Aus der Augenklinik und Poliklinik
der Universitätsmedizin der Johannes Gutenberg-Universität Mainz

Proteomic analysis of human cerebrospinal fluid of patients with major depressive disorder and schizophrenia

(Proteomische Analyse humanen Liquor cerebrospinalis von Proband:innen mit Depressionen und Schizophrenie)

Inauguraldissertation
zur Erlangung des Doktorgrades der Zahnmedizin
der Universitätsmedizin
der Johannes Gutenberg-Universität Mainz

Vorgelegt von

[redacted]
aus [redacted]

Mainz 2022

Wissenschaftlicher Vorstand: [redacted]

1. Gutachter: [redacted]

2. Gutachter: [redacted]

Tag der Promotion: 06. Dezember 2022

Table of Contents

List of Abbreviations

%	Percentage
°C	Degree Celsius
µg	Microgram
µl	Microlitre
1DE	One-dimensional gel electrophoresis
2DE	Two-dimensional gel electrophoresis
AA	Atypical antipsychotic
ABC	Ammonium bicarbonate
AD	Alzheimer's disease
ACN	Acetonitrile
APR	Acute phase respond
B	Benzodiazepine
BPD	Bipolar disorder
CNS	Central nervous system
A2M	Alpha-2 Macroglobulin
ALDOA	Fructose-bisphosphate aldolase A (),
ALDOC	Fructose-bisphosphate aldolase C
ANKRD36C	Ankyrin repeat domain-containing protein 36C
APOA1	Apolipoprotein A1
APOA2	Apolipoprotein A2
APOA4	Apolipoprotein A4
APOE	Apolipoprotein E
APP	Amyloid-precursor-protein/Amyloid beta A4
BCAN	Brevican core protein
C1R	Complement C1r subcomponent
C5	Complement C5
C7	Complement component C7
CACNA2D1	Voltage-dependent calcium channel subunit alpha-2/delta-1
CARTPT	Cocaine- and amphetamine-regulated transcript
CBLN4	Cerebellin-4
CD14	Monocyte differentiation antigen CD14
CFD	Complement factor D
CFH	Complement factor H
CFHR1	Complement factor H-related protein 1
CHGA	Chromogranin A

CHGB	Secretogranin I
CHI3L1	Chitinase-3-like protein 1
CHL1	Neural cell adhesion molecule L1-like protein
CLSTN1	Calsyntenin-1
CNDP1	Beta-Ala-His dipeptidase
CNTN2	Contactin-2
CNTNAP4	Contactin-associated protein-like 4
COL6A1	Collagen alpha-1(VI) chain
CPE	Carboxypeptidase E
CSF	Cerebrospinal fluid
CST3	Cystatin C
CTRL	Healthy suspects
CTSD	Cathepsin D
Da	Dalton (molecular mass)
dH2O	Distilled water
DHT	Dihydrotestosterone
DTT	Dithiothreitol
e.g.	For example
ELISA	Enzyme-Linked Immunosorbent Assay
ESI	Electrospray ionization
FDR	False discovery rate
FN1	Fibronectin
g	Gram
GC	Vitamin D-binding protein
GOT1	Aspartate aminotransferase, cytoplasmic
GRIA4	Glutamate receptor 4
HRG	Decreased histidine-rich glycoprotein
HSPG2	Basement membrane-specific heparan sulfate proteoglycan core protein
i.e.	Id est (“that is to say”)
IAA	Iodoacetamide
iBAQ	Intensity-based absolute quantification
IGFBP7	Insulin-like growth factor-binding protein 7
IGHA1	Ig alpha-1 chain C region
IL-6	Interleukin-6
IL-8	Interleukin-8
IPS	Initial prodromal state of psychosis
ITIH1	Inter-alpha-trypsin inhibitor heavy chain H1

ITIH2	Inter-alpha-trypsin inhibitor heavy chain H2
ITIH4	Inter-alpha-trypsin inhibitor heavy chain H4
iTRAQ	Isobaric tags for relative and absolute quantitation
kDa	Kilodalton (molecular mass)
kV	Kilovolt
LC	Liquid chromatography
LFQ	Label-free quantitative
LGALS3BP	Galectin-3-binding protein
Li	Lithium
LINGO1	Leucine-rich repeat and immunoglobulin-like domain-containing nogo receptor-interacting protein 1
LLD	Late-life depression
LSAMP	Limbic system-associated membrane protein (
LTQ	Linear quadrupole ion trap
m/z	Mass to charge ratio
MALDI	Matrix assisted laser desorption ionization
MAOI	Monoamine oxidase inhibitor
MDD	Major depressive disorder
MDH1	Malate dehydrogenase
MEGF8	Multiple epidermal growth factor-like domains protein 8
mg	Milligram
Mi	Mirtazapine
min	Minute
ml	Millilitre
MMP2	72 kDa type IV collagenase
MOPS	3-(*N*-Morpholino)propansulfon acid
MS	Mass spectrometry
MS	Mood stabilizer
MS/MS	Tandem mass spectrometry
NEFH	Neurofilament heavy polypeptide
NEFL	Neurofilament light protein
NELL2	Protein kinase C-binding protein NELL2
NFASC	Neurofascin
NPTX1	Neuronal pentraxin-1
NPTXR	Neuronal pentraxin receptor
NRXN3	Neurexin-3
NRXN3	Neurexin-3-beta

NXPH1	Neurexophilin-1
OCD	Obsessive-compulsive disorder
OPCML	Opioid-binding protein/cell adhesion molecule
PAGE	Polyacrylamide gel electrophoresis
PAM	Peptidyl-glycine alpha-amidating monooxygenase
PCR	Polymerase chain reaction
PCSK1N	ProSAAS
PD	Parkinson’s disease
PLEKHH2	Pleckstrin homology domain-containing family H member 2
PLG	Plasminogen
PON1	Serum paraoxonase/arylesterase 1
ppm	Parts per million
PTGDS	Prostaglandin-H2 D-isomerase
PTM	Post-translational modification
R	Replicate
RNA	Ribonucleic acid
S100beta	Protein S100beta
SAD	Schizoaffective disorder
SCG2	Secretogranin II
SCZ	Schizophrenia
SDS	Sodium dodecyl sulfate
SELDI	Surface enhanced laser desorption ionization
SERPINF1	Pigment epithelium-derived factor
SERPINI1	Neuroserpin
SEZ6	Seizure protein 6 homolog
SEZ6L	Seizure 6-like protein
SHANK2	SH3 and multiple ankyrin repeat domains protein 2
SIRPA	Tyrosine-protein phosphatase non-receptor type substrate 1
SNRI	Selective serotonin and noradrenalin reuptake inhibitor
SRM	Selected reaction monitoring
SSRI	Selective serotonin reuptake inhibitor
T3	Triiodothyronine
T4	Thyroxine
TA	Typical antipsychotic
TCA	Tricyclic antidepressant
TFA	Trifluoroacetic acid
TOF	Time-of-flight

TSH	Thyroid stimulating hormone
TTR	Transthyretin
V	Volt
VGF	Neurosecretory protein VGF
vs.	Versus

List of Figures

List of Tables

1 Introduction

Cerebrospinal fluid (CSF) is a crystal-clear liquid that surrounds the spinal cord and the brain (1). As the only body fluid in direct contact with the brain, it is considered to present a promising source in the investigation of neurological diseases (2). Since most of the biological functions in the human body are regulated by proteins which are very complex and dynamic, the analysis of the proteome is a promising approach in the study of pathological processes (3-5). Proteins can undergo various forms of modifications. They can be synthesized or degraded, adjusting to the current state of the organism (6). Alterations in the CSF proteome may thus reflect pathologies of the brain (7).

The field of proteomics has been progressing rapidly due to major developments in technology. Great improvements in mass spectrometry (MS) allow the identification of hundreds to thousands of proteins of a complex sample and their modifications (8). The detection of alterations in protein expression levels may help in understanding pathologies behind diseases and may enable the identification of potential biomarker candidates (9). Biomarkers are proteins that are linked to the pathophysiology of a disease and therefore may also hold information on the fundamental molecular mechanisms behind the illness (10).

Mental health issues are considered as the main cause of disability globally (11). The immense burden to wellness and lifestyle of individuals affected by neurological diseases as well as the global economic consequences necessitate progress in diagnosis and treatment (12). In the past, Alzheimer's disease has been the subject of several studies with promising results regarding the identification of potential biomarkers (7, 13-15). However, to date, the mechanisms underlying neurological conditions like schizophrenia (SCZ) or major depressive disorder (MDD) remain largely unknown (16, 17). The proteomic analysis of CSF of patients that suffer from SCZ or MDD may reveal protein alterations that can help to better understand the diseases and, thus, may contribute to higher accuracy in diagnosis and treatment in the future.

In this study, discovery proteomics strategy was employed to determine the changes in the proteome of CSF taken from patients diagnosed with SCZ and MDD. A group of individuals with no known mental diseases serves as control. One-dimensional gel electrophoresis approach was employed to reduce the complexity of CSF proteins utilizing capillary-LC-ESI-MS/MS. Generally, the aims were to investigate disease associated alterations in protein expression levels in the CSF of patients with SCZ and MDD, to define potential biomarker candidates and to learn more about the pathological processes by studying the functions of altered proteins.

Furthermore, in a second discovery study CSF from schizophrenic patients was analysed employing a different proteomic approach of in-solution digestion and nano-LC-ESI-MS/MS.

This allows the comparison of numbers of detected proteins between two established proteomic techniques in relation to the total quantity of work required for sample preparation and measurement. Moreover, it makes the investigation of potential differences of abundance in gender possible due to individual sample analysis.

2 Literature Discussion

2.1 Cerebrospinal Fluid

Cerebrospinal fluid (CSF) is a clear body fluid which surrounds the surface of the brain, the spinal cord and the ventricles (1, 18). One of its main functions is the transport of nutrition to the brain and the removal of metabolism products, thereby maintaining a constant chemical milieu. Furthermore, it protects the brain from the impact of severe mechanical forces and provides buoyancy, reducing the brain's weight (18-20).

