic mild cognitive impairment.* Proteomics, 2013, 13(16):2526-2533.

L. IJsselstijn, M.P. Stoop, C. Stingl, P.A.E. Sillevis Smitt, T.M. Luider, L.J.M. Dekker. *A comparative study of targeted and label-free mass spectrometry methods for protein quantification.* Journal of Proteome Research, 2013, 12(4):2005-2011.

L. IJsselstijn, L.J.M. Dekker, C. Stingl, M.M. van der Weiden, A. Hofman, J.M.Kros, P.J. Koudstaal, P.A.E. Sillevis Smitt, M.A. Ikram, M.M.B. Breteler, T.M. Luider. *Serum levels of pregnancy zone protein are elevated in presymptomatic Alzheimer's disease.* Journal of Proteome Research, 2011, 10(11):4902-4910.

L. IJsselstijn, L.J.M. Dekker, P.J. Koudstaal, A. Hofman, P.A.E. Sillevis Smitt, M.M.B. Breteler, T.M. Luider. *Serum clusterin levels are not increased in presymptomatic Alzheimer's disease.* Journal of Proteome Research, 2011, 10(4):2006-2010.

A.A. Dihal, H. van der Woude, P.J. Hendriksen, H. Charif, L.J.M. Dekker, **L. IJsselstijn**, V.C. de Boer, G.M. Alink, P.C. Burgers, I.M. Rietjens, R.A. Woutersen, R.H. Stierum. *Transcriptome and proteome profiling of colon mucosa from quercetin fed F344 rats point to tumor preventive mechanisms, increased mitochondrial fatty acid degradation and decreased glycolysis.* Proteomics, 2008, 8(1):45-61.

PhD portfolio

Name PhD student:	Linda IJsselstijn
Erasmus MC departement:	Neurology
PhD period:	April 1, 2007 – July 1, 2011
Research school:	Erasmus Postgraduate school Molecular Medicine
Promoter:	prof. dr. P.A.E. Sillevis Smitt
Co-promotor:	dr. T.M. Luider

In-depth courses

- Bioinformatics for Protein Identification, 59th Conference of the American Society for Mass Spectrometry, 2011 (0.6 ECTS).
- Metabolomics, 58th Conference of the American Society for Mass Spectrometry, 2010 (0.3 ECTS).
- Regression Analysis, Erasmus Summer Programme, 2009 (1.1 ECTS).
- Introduction to Data-analysis, Erasmus Summer Programme, 2009 (0.9 ECTS).
- Principles of Research in Medicine and Epidemiology, Erasmus Summer Programme, 2009 (0.5 ECTS).
- Principles and Applications of Fourier Transform Mass Spectrometry, 57th Conference of the American Society for Mass Spectrometry, 2009 (0.6 ECTS).
- Case Studies in Quantitative Proteomics, 56th Conference of the American Society for Mass Spectrometry, 2008 (0.6 ECTS).
- Biomedical English Writing and Communication, Erasmus MC, 2008 (4 ECTS).
- Ultraflex III (MALDI-TOF/TOF) Operator Training: WARP-LC and flex-Imaging, Bruker Daltonik, 2007 (0.8 ECTS).
- Quadrupole Ion Trap Mass Spectrometry, 55th Conference of the American Society for Mass Spectrometry, 2007 (0.6 ECTS).

(Inter)national conferences and presentations

- 59th Conference of the American Society for Mass Spectrometry, Denver, CO, USA, 2011. Poster presentation (0.4 ECTS).
- 15th Molecular Medicine Day, Rotterdam, 2011. Poster presentation (0.4 ECTS).
- 58th Conference of the American Society for Mass Spectrometry, Salt Lake City, UT, USA, 2010. Poster presentation (0.4 ECTS).
- 9th Alzheimer's Association International Conference on Alzheimer's Disease, Vienna, Austria, 2009. Poster presentation (0.4 ECTS).

- 57th Conference of the American Society for Mass Spectrometry, Philadelphia, PA, USA, 2009. Poster presentation (0.4 ECTS).
- 7th HUPO World Congress, Amsterdam, 2008. Poster presentation (0.4 ECTS).
- 56th Conference of the American Society for Mass Spectrometry, Denver, CO, USA, 2008. Poster presentation (0.4 ECTS).
- 55th Conference of the American Society for Mass Spectrometry, 2007 (0.4 ECTS).

Curriculum Vitae

Linda IJsselstijn was born on March 1, 1982 in Rotterdam, the Netherlands. She finished her pre-university education at the Sint Laurens College in Rotterdam in 2000, and started university in the same year. She studied Biomedical Engineering at the Eindhoven University of Technology. She obtained her bachelor degree in 2003 and for her master she focused on Molecular Bioengineering. She performed an internship at the department of Pharmacology and Toxicology at the University of Maastricht, where she examined the anti-oxidant capacity of erythritol. In 2004, she spent three months in Athens, Greece, for an internship at the University of Athens, studying the aggregation of polystyrene with amine end groups. For her master thesis she worked on amyloid β-dendrimer complexes under guidance of dr. M. Merkx. She graduated from the Eindhoven University of Technology in 2006 under final supervision of prof. dr. ir. E.W. Meijer.

In 2006 she started working with mass spectrometric techniques in the group of dr. T.M. Luider at the department of Neurology at the Erasmus University Medical Center, Rotterdam. She started her PhD research in April 2007 of which the results are described in the current thesis under supervision of dr. T.M. Luider and prof. dr. P.A.E. Sillevis Smitt.

Since December 2011 she is working as a resident in Clinical Chemistry at the Maasstad Hospital in Rotterdam.