Cerebrospinal fluid is mainly produced by the choroid plexus of the third, fourth and lateral ventricles *via* secretion (about 70% to 85%) and partly by diffusion through the brain's capillary bed and production of metabolic water (15% to 30%) (2, 20, 21). The composition of CSF is thereby regulated by the choroid plexus, a branch-like complex of rich vascularised tissue covered by a single layer of epithelial cells. It forms numerous villi which project in the lateral, third, and fourth ventricles of the brain (20, 22). The body fluid runs from the fourth ventricle into the subarachnoid spaces which are located between the pia mater and arachnoid, the two innermost layers of the membranes surrounding the brain (20). It circulates with a total volume of 100 to 160 ml. While circulating the volume is being replaced completely every six to eight hours, i.e. approximately three times a day (18). Cerebrospinal fluid's reabsorption takes place in the arachnoid granulations and villi into the venous circulation. Recent studies indicate that the lymphatic system may also partly play a role in the reabsorption process (23-25). The flow of CSF is visualized in **Figure 1**.

The blood-CSF barrier separates CSF from blood. It is a semipermeable border formed by the junctional complexes between the choroid plexus' epithelial cells (22, 26). Compared to the blood-brain barrier, the blood-CSF barrier is relatively pervious. It does allow the exchange with blood substances that may not pass the brain-blood barrier. Thereby, the rate of entering depends on the molecular weight of the molecules. For example, plasma proteins like albumin which are unable to pass the blood-brain barrier may cross the choroid plexus to enter CSF (26, 27). The protein concentration in CSF is determined by the molecule's concentration in serum, its hydrodynamic radius and the flow rate of CSF (28). However, CSF also includes polypeptides and proteins which are produced locally. Even though, it is widely believed that the majority of CSF proteins derive from blood and CSF is sometimes referred to as "ultrafiltration of blood", Zougman *et al*. (2008) suggest that the locally produced intrinsic proteins form the CSF's major proteome (19, 29). The proteins found in CSF thus are unique to the brain compartment and can reflect the state of the central nervous system (CNS) (2). Due to its proximity to brain tissue, CSF is considered a promising reservoir in the investigation of brain associated diseases (30).

Figure 1: Flow of cerebrospinal fluid *(31)*

2.1.1 CSF for the Investigation of Mental Disorders

There are not many sources available when studying pathologies of the brain in the living since working with brain tissue usually implies the usage of post-mortem samples. Moreover, brain tissue has some disadvantages. Unlike CSF, it is a static source and holds factors like age, medication or chronicity of the disease that are potentially confounding (32, 33). Furthermore, in many cases the tissue cannot be obtained and fixed directly after the patient's death. This postponement has an impact on the protein degradation which begins right at the time of death, hampering the reproducibility of sample collection (34). In contrast, CSF gives the opportunity to investigate the pathologies in neurological diseases on the living (35). According to Hühmer *et al*. CSF as the target of analysis enables a highly innocuous and direct approach to evaluate the CNS' cellular and chemical environment. An exchange of small molecules between extracellular fluid and CSF has been observed recently, implying the function of CSF as a medium to return molecules from the CNS to blood plasma (23).

Cerebrospinal fluid is the only body fluid in direct contact with the brain's extracellular space and contains many molecules produced in the ventricles. Thus, it can play an important role in the search of potential biomarkers in neurological diseases such as Alzheimer's disease, SCZ und MDD (2, 35).

2.1.2 Analysis of CSF

Cerebrospinal fluid is standardly collected by lumbar puncture (36). The patient can either be sitting or be lying in lateral position (37). In this procedure, the needle for liquor collection is standardly inserted in the spinal canal between the lumbar vertebrae L3/L4, L4/L5, as shown in **Figure 2**. The patient is asked to form a round back by bending the neck and allowing the chin to come close to the patient's chest to compensate the natural lordosis of the spine. After disinfecting the skin area, the needle is inserted under sterile conditions in cranial direction with an angle of approximately 70 degrees to the skin until it reaches the subarachnoid space. Generally, a volume up to 15 ml CSF can be collected utilising this strategy. Note that the first three drops of CSF are preferably to be discarded (29).

Importantly, after CSF collection, the sample should be checked visually for any turbidity, precipitation or changes in colour which can occur for example due to contamination with blood (29). Cerebrospinal fluid is a colourless, crystal clear liquid (38).

There are different methods and indications for the analysis of CSF in laboratory, the detailed information pertaining to the sampling method is as described in the catalogue of the German association of CSF diagnosis and clinical neurochemistry (38). A basic CSF analysis for diagnostic purposes includes the determination of glucose and lactate contents, number of cells, proteins and cytological differentiation (29). Techniques for the detection of antibodies like Enzyme-Linked Immunosorbent Assay (ELISA) can be applied for clinical purposes (29). Polymerase chain reaction (PCR) has become a helpful method in the diagnosis of viral meningitis (39).

Generally, for the detection of inflammatory processes in the CNS the comparison of the albumin, IgG, IgA or IgM quotient in CSF/Serum are frequently used (28). Furthermore, an increased total amount of proteins in CSF was found in multiple sclerosis and Guillain Barré syndrome (39). Protein levels can be reduced for instance due to a chronic leak (39). This implies a diagnostic value for alterations in protein quantification (32).

Even though the analysis of CSF has been of great use regarding the detection of acute neurological conditions (37) none of the markers proposed for MDD and SCZ in previous studies could be integrated in clinical use yet. Nevertheless, the successful identification of biomarkers for Alzheimer's disease due to proteome changes in CSF proves the body fluid's diagnostical value (7, 40).

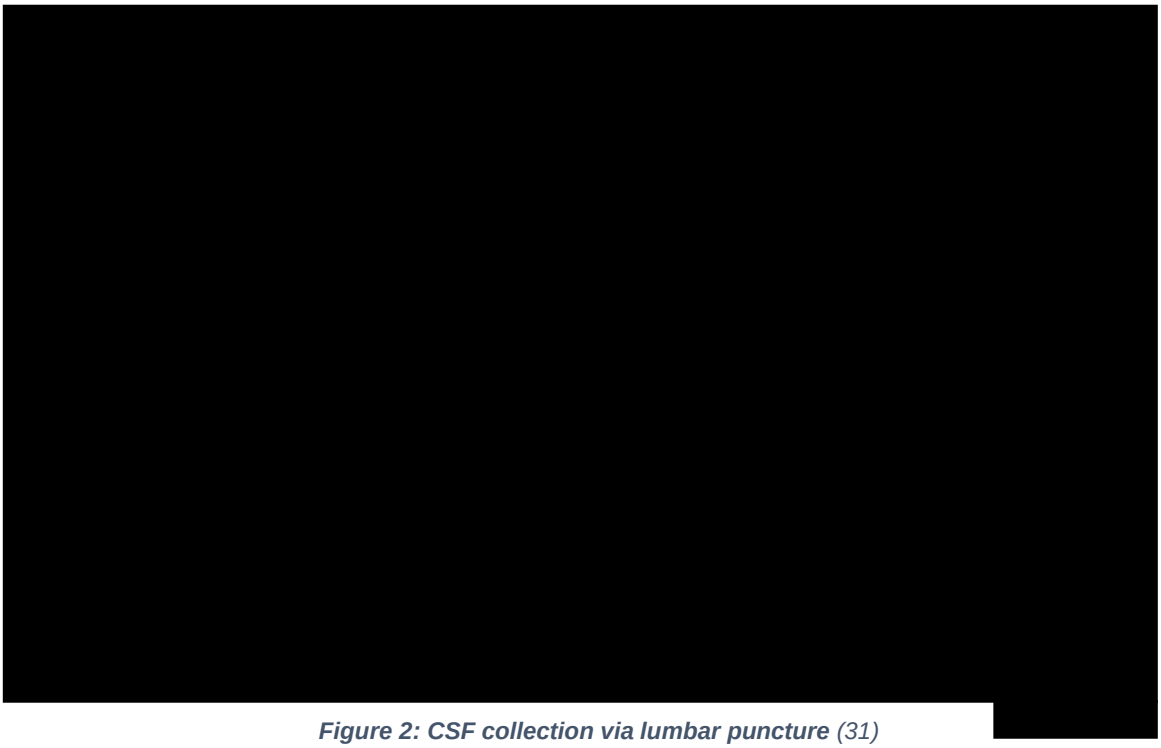

Figure 2: CSF collection via lumbar puncture *(31)*

2.2 Cerebrospinal Fluid: Proteomics

In 1995 the term “proteomics” was firstly introduced as a large-scale analysis of the entire proteins of a cell line, tissue or organism. Generally, there are two main goals in the field of proteomics: firstly, to gain a global view at protein levels and to identify all the proteins in a cell. Secondly, to understand the proteins' subcellular localization (6). The National Research Council Steering committee describes proteomics as “the effort to establish the identities, quantities, structures, and biochemical and cellular functions of all proteins in an organism, organ, or organelle, and how these properties vary in space, time, or physiological state“ (41).

The term “proteome” describes the total set of all proteins expressed in a tissue, cell or organism (42). The proteome determines the phenotype and reflects molecular modifications influenced by the past and present environment. Unlike the static genome the proteome is a dynamic system (32, 42, 43).
Most of the human’s biological functions are regulated by proteins which are dynamic and complex in their modulation (4). Proteins can undergo various forms of modifications. They can be synthesized or degraded, adjusting to the current state of the organism. Likewise, post-translational modifications (PTMs) and translocations within the cell play an important role. These PTMs can be a reaction to different extra- or intracellular signals. The proteome of a cell reflects the environment of the timepoint when the sample was taken (6). These variations in the proteome allow the investigation of alterations in the quantitative and qualitative expression in pathological states, evaluating progression in disease and finding potential biomarkers (9, 32). In proteomic studies this is usually done by comparing protein expression levels in disease state to a group of healthy control (34).

In the past, CSF has not been extensively studied in comparison to blood. This is mainly due to the differing accessibility of the two body fluids. The sample collection of CSF *via* lumbar puncture is a more invasive method compared to blood slimming. Moreover, the usage of CSF in research is limited to brain associated diseases (36). In addition, the total amount of proteins in CSF ranges between 0.2-0.7 mg/ml, which is a lot less than in blood plasma (60-70 mg/ml) (44, 45). Furthermore, the complexity of CSF makes it a difficult object of study (36, 46). The high abundance of albumin (most abundant protein in CSF) and immunoglobulins make the analysis of CSF a challenging task regarding the identification of low-abundant proteins (19, 36). The abundance of albumin and less abundant CSF proteins vary within eight orders of magnitudes, between dg/L and ng/L range. Due to the high dynamic range of proteins, there can be high abundant proteins like albumin masking low abundant proteins of interest (36, 47). For a long time, the progression of the proteomics field has been slow due to the limitation of the technology compared to those frequently used in the field of genomics. This has changed

recently. Due to the developments in mass-spectrometry, a new era of proteomics has been introduced. The great advantage of MS-based proteomics is that it enables the identification of the primary protein structure and allows the analysis of protein modifications at posttranscriptional and posttranslational level. It also contributes immensely to the research in the field of protein interactions (42). As discussed in chapter 2.1.1, CSF is a great target in the scientific investigation of biomarkers in the field of brain associated diseases. Due to modern techniques great improvement regarding total numbers of identified proteins in CSF could be achieved, starting with only 21 detected proteins in 2000, reaching 3379 identified proteins in 2018 (36).

Proteomics has become an important tool in the research of cancer, heart and neurological diseases and the analysis of CSF is an established approach for the detection of diseases like multiple sclerosis these days (9, 48).

2.2.1 Advantages of Proteomics Compared to Genomics

In the past, studies of the pathophysiology in SCZ disease have mainly focused on the genetic influence, studying family and twin relations. Given a heritability up to 80% for SCZ, scientists were looking for rare genetic variations. Even though it is known that genetics are a part of SCZ's pathophysiology, the cause of neurological diseases cannot be reduced to the genetic level only. The disease is of high complexity with multiple factors contributing (32).

Unlike proteomics, techniques used in genomics do not allow the investigation of post-translational modifications or alterations in the position or stability of the proteins. Previous genomic studies have shown that the abundance of proteins cannot be deduced by the sole analysis of messenger RNA (3, 49). Since the proteome is very dynamic, protein modifications need to be taken into account (32). These modulations have an impact on the protein functions and in the greater picture also on the cellular performance. Previous studies indicate that the same proteins or proteins with equivalent functions alter in expression in different neurodegenerative diseases. Thus, common inciting responses or causes in these diseases might be conceivable. The great advantage of proteomic technology is that it can be employed to analyse changes in protein modifications, protein abundance or protein interactions. These are not detectable using any other biological or genomic approach (3). It investigates the outcome of the gene expression cascade which is corresponding to the biological function (42). Furthermore, a great advantage of the proteomic method compared to genomics is that it allows a more detailed look at specific questions by analysing only fractions of the proteome, i.e. by using approaches like affinity methods (34). Proteomics can thus be a promising tool to evaluate the protein expression and to understand the pathophysiology underlying neurodegenerative diseases (3).

2.3 Proteomics Workflow

It is notable that various approaches exist in the field of proteomics using different analysing methods. However, they all compose the main aspects protein isolation, separation/fractionation, identification and quantification (50). In the field of proteomics, the samples are nearly always protein mixtures comprising of 100 to 10 000 proteins (42). The analysis of these protein mixtures is done in two steps. Firstly, the proteins need to be separated by one-dimensional (1DE) or two-dimensional (2DE) gel electrophoresis. The second step is the identification and quantification of the previously separated proteins. There are various techniques to identify proteins of which mass-spectrometry is the most commonly used method these days (9, 51). **Figure 3** provides a basic overview on the different work steps in the proteomic analysis of CSF.

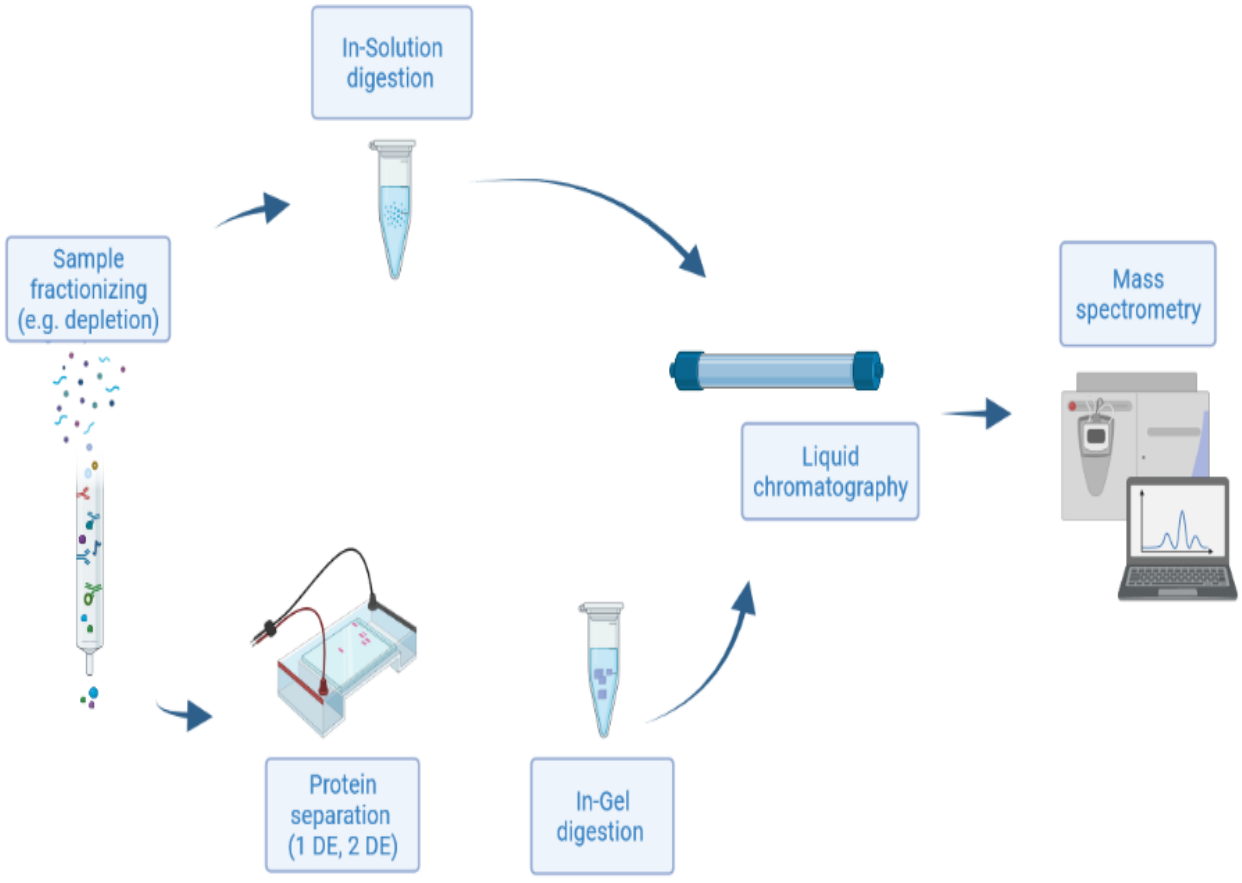

Figure 3: Simplistic proteomics workflow

2.3.1 Sample Preparation

2.3.1.1 Sample Fractionizing

The analysis of CSF is complicated by its high dynamic range. The protein concentrations in CSF are estimated to span in the range of 8-10 orders of magnitude, making the analysis of CSF challenging in regards of methodology (23, 36). In CSF a high concentration of proteins originating from blood can be detected. Albumin and immunoglobulins represent approximately 65% of the total protein content in CSF (23, 52, 53). The most abundant protein is albumin (130-350 mg/l) (23). These high concentration levels hamper the identification of low abundant proteins of interest. A prefractionation of CFS samples can hence be reasonable when looking for low abundant proteins as putative biomarkers in neurological diseases (23).

Depletion vs. Non-Depletion CSF

Affinity depletion methods aim to remove high abundant proteins before protein separation to enable a closer analysis of the proteins in low concentration levels (23). These techniques work with molecules such as antibodies with a high affinity and specify for the high abundant proteins. While the targeted proteins are captured by those molecules in a column, the rest of the sample including low abundant proteins can run through (44). However, the consideration of protein-protein interaction is inevitable. By removing high abundant proteins like albumin from the complex protein interplay, some proteins at lower concentration levels may become detectable. Yet, at the same time other low abundant proteins of interest that are bound to albumin are removed from the sample together with albumin (45). Due to this co-depletion of non-targeted proteins, there is an ongoing debate on the actual benefits of depletion (44).

Furthermore, most of the depletion kits available are designed for serum. These kits remove high abundant proteins like albumin which can not only be found in blood but also in CSF. Nevertheless, other proteins like prostaglandin D2 synthase or cystantin C which are synthesized in CNS and only exist in high concentrations in CSF (not in blood) are not targeted by these methods. A depletion kit especially developed for CSF might result in better output (52).

In 2019, Jankovska *et al*. demonstrated two different fractionizing methods (affinity depletion: MARS 14 and relative protein enrichment with a peptide ligand library: ProteoMiner) along with the analysis of crude CSF. Their results showed an increase in total protein detection with both depletion approaches. Nevertheless, they also noted a great variety between the identified proteins with the different strategies. A specific cluster of proteins could not be detected with one or the other fractionizing methods (including proteins which had been claimed before to be brain enriched) indicating a potential loss of biomarkers. Furthermore, they have investigated the potential protein loss by analysing the waste fraction after depletion. Jankovska *et al.* found many low abundant proteins to be co-depleted. It becomes clear that

the depletion procedure intervenes in the complex molecular interactions of proteins since many proteins found in the waste fraction could not be found in the depleted CSF. Interestingly, most of these co-depleted proteins were also detected in crude CSF. Jankovska *et al.* suggest to analyse not only the depleted CSF but also the waste fraction in future studies to prevent protein losses (44).

In summary, depletion is a feasible strategy to reduce the complexity of a sample but is an expensive procedure that holds the risk of unwanted co-depletion of proteins of interest.

2.3.1.2 Protein Separation

In order to analyse the CSF proteome, it is recommended to reduce the complexity of the proteins based on specific criteria, namely, molecular weight and charge. The most common techniques used are one-dimensional (1DE), two-dimensional (2DE) gel electrophoresis and various chromatography (LC) techniques (6, 50).

One-dimensional polyacrylamide electrophoresis:

1DE gel electrophoresis is a well-established method to separate protein mixtures according to their molecular mass (6). An electric field is applied to a polyacrylamide gel. The migration of charged proteins through the gel pores towards the positively charged anode depends on their size, shape and charge as well as on the gel pore size. The higher the concentration of acrylamide in the gel, the smaller the pores. While compact proteins run through the gel easily, the migration of bigger or extended molecules is slower (54). For the SDS-PAGE method proteins react with sodium dodecyl sulfate (SDS) prior to electrophoretic separation. SDS binds to the proteins according to their size (about 1.4g SDS per 1g protein) thereby charging them negatively and denaturing the proteins. The resulting negatively charged molecules move towards the positive pol and are hence separated according to their size and charge (55). The strengths of this technique are its reproducibility and the simplicity of its performance (6).

Two-dimensional polyacrylamide electrophoresis:

This technique was firstly introduced in 1972 by O'Farrell and Klose and allows the separation of complex protein mixtures including thousands of proteins in two dimensions by applying an electric field (50). The proteins (including post-translational modifications and protein isoforms) are separated in two manners. In the first dimension, the protein mixture is separated by isoelectric focusing, meaning the separation according to the protein's isoelectric point. Proteins can act as an acid or base depending on the surrounding pH. There is a specific pH value at which all charges of the amino acids add up to zero, the isoelectric point. By using a pH gradient gel, proteins can be separated according to their net charge since they will stop migrating in the solution when the net charge is zero at the protein's isoelectric point. In the

second dimension, the proteins which have been separated by isoelectric focusing migrate into the SDS-PAGE gel which is used for the separation by electrophoresis according to the protein's relative molecular mass. The proteins are then made visible as numerous "spots" by various staining approaches (34, 50). Softwares like Melanie and PDQuest 2-D are helpful tools in the comparison of gel pictures as well as in the identification and the evaluation of protein expression levels. The great advantages of 2DE gel are its ability to detect posttranslational modifications as well as variations of proteins that exist due to alternative mRNA splicing (6). A drawback in this technique may occur due to proteins whose isoelectric point is outside the regular range or which are not well resolved in the buffer (34). Additionally, 2DE is quite time consuming since one gel per sample is required (6). Gel electrophoretic separation of proteins is sensitive to salt. Since the salt concentration in CSF is comparably high, a prior desalting (>80%) is advisable (45).

2.3.1.3 Liquid Chromatography

After the separation of the protein mixture, proteins of interest are further cleaved into peptide fractions by in-gel or in-solution trypsin digestion. For the enzyme digestion a protease is employed, namely trypsin due to its cleavage specificity (34). It breaks up the bonds at arginine or lysine remnants (56).

A further fractionation is necessary to further reduce the complexity of all the peptides utilizing the liquid chromatography (LC) system. In the LC method the sample is eluted with a liquid solvent through a column that contains a stationary phase consisting of C18 material (57). The different components of the liquid phase interact individually with the stationary phase. Some components are absorbed stronger than others leading to differences in the flow rates and the separation of the sample (58). The separation of proteins employing LC is based on their physical attributes, namely ion exchange, affinity, reverse phase as well as exclusion by size. The LC column can be directly interfaced to a mass spectrometer (50).

2.3.2 Mass Spectrometry

The field of proteomics has been revolutionized due to great progress in mass spectrometry technologies (42). Its ability to detect sequences of amino acids and to determine PTMs allows the identification of thousands of proteins and their modifications from a complex protein mixture (59, 60). In a first step, the molecules are ionized, then secondly separated according to the mass to charge ratio (*m/z*) and afterwards detected in the third step. Identification is made possible with tandem mass spectrometry (MS/MS) by breaking precursor peptides that were selected during MS process into smaller peptide fractions (50). The most frequently used

techniques for ionization in the field of proteomics are matrix assisted laser desorption ionization (MALDI) or electrospray ionization (ESI) (50, 61).

In MALDI proteins are co-crystallized with matrix material on a metal plate. Laser pulses of a set wavelength can desorb and ionize proteins which pass in gas phase directly to the mass analyser (34). In ESI the proteins are dissolved in a solution. A charged needle spreads the liquid in form of small droplets. When the solvent evaporates the droplets get smaller and smaller until some ions finally leave the droplet and fly directly into the mass analyser (50).

There are multiple mass analyser types which divide the ions according to *m/z*, namely linear quadrupole ion trap (LTQ), orbitrap and time-of-flight (TOF) (59). The hybrid mass spectrometer Thermo Fisher's LTQ-Orbitrap-XL™ has proven to be one of the well-established systems for characterizing proteins largely due to its sensitivity, high resolution and precise mass determination. An Orbitrap system consists of a central spindle-shaped electrode and an outer electrode. For protein detection in Orbitrap the ions are electrostatically trapped due to the balance between attraction to the electrode and centrifugal forces. They are rotating around the central electrode, shifting back and forth. While the amplitude of the axial, harmonic oscillation does not differ between the ions, the frequency is specific for a certain *m/z* value. The frequency can thus be enhanced enabling the determination through an image current (62).

In the mass spectrum the relative intensity of the protein is presented on the *y*-axis and the *m/z* on the *x*-axis (34). After measuring the deviations in mass between the fragments the “decoded protein sequence” can be search against a database for the identification of the protein (50, 63).

Figure 4: Thermo Fisher Scientific LTQ Orbitrap XL schematic representation
The LTQ-Orbitrap-XL™ holds four components – a linear ion trap, a storage device (curved linear trap), an Orbitrap™ analyser and a collision cell. (64)

2.3.2.1 Data Analysis

One of the biggest challenges in proteomics is the enormous set of data that is generated and which needs to be comprehensively analysed utilizing various bioinformatics tools (65). There are several techniques for the determination of protein expression levels of which label-free quantification stands out as a simple, inexpensive and unlimited method regarding the number of samples (59, 66). The usage of databases is the conventional way of dealing with the structural information gained from MS in order to identify the proteins of a sample. A wide range of programs is available these days that can process uninterpreted MS/MS data working with search algorithms, e.g. Mascot or SEQUEST (6).

2.4 Targeted Proteomics for Clinical Study

While discovery proteomics aim to determine all proteins of a complex sample, the targeted approach focuses on a selection of proteins of interest. With this approach scientists and members of the health system can specifically analyse proteins that are important to their research questions and hypotheses (67). In contrast to the discovery approach a prior definition of the molecules in question is necessary. Thus, data for the proteins in focus is vital for the analysis (68). When investigating disease associated proteins sample pooling is not always the best option because the heterogenous of the sample and the individuality of proteins are compensated in the pool (69).

The field of targeted proteomics is constantly developing, and scientists have recently come up with multiple protocols (70). However, the principle working steps in targeted proteomics do not differ much from the workflow described above. Proteins need to be fractioned into peptides prior to analysis *via* mass spectrometry (67).

With the “selected reaction monitoring” (SRM) assays (the most common technique) the mass spectrometer can be programmed to solely detect preselected molecules enabling the exclusive analysis of targeted proteins. By specifically profiling fragments that were found in low numbers in discovery proteomic studies, the analysis of low abundant proteins becomes possible unimpaired by the high dynamic range of the sample (67). Due to the great development regarding software and instrumentation in the field of targeted data analysis by SRM a sensitive and robust protein quantification becomes possible (71). An advantage of the targeted approach is the reduced time of measurements enabling the analysis of a bigger patient cohort. Moreover, it is an adequate method to validate potential biomarker identified employing a discovery approach (67).

This strategy is of high relevance regarding the analysis of body fluids in clinical use since it allows a direct and fast verification of a suspected diagnosis by profiling known biomarkers. Once a biomarker set for SCZ and MDD will be specified, targeted proteomics can be implemented in clinical routine facilitating the diagnosis of the mental diseases.

2.5 Critical View on Analysing CSF Employing Proteomics Strategy

As discussed in chapter 2.3.1.1 a prior protein fractionizing in form of affinity depletion can be reasonable when analysing CSF due to its high dynamic range and the masking effect of high abundant proteins. However, the process of depletion intervenes in the complex protein interplay. It holds the risk of co-depleting potential biomarker proteins that are of particular interest in the present study (44). Moreover, the employment of depletion techniques is always accompanied by changes in protein concentrations (7). In order to avoid the accidental loss of these precious proteins no depletion steps were performed in the present study. Nevertheless, a method to deal with the high abundant proteins such as albumin in CSF is inevitable. 1DE gel electrophoresis is a reliable technique to separate the proteins by size. In order to minimise the masking effect, it is advisable to cut the resulting gel in very small bands, particularly in the intensely stained regions of the gel where albumin is located. 1DE is a basic, well- established technique for protein separation and is in this case preferred over 2DE due to its great reproducibility and the reduced workload (6). Thereby minimizing the risk for technical errors.
Mass spectrometry has become a major tool in proteomics enabling the analysis of amino acid sequences as well as PTMs convincing with its speed, versatility, specificity, sensitivity and accuracy (59, 72). Thermo Fisher's LTQ-Orbitrap-XL™ is a MS system of high sensitivity and accuracy. However, a prior protein separation is necessary when working with Thermo Fisher's LTQ-Orbitrap-XL™ and results in better outcomes at peptide level.

2.6 Mental Disorders

The term “mental disorder” covers a wide range of diseases including depression, bipolar disorder (BPD), schizophrenia, dementia, developmental disorders and other psychoses (73). These diseases are expected to be related to 40% of the disabilities in developed countries (74, 75). Mental disorders have major consequences on social, human rights, health and the economy with a clear growing tendency (73). Our knowledge of the pathophysiology behind these diseases is limited due to the complex interactions of various factors such as psychological, social, environmental and etiological components (76).

Proteomic analysis of various body fluids has recently become an important tool in the search of potential biomarkers for mental diseases. While blood has been the body fluid of choice in many studies, CSF, a fluid in direct contact with the central nerve system, has not been equally well studied (77). In this study we are looking at proteome alterations in CSF of patients who suffer from schizophrenia or major depressive disorder in search for potential biomarkers. It is noteworthy that mental disorders are complex and the pathologies cannot be reduced to a gene mutation resulting in a single altered protein only. In neurological illnesses multiple proteins with various interactions as well as variations in gene products may play a role in the cause of the diseases (34).

2.6.1 Schizophrenia

Schizophrenia is a complex mental disorder that affects about 1% of the population worldwide (78). A certain family predisposition is known, e.g. children of schizophrenic parents have a higher risk of developing the disease (lifetime prevalence 12%), likewise siblings (lifetime prevalence 10%) (79). Affected patients show deviations in thoughts and behaviour that can be accompanied by different forms of hallucinations, delusional or paranoid ideations as well as movement and mood disorders (“positive symptoms”) (12, 30, 80). “Negative symptoms” may include reduced expression of emotions (“flat effect”), reduced feelings and speaking, social retreat and listlessness (79, 80). The causes of this impairing neuropsychiatric disease are heterogeneous and to date not fully understood. However, they are considered to be related to genetic and environmental factors (78, 81). While in the past many studies on SCZ have focused on genetic variations, it is clear now that the aetiology cannot be explained by genetics alone (32). Scientists hypothesize that changes of the dopamine system in the brain may lead to this disabling disorder (78). To date, the assessment of SCZ is solely based on clinical criteria (ICD-10 classification), making it a subjective diagnose that is complicated by the great variety of symptoms, the similarity to other neuropsychiatric diseases and the lack of disease specific biomarkers (12, 79, 82).

2.6.1.1 Potential Biomarkers in Human CSF for SCZ

Table 1 presents 12 studies that analysed the proteome of human CSF taken from schizophrenic patients between 2003 and 2020 (32). Overall, the numbers of proteins identified in the listed studies were relatively low, an exception is a study presented by Al Shweiki *et al.* with a total detection of 153 differently expressed proteins in SCZ vs. control (CTRL).

Different proteomic techniques were utilised to investigate the CSF proteome. Three studies employed 2DE for prior protein separation in combination with Matrix-Assisted Laser Desorption and Ionisation (chapter 2.3.2) coupled with a time-of-flight detector and a mass spectrometer (MALDI-TOF-MS). Jiang *et al.* found apolipoprotein A4 (APOA4) to be significantly downregulated in CSF of SCZ patients employing this technique (30). Wan *et al.* published a paper in 2006 presenting 80 identified proteins out of which they found transthyretin (TTR) and apolipoprotein E (APOE) to be significantly altered in SCZ CSF, suggesting these two proteins as potential biomarkers for diagnosis and treatment in future (69). MALDI-TOF-MS analysis performed by Martins-de-Souza *et al.* highlighted apolipoproteins E and A1 (APOA1) together with prostaglandin-H2 D-isomerase (PTGDS) as a possible set of markers in CSF (83).

Four other studies listed in **Table 1** also employed TOF-MS in combination with a different ionization method, namely Surface-Enhanced Laser Desorption and Ionisation (SELDI). SELDI is a further development of MALDI, in which a chip is integrated that allows the reduction of sample complexity due to its affinity chromatographic surface (84). This method has been frequently employed analysing human CSF by Huang *et al.* In the first approach 2006 they found neurosecretory protein VGF (VGF)-derived peptide to be upregulated and TTR to be downregulated and were able to distinguish SCZ from controls with a specificity of 95% and a sensitivity above 80% (10). The result of elevated VGF-derived peptide and downregulated TTR was verified by the study group one year later (85). Moreover, in 2008 Huang *et al*. employed SELDI-TOF-MS detecting decreased levels of APOA1 not only in CSF but also in liver, serum, red blood cells and post-mortem brain tissue proving an alteration in the CNS as well as in peripheral tissue (78). Albertini *et al.* reported a clear reduction of Aβ1–42 peptide levels (fractions of amyloid-precursor-protein (APP)) in SCZ patients as well as a slight increase of sAPPα, likewise utilising SELDI-TOF-MS (86).

Johansson *et al.* analysed the CSF of 17 twin pairs, aiming to investigate how family and environment influence CSF markers. Therefore, they employed different immunoassays. In the course of the study, they found monocyte differentiation antigen CD14 (CD14) to be elevated in individuals with SCZ or BPD in comparison to their healthy co-twins (87). A potential biomarker to differentiate between SCZ, control and Parkinson's disease was discovered by Gupta *et al.* in 2019. The group attached chemical tags (iTRAQ) to the peptides prior to QTOF-MS analysis. For validation purposes Enzyme-linked Immunosorbent Assay (ELISA) was

employed. Alpha-2 Macroglobulin (A2M) was found to be downregulated in SCZ patients compared to control, while elevated levels of A2M were observed in Parkinson's disease (88). Al Shweiki *et al*. likewise performed iTRAQ-based MS analysis of CSF, however, in combination with liquid chromatography. With this approach the group identified a total number of 1795 proteins analysing samples from three cohorts (SCZ, MDD and CTRL) of which 153 were differently abundant in SCZ compared to CTRL.

These studies, all presenting different potentially involved proteins in the pathophysiology of SCZ, prove again the complexity of the mental diseases and stress the need of research in this field in future to come closer to an understanding of SCZ and its biological causes and manifestations.

2.6.2 Major Depressive Disorder

Depression is a mental disorder that affects more than 264 million people globally and can be found in all age groups with a lifetime prevalence of 16–20% (89, 90). It presents one of the major global causes of disability resulting in a great contribution to the disease burden worldwide (90). The total costs of adults suffering from MDD in the US have increased from US$ 236.6 billion in 2010 to US$ 326.2 billion in 2018 (91). The leading symptoms are a depressed mood, loss of interest and listlessness. Additionally, sleep disturbances, reduced self-confidence, feelings of guilt, loss of appetite and concentration difficulties may occur. Self-injury and thoughts of suicide are possible (89).

To date, 20% of patients do not show respondents to any available treatment (92, 93). Due to factors like phenotype heterogeneity patient's response to drug treatment with antidepressant and numbers of remission differ (94). The causes of MDD are still not completely understood. There are several hypotheses regarding the background of this disease. One common theory is the deficiency of serotonin and norepinephrine that was concluded due to the correspondence to medication, namely tricyclic antidepressants (TCAs) and selective serotonin reuptake inhibitors (SSRIs), that cause an increase of serotonin and norepinephrine in the neuronal synapse. However, this theory conflicts with newer antidepressants (92, 95). At the present time, MDD is perceived as a disease of multifactorial origin with genetic, neurobiological, social-psychiatric and further environmental factors contributing (89). The diagnosis of MDD is (like SCZ) mainly based on the evaluation of clinical symptoms, making it a subjective and error-prone diagnose, emphasising the need for reliable biomarkers (92).

2.6.2.1 Potential Biomarkers in Human CSF for MDD

Table 2 presents 12 studies that analysed CSF from humans diagnosed with MDD from 2001 to 2020 (92). Differences in methodology were not as distinctive as in previous SCZ studies (chapter 2.6.1.1). Eight studies employed various immunoassays that function by means of a specific reaction between antigen and antibody, most frequently ELISA Kits were utilized, e.g. Diniz *et al.* presented reduced levels of brain-derived neurotrophic factor in late life depression employing ELISA technique (96). Higher Interleukin-6 (IL-6) and Interleukin-8 (IL-8) levels were observed by Kern *et al*. in depressed patients aged between 70-84 years (97). Elevated levels of neurofilament light protein (NEFL), CSF/serum albumin ratio and Aβ42 were monitored by Gudmundsson *et al.* (98). However, deviating results of decreases Aβ42 levels were described by Pomara *et al*. employing the same method, namely ELISA (99). ELISA is a simple, highly efficient procedure that is based on a reaction between antigen and antibody. However, it holds several disadvantages, e.g. the procedure is expensive and labour-intensive. Moreover, the instability of antibodies holds certain requirements regarding transport and storage and a relatively high risk of false results occurs due to immobilized antigens that haven't been sufficiently blocked (100).

Grabe *et al.* analysed specifically the abundance of protein S100beta (S100beta) in depressive patients utilizing an immunoluminometric assay. Protein S100beta is believed to be involved in the regeneration of serotonergic synapses and showed significantly higher expression levels in disease (101). Yoon *et al*. investigated the levels of cocaine- and amphetamine-regulated transcript (CARTPT) peptide in CSF of depressed patients. They suggest CARTPT to play a role in the development of depression due to reduced CSF levels in MDD cohort and a negative correlation with dose of antidepressant (102).

In 2013, Maccarrone *et al*. performed an interesting study using protein biosignatures rather than a single protein biomarker to differentiate between CSF of depressed, bipolar or schizophrenic patients and healthy individuals. They managed to distinguish MDD from control with classification rates of 100% and 97% using a panel of 24 proteins that had been classified as "significant biomarkers". Although the protein variations within the three psychiatric patient groups were not as significant as each compared to healthy control, they named chromogranin A (CHGA) as a potential protein to distinguish between MDD and SCZ patients (103). Besides, other studies discussed tau protein and pTau as biomarkers to differentiate between MDD and Alzheimer's disease or mild cognitive impairment (104, 105).

Solely three previous studies on MDD utilized MS analysis. Huang *et al.* described altered VGF23-62 peptide and secretogranin II (SCG2) 529-566 levels employing SELDI-TOF-MS (10). Ditzen *et al.* compared the CSF of twelve individuals with depressive symptoms and healthy controls using 2DE and MALDI-TOF-MS analysis and found 39 proteins to be differently abundant including elevated levels of cystatin C (CST3), prostaglandin-D2 synthase,

APOE and neurofilament heavy polypeptide (NEFH) as well as decreased levels of secretogranin I (CHGB) and TTR (94).

An even higher number of differently abundant proteins in MDD was presented in 2020 by Al Shweiki *et al.* employing iTRAQ and LC-MS. In the comparison MDD vs. CTRL they determined 161 significantly differently expressed proteins which all appeared to be downregulated. Subsequently, twelve of the identified markers were validated *via* targeted multiple reaction monitoring analysis including VGF, leucine-rich repeat and immunoglobulin-like domain-containing nogo receptor-interacting protein 1 (LINGO1), contactin-associated protein-like 4 (CNTNAP4), neurexin-3-beta (NRXN3), neurexophilin-1 (NXPH1), proSAAS (PCSK1N), neuroserpin (SERPINI1), CARTPT, glutamate receptor 4 (GRIA4), neuronal pentraxin receptor (NPTXR) and cerebellin-4 (CBLN4) (16).

Mental disorders are heterogeneous and there is a great variety in expression. Yet, at the same time there are shared symptoms between the diseases (76, 92). The complexity of these neurological diseases is not likely to be expressed in the different abundance of a single protein, looking for biomarker panels might thus be a more promising approach (92).

Table 1: Altered proteins in SCZ discussed in literature

Author	Year	Patients	Method	Altered proteins	Observations	Reference
Jiang et al.	2003	10 SCZ, 10 CTRL	2DE, MALDI-MS	↓APOA4		(30)
Huang et al.	2006	58 SCZ, 16 MDD, 5 OCD, 10 AD,90 CTRL	SELDI-TOF-MS	↑VGF32-62 peptide, ↓TTR	Differentiation between SCZ and CTRL, 80% sensitivity, 95% specifity	(10)
Wan et al.	2006	35 SCZ, 36 CTRL	2DE, MALDI-MS	↓APOE, ↓TTR tetramer, ↓TRFE, ↓RET4, ↓Ig Kapa, ↓Ig Gama, ↓HPT, ↓APOA1, ↓A1AG2, ↑TTR monomer, ↑TETN, ↑APOJ, ↑A1AT, ↑ALBU		(69)
Huang et al.	2007	54 first-episode SCZ, 24 IPS, 70 CTRL	H-NMR, SELDI-TOF-MS	↑VGF32-62 peptide, ↓TTR	Only significant difference between SCZ and CTRL, not between IPS and CRTL or IPS and SCZ	(85)
Huang et al.	2007	10 SCZ, 10 CTRL	label-free nano-LC MS		Differentiation between SCZ and CTRL in PLS-DA scores	(7)
Huang et al.	2008	41 SCZ, 40 CTRL	SELDI-TOF-MS	↓APOA1	↓APOA1 also in peripheral tissue (liver, RBC, serum, brain tissue)	(78)
Martins-De-Souza et al.	2010	17 first-onset SCZ, 10 CTRL	2DE, MALDI-TOF, TOF	↑APOE, ↑PTGDS, ↑APOA1, ↓TTR, ↓TGFR1, ↓CCDC3		(83)
Albertini et al.	2012	11 SCZ, 20 AD, 20 CTRL	SELDI-TOF-MS	sAPPα↑; Aβ1-42↓		(86)
Johansson et al.	2017	17 twin pairs, 1 twin sibling. 2 pairs concordant SCZ, 11 pairs discordant SCZ/SAD/BPD, 4 pairs CTRL	Immunoassays, ELISA	↑CD14	Several neurodegenerative markers with high heritability. Influence of dominant genetic variation on AβX-42, Aβ142, P-tau and CSF/serum albumin ratio	(87)
Gupta et al.	2019	36 PD, 7 SCZ, 15 CTRL	iTRAQ, LC-QTOF-MS, ELISA	↓A2M	↑A2M in Parkinson's disease	(88)

Gupta et al.	2019	36 PD, 10 SCZ (6 treated, 4 drug-naive), 15 CTRL	ELISA	Tendency: ↓APOE in drug-naïve ↑APOE in treated	APOE and α-synuclein concentrations both inversely correlate with dopamine concentrations	(106)
Al Shweiki et al.	2020	Discovery: 12 MDD, 6 BPD, 6 SCZ, 14 CTRL Targeted: 40 MDD, 11 BPD, 13 SCZ, 27 CTRL	iTRAQ, LC-MS	SCZ vs. CTRL: ↓47 proteins ↑106 proteins	Validation of 12 differently expressed proteins *via* targeted approach	(16)

Table 2: Altered proteins in MDD discussed in literature

Author	Year	Patients	Method	Altered proteins	Observations	Reference
Grabe et. al.	2001	11 mild/moderate D, 11 CTRL	Immunoluminometric assay	↑S100β		(101)
Brunner et al.	2005	14 MDD	2DE		characteristic protein for suicide attempter 33kDa and pI5.2, no identification of protein	(107)
Huang et al.	2006	58 SCZ, 16 MDD, 5 OCD, 10 AD, 90 CTRL	SELDI-TOF-MS	↑VGF23-62 peptide, ↓SCG2 529-566		(10)
Gudmundsson et al.	2008	78 MDD	ELISA	↑NFL-L, ↑Aβ42	↑ CSF/serum albumin ratio	(98)
Ditzen et al.	2011	12 MDD, 12 CTRL	2DE, MALDI-TOF-MS	↓B2GPI, ↓DBP, ↓TTR, ↓PTGD2S, ↓/↑A1BG, ↓HNRNPH 1, ↓/↑SERPINF 1, ↓HSPA8, ↓APOE, ↑HPX, ↓/↑CYSC, ↓ALDOC, ↓PRKACA, ↓ORMI, ↓A2M, ↑APP, ↑AGT, ↓APOA1, ↓APOA2, ↓APOD, ↑B2M, ↓CD44, ↓CLU, ↓COL6A1, ↓C3, ↓NPC2, ↓GSN, ↓CNDP1, ↑HRG, ↑HOXD12, ↑NEFH, ↑NRCAM, ↓PEX5, ↓SERPING1, ↓PLA2G7, ↓F2, ↓CHGB, ↓SOD1, ↓VTN		(94)
Pomara et al.	2012	28 MDD, 19 CTRL	ELISA	↓Aβ42, (↓Aβ40),		(99)
Maccarrone et al.	2013	36 MDD, 27 BPD, 35 SCZ, 35 CTRL	Protein Microarray		Differentiation CTRL from MDD, SCZ, BPD using panel of proteins	(103)
Diniz et al.	2014	25 LLD, 25 CTRL	ELISA	↓BDNF		(96)

Deuschle et al.	2014	10 AD, 10 MDD; confirmation: 17 AD, 8 MDD	ELISA-RIA		hypocretin-1 concentrations not differ between AD and MDD, hypocretin concentrations in CSF are related to Tau and phosphorylated Tau (pTau) in CSF	(108)
Kern et al.	2014	19 minor or MDD, 67 CTRL	ELISA	↑ IL-6, ↑ IL-8		(109)
Yoon et al.	2018	24 MDD, 25 CTRL	Immunoassay Kit	↓ CARTPT peptide	Negative correlation of CARTPT levels with antidepressant dose, psychomotor retardation, somatic anxiety, and general somatic symptoms. Positive correlation with obsessive and compulsive symptoms.	(102)
Al Shweiki et al.	2020	Discovery: 12 MDD, 6 BI, 6 SCZ, 14 CTRL Targeted: 40 MDD, 11 BPD, 13 SCZ, 27 CTRL	iTRAQ, LC-MS, MRM, Immunoblotting	MDD vs. CTRL: ↓ 161 proteins	Validation of 12 differently expressed proteins *via* targeted approach	(16)

3 Materials and Methods

3.1 Equipment, Consumables and Chemicals

Equipment

Intelli Mixer	NeoLab, Heidelberg, Germany
Ultrasonic bath	Bandelin Sonorex RK31, Berlin, Germany
Biofuge primo R	Heraeus, Fisher Scientific GmbH, Schwerte
Eppendorf pipettes	Eppendorf, Hamburg, Germany
Eppendorf 0,5, 1,5, 2,0 ml containers	Eppendorf, Hamburg, Germany
Bio-one, pp-tubes, 15 ml	Greiner, Frickenhausen, Germany
Covering foil PCR, self-adhesive, PP	Ratiolab GmbH, Dreieich, Germany
96 well cell culture cluster, Costar	Merck KGaA, Darmstadt, Germany
Well cell culture cluster, v-bottom	Greiner bio-one, Frickenhausen, Germany
SOLAμ™ HRP SPE well plate	Thermo Fisher Scientific Inc., Waltham, USA
EASY-nLC 1200 system	Thermo Scientific, Rockford, USA
Acclaim PepMap RSLC, nano column	Thermo Scientific, Rockford, USA
LTQ Orbitrap XL	Thermo Fisher Scientific Inc., Waltham, USA

Chemicals and Consumables

HPLC water	AppliChem GmbH, Darmstadt, Germany
Acetonitril, LC-Grade	AppliChem GmbH, Darmstadt, Germany
TFA for Proteinsequence analysis	Merck KGaA, Darmstadt, Germany
Pierce BCA-Protein Assay Kit	Thermo Fisher Scientific Inc., Waltham, USA
Ammonium bicarbonate	Carl Roth GmbH+CoKG, Karlsruhe, Germany
Acetic acid	Carl Roth GmbH+CoKG, Karlsruhe, Germany
Trifluoroacetic acid (TFA)	Merck KGaA, Darmstadt, Germany
DTT 1,4-Dithiotreit	Carl Roth GmbH+CoKG, Karlsruhe, Germany
IAA Iodoacetamide	Merck KGaA, Darmstadt, Germany

<u>Protein digestion:</u>

Trypsin	Promega, Madison, WI 53711, USA
HPLC-water	AppliChem GmbH, Darmstadt, Germany
Acetonitril, LC-Grade	AppliChem GmbH, Darmstadt, Germany
Ammonium bicarbonate	Carl Roth GmbH+CoKG, Karlsruhe, Germany

Gel electrophoresis:

NuPage 4-12 % Bis-Tris Gel	Novex, life technologies, Carlsbad, CA, USA
NuPage Reducing Agent (10x)	Novex, life technologies, Carlsbad, CA, USA
Seeblue Plus2 Prestained Standard	Invitrogen AG, Carlsbad, USA
NuPage MOPS-buffer	Novex life technologies, Carlsbad, CA, USA
Antioxidant	Invitrogen AG, Carlsbad, USA
Novex ® Colloidal Blue Staining Kit	Invitrogen, Carlsbad, CA, USA
NuPage LDS Sample Buffer (4x)	Life Technologies, Carlsbad, CA, USA
Methanol MS-Grade	Fisher Scientific, Waltham, USA

Zip-tipping / Sola plate:

C18 ZipTips®	Millipore, Merck KGaA, Darmstadt, Germany
Acetonitrile (ACN), LC-Grade	Merck KGaA, Darmstadt, Germany
Trifluoroacetic acid (TFA)	Merck KGaA, Darmstadt, Germany
HPLC-water	AppliChem GmbH, Darmstadt, Germany

LC/Orbitrap – sample preparation:

Acetonitrile (ACN), LC-MS grade	AppliChem GmbH, Darmstadt, Germany
Water, LC-MS grade	AppliChem GmbH, Darmstadt, Germany
Methanol, LC-MS	AppliChem GmbH, Darmstadt, Germany
Formic acid, LC-MS grade	Fisher Scientific, Waltham, USA

Software

MaxQuant v. 1.6.5.0 / 1.6.17.0	Max Planck Institute of Biochemistry (Cox and Mann 2008), Martinsried, Germany
Proteome Discoverer v. 1.1.0.263	Thermo Fisher Scientific Inc., Waltham, USA
Perseus Version 1.6.5.5	Computational Systems Biochemistry, Max Planck Institute of Biochemistry, Martinsried, Germany

3.2 Discovery Study 1: CSF Analysis *via* 1DE, Capillary-LC-MS

In the present study, we analysed CSF samples taken from patients diagnosed with schizophrenia or major depressive disorder and compared the proteome of the two cohorts with a group of healthy controls aiming to detect protein alterations and to identify potential biomarker candidates. One-dimensional gel electrophoresis and capillary-LC-ESI-MS/MS approach was chosen for the label-free quantification of peptides.

3.2.1 Study Samples

CSF samples were provided by Dr. Michael van der Kooij as part of the collaboration with Prof. Dr. Marianne Müller, Department of Psychiatry and Psychotherapy, University Medical Center Mainz. Cerebrospinal fluid was extracted by lumbar puncture from the individuals that were categorised into three different groups. Each cohort comprised of 21 patients. Since the composition of the CSF proteome is age dependent (23), a wide range of age was chosen to gain an overview of protein variations independent of age. The clinical attributes of the patient's characteristics of the included patients are listed in **Table 3**.

The first group consisted of eleven female and ten male individuals diagnosed with SCZ with an average age of 37.76 ± 12.4 years. All patients were under medication during sample collection. Applied drugs were atypical antipsychotics, benzodiazepines, selective serotonin and noradrenalin reuptake inhibitors, typical antipsychotics and tricyclic antidepressant, as presented in **Table 4**.

The second cohort comprised of patients suffering from MDD. Eleven female and ten male individuals with an average age of 51.4 ± 16.1 years were included. Participants were treated with different drugs such as atypical antipsychotics, benzodiazepines, lithium, mirtazapine, mood stabilizer, selective serotonin and noradrenalin reuptake inhibitors, selective serotonin reuptake inhibitors, typical antipsychotics and tricyclic antidepressants, as shown in **Table 5**.

As control, CSF from a pool of 21 healthy patients with no known neurological disorder was analysed. The average age of participants was 42 ± 16.6 years. Fifteen female und six male individuals were included. Details can be obtained from **Table 6**. All samples were stored at -80 °C until usage.

Table 3: Overview of the included patients in all cohorts

	SCZ	MDD	CNTRL
Age	37.8 ± 12.4	51.4 ± 16.1	42 ± 16.6
Gender	11 female, 10 male	11 female, 10 male	15 female, 6 male
Total proteins (µg / 100 µl)	37.46 ± 9.38	35.54 ± 8.37	33.05 ± 8.35
Medication	Under medication	Under medication	Without medication

Table 4: Characteristics of patients included in SCZ cohort

Group	Age	Gender	Medication	Total proteins (µg / µl)	Total proteins (µg / 100 µl)	Standardized volume per sample = ~140 µl (~40 µg)	Replicates	Standardized volume per replicate = ~140 µl
SCZ	28	f	SSRI, AA	0.24	24.00	20.00	R1	140.00
SCZ	47	f	AA, TCA	0.27	27.40	20.00	R1	
SCZ	35	m	TA	0.31	30.60	20.00	R1	
SCZ	60	f	AA	0.34	34.00	20.00	R1	
SCZ	28	m	AA	0.41	41.20	20.00	R1	
SCZ	43	f	B, TCA	0.44	44.00	20.00	R1	
SCZ	31	m	AA	0.51	50.80	20.00	R1	
SCZ	37	f	AA	0.26	25.80	20.00	R2	140.00
SCZ	39	m	AA	0.28	28.20	20.00	R2	
SCZ	50	f	AA	0.33	33.00	20.00	R2	
SCZ	32	f	AA	0.35	34.70	20.00	R2	
SCZ	33	f	AA, B	0.42	41.70	20.00	R2	
SCZ	27	m	AA	0.45	45.00	20.00	R2	
SCZ	22	m	TA	0.53	53.00	20.00	R2	
SCZ	18	f	AA	0.27	27.00	20.00	R3	140.00
SCZ	43	f	AA	0.31	30.60	20.00	R3	
SCZ	18	m	AA	0.33	33.10	20.00	R3	
SCZ	57	f	TA	0.35	35.00	20.00	R3	
SCZ	55	m	AA, B	0.42	42.00	20.00	R3	
SCZ	35	m	SSRI, AA	0.49	48.80	20.00	R3	
SCZ	55	m	AA	0.57	56.80	20.00	R3	

Table 5: Characteristics of patients included in MDD cohort

Group	Age	Gender	Medication	Total proteins (µg / µl)	Total proteins (µg / 100 µl)	Standardized volume per sample = ~140 µl (~40 µg)	Replicates	Standardized volume per replicate = ~140 µl
MDD	75	f	SSRI, TA	0.26	26.20	20.00	R1	140.00
MDD	57	f	SNRI, MS, AA	0.29	28.90	20.00	R1	
MDD	67	m	Mi	0.31	31.00	20.00	R1	
MDD	24	m	SSRI	0.33	32.80	20.00	R1	
MDD	72	m	SSRI, AA	0.35	34.70	20.00	R1	
MDD	18	f	SSRI	0.38	38.00	20.00	R1	
MDD	29	m	SSRI, AA, B,TCA	0.42	41.60	20.00	R1	
MDD	60	f	TCA, B	0.26	26.30	20.00	R2	140.00
MDD	62	m	SSRI	0.29	29.30	20.00	R2	
MDD	38	m	SSRI	0.31	31.20	20.00	R2	
MDD	34	f	TCA, Mi	0.34	33.80	20.00	R2	
MDD	40	f	TCA	0.36	36.20	20.00	R2	
MDD	58	m	B	0.40	40.10	20.00	R2	
MDD	48	m	AA, SSRI	0.57	57.00	20.00	R2	
MDD	61	m	SSRI, TCA	0.28	27.60	20.00	R3	140.00
MDD	60	f	B	0.31	30.60	20.00	R3	
MDD	50	f	SSRI	0.32	31.70	20.00	R3	
MDD	42	f	TCA, MS, Li	0.34	34.00	20.00	R3	
MDD	65	f	Mi	0.37	36.60	20.00	R3	
MDD	74	m	SSRI, Mi	0.41	40.60	20.00	R3	
MDD	45	f	TA, Mi, B	0.58	58.10	20.00	R3	

Medication:
AA: atypical antipsychotic, B: benzodiazepine, Li: Lithium, MAOI: monoamine oxidase inhibitor, Mi: mirtazapine, MS: mood stabilizer, SNRI: selective serotonin and noradrenalin reuptake inhibitor, SSRI: selective serotonin reuptake inhibitor, TA: typical antipsychotic, TCA: tricyclic antidepressan

Table 6: Characteristics of patients included in the control group

Group	Age	Gender	Medication	Total proteins (µg / µl)	Total proteins (µg / 100 µl)	Standardized volume per sample = ~140 µl (~40 µg)	Replicates	Standardized volume per replicate = ~140 µl
CTRL	25	f	none	0.21	20.80	20.00	R1	140.00
CTRL	53	f	none	0.23	23.00	20.00		
CTRL	51	f	none	0.26	26.30	20.00		
CTRL	52	m	none	0.31	30.50	20.00		
CTRL	37	f	none	0.36	36.20	20.00		
CTRL	43	f	none	0.40	40.30	20.00		
CTRL	76	f	none	0.42	42.40	20.00		
CTRL	21	f	none	0.23	22.50	20.00	R2	140.00
CTRL	22	f	none	0.25	24.60	20.00		
CTRL	30	f	none	0.29	28.90	20.00		
CTRL	18	m	none	0.33	32.80	20.00		
CTRL	36	f	none	0.39	38.60	20.00		
CTRL	63	m	none	0.41	40.60	20.00		
CTRL	57	m	none	0.45	44.60	20.00		
CTRL	43	f	none	0.23	23.00	20.00	R3	140.00
CTRL	21	m	none	0.25	25.00	20.00		
CTRL	41	f	none	0.29	29.00	20.00		
CTRL	34	f	none	0.34	33.80	20.00		
CTRL	48	f	none	0.39	38.90	20.00		
CTRL	35	f	none	0.42	42.40	20.00		
CTRL	76	m	none	0.50	49.90	20.00		

3.2.2 Sample Preparation

Twenty µl CSF per sample of seven individuals from the same cohort were pooled to 140 µl within one biological replicate resulting in three biological replicates (R1, R2, R3) within the three designated groups of MDD, SCZ and CTRL, as shown in **Table 4**,**Table 5** and **Table 6**, respectively.

To enable the identification of proteins employing capillary-LC-MS, an amount of 40-50 µg of total protein per biological replicate is required. In order to reach an adequate amount of total proteins for the analysis a prior pooling of samples was inevitable. Cerebrospinal fluid from seven individuals (20 µl per sample) summed up to a total amount of 140 µl. Standardisation of the sample volume, rather than the amount of total proteins, is reasonable when analysing CSF due to the huge role albumin plays in the composition of CSF, it accounts for approximately 60% of total CSF proteins. The natural composition of CSF may be disturbed by standardisation of the total amount of proteins because smaller proteins that are interacting with albumin could be either masked or be higher abundant than they would normally appear due to their relation to albumin. Since this study is interested in the composition of CSF and its complex molecular structure in disease, the method refrains from adjustment of the amount of total proteins to avoid a disturbance of the molecular interplay and standardizes the volume instead.

3.2.3 First Dimensional Gel Electrophoresis

Since high abundant proteins can mask low abundant proteins of interest, the need for a fractionation method to reduce the complexity of the CSF proteins prior to the MS analysis occurs (110-112). One-dimensional gel electrophoresis is a fractionation method based on the separation according to variations in protein mass (113, 114).

CSF samples were prepared for 1DE as described by Perumal *et al.* (115, 116). Per replicate 135 µl pooled CSF were divided into two equal amounts of 67.5 µl each. This division resulted in two lanes with, in comparison to one bigger band in a single lane, smaller and more accurate bands. This is particularly sensible with regard to the thick band of albumin since the creation of smaller, more accurate albumin bands reduces the probability of smaller protein bands in close positional relation to albumin to be masked.

For 1DE 4-15% NuPAGE Bris-Tris gel and MOPS running buffer were employed. The voltage was set on 150 V for approximately 60 min at 4°C. As a marker for protein molecular mass SeeBlue Plus 2 was utilized. Colloidal Blue Staining Kit was employed according to the manufacturer’s instructions to stain the gels. After staining the gels overnight, the staining solution was replaced by deionized water.

3.2.4 In-Gel Tryptic Digestion & Extraction

As described in section 3.2.3, the separation of the pooled CSF into two lanes results in smaller, more precise protein bands in the gel. Especially, the complexity of the big albumin band can hence be reduced and masks less low abundant proteins. Moreover, a cutting scheme of 24 bands was employed, which divided the big band of albumin in three parts. Each band was then carefully sliced into small pieces with a size of approximately 1 mm x 1 mm. The two identical bands of the same replicate were combined in 1.5 mL microcentrifuge tubes. The gel pieces were destained by applying 500 µl of destaining solution (50 mM ammonium bicarbonate in 50% acetonitrile) for approximately 30 min and the solution was then carefully removed with a pipette. The following reduction of the disulfide bonds was achieved with 400 µl of 10 mM 1,4-dithiotreit (DTT). The tubes were kept for half an hour at 65 °C. The DTT solution was carefully removed before alkylation. Alkylation prevents the disulfide bonds from rearranging (117). For the alkylation step 400 µl 55 mM iodoacetamide solution was utilized. Following the dehydration with 500 µl acetonitrile, the proteins were digested for 16 hours at 37°C using 50 µl trypsin solution (0.013 µg/µl trypsin in 10 mM ammonium bicarbonate in 10 % acetonitrile).

In order to extract the remaining peptides from the gel pieces, the extraction buffer (150 µg of 1:2 (v/v) of 5 % formic acid:acetonitrile) was carefully added. Samples were incubated in a shaker for 30 min at 37°C. The supernatant containing the peptides of interest was collected, dried in a speed-vac-centrifuge and stored at -20 °C.

3.2.5 Peptides Purification

Peptides were purified utilizing ZipTip® C18 columns, special pipette tips that hold a C18 reversed-phase media and enable a desalting, purifying and concentrating of the samples (118). Samples were retrieved from the freezer and 10 µl of 0.1% TFA was added in each tube. Sonification for 5 min and centrifugation followed.

According to the manufacturer's instruction the employed ZipTip was cleaned using 100% ACN three times, then the running solution (0,1% TFA) was aspirated and discarded three times, before resuspending the sample in its tube ten times. After repeating the second step, the samples were eluted in a 0.5ml microcentrifuge tube using two times 10 µl of the elution solution (0.1% TFA in 60% ACN) and resuspended five times. This cycle was repeated four times. After completing the ZipTip process, the samples were dried in SpeedVac and then stored at -20°C.

3.2.6 Liquid Chromatography

The peptides were dissolved in 10 µl of 0.1% TFA for subsequent LC-MS-analysis. The LC system employed consisted of a Rheos Allegro pump and an HTS PAL autosampler provided with a BioBasic C18, 30 x 0.5 mm precolumn which was linked to a BioBasic Phenyl, 100 x 0.5 mm analytical column in combination with a Jupiter 4 µ Proteo analytical, 150 x 0.5 mm column. LC-MS grade water with 0.1% (v/v) formic acid was used as solvent A and LC-MS grade acetonitrile also with 0.1% (v/v) formic acid as solvent B. The total time for the run of the resulting gradient per sample summed up to 60 min. 0-5 min: 10% B, 5-45 min: 10-45% B, 45-50 min: 45-90% B, 50-55 min: 90% B, 55-60 min: 10% B. The LC system was directly coupled with ESI-LTQ-Orbitrap-XL-MS system (59).

3.2.7 Mass Spectrometry

The system utilized in this study to obtain continuum data in form of mass spectra was an ESI-LTQ-Orbitrap-XL-MS. The system was set on a spray voltage of 2.15 kV in positive-ion electrospray ionization mode, heated capillary temperature was determined at 220°C. The orbitrap was programmed in the “data-depending mode”, enabling the transition between Orbitrap MS and LTQ MS/MS. Mass spectra in the range of 300 to 1600 *m/z* were obtained with a resolution set to 30000 at *m/z* 400. “Target automatic gain control” was fixed at 1 x 10^6 ions (59). The option lock mass was selected in MS mode. For “internal recalibration polydimethlycyclosiloxane” ions at *m/z* 445.120025 were applied in real time (119). Using CID fragmentation, the five ions with the highest intensity were isolated and fragmentated in the LTQ operating with the following settings: “normalized collision energy” (NCE): 35%, “activation time”: 30 ms and “dynamic exclusion”: 60 s. The fragments were then processed in the LTQ (59).

3.2.8 Label-free Quantification Analysis

The resulting mass spectra were processes by MaxQuant software version 1.6.5.0. It uses the search engine Andromeda on the basis of intensity-based absolute quantification (iBAQ) algorithm to identify and quantify peptides with high accuracy. The data from MS was search against Uniprot Human database (date 05/04/2019). The following settings were defined for the search: “fragment mass tolerance” of ± 0.5 Da, “peptide mass tolerance” of ± 20 ppm. A “false discovery rate” (FDR) of 0.01 for “peptide and protein identification with ≥ 6 amino acid residues” was allowed. Only “unique plus razor peptides” were accepted for the identification (59, 120). As a fixed modification “Carbamidomethylation of cysteine” was determined, as

variable modifications “protein N-terminal acetylation and oxidation of methionine”. The enzyme trypsin was selected and the “maximum number of missed cleavages” was set at 2. “Match between the runs” was enables with match time window of 5 min. The resulting table “proteingroups.txt” was “filtered for contaminants” (59).

The generated data was further processed in Perseus software version 1.6.5.5 which enables a statistical analysis of enormous proteomic data sets. Firstly, $\log_2$ transformation of normalized LFQ protein intensities was performed to generate a standard distribution with ascertainable values. Working with Perseus software “reverse hits and minimum number of values in at least one group” is 3 and the normal distribution was used to replace the missing values. For the statistical assessment, a “two-samples t-test” with p-values < 0.05 was employed to determine significantly differently abundant proteins. Unsupervised hierarchical clustering analysis was performed according to Euclidean distance (linkage = average; preprocess with k-means) to illustrate the heat map that presents the differentially expressed proteins based on the z-scored values (59).

3.2.9 Functional Annotation and Pathway analysis

The identified proteins and gene names were transferred to IPA software for subsequent functional annotation and pathway analysis. The functional annotation and pathway analysis enables the network analysis by determination of molecular functions and protein-protein interactions as well as the localization of the protein within the cell and the identification of upstream regulators (121). Top diseases and canonical pathways of those proteins that appeared to be differently expressed in this study were listed. Their p-values were calculated employing the Benjamini-Hochberg (B-H) multiple testing correction (122).

An excel list of all significantly differently expressed proteins was processed in IPA program. The protein-protein interaction networks were presented in different shapes representing the functional classes of proteins and are organized by their location within the cell. The colours and the colour intensities represent the protein abundance (red: upregulation, green: downregulation). The connecting lines between the proteins show known direct and indirect protein interactions.

3.3 Discovery Study 2: CSF Analysis *via* In-Solution, Nano-LC-MS

In a second proteomic discovery study, CSF of 13 patients diagnosed with schizophrenia was analysed employing a nano-LC column which enables a measurement of very high accuracy of individual samples. No sample pooling and no prior protein separation *via* gel is necessary employing this method due to the small diameter of the nano column. The small size of the column results in a much higher pressure within the column that allows an outstandingly sensitive protein separation. Overlapping signals can thus be avoided and proteins can be detected at much lower concentration rates.

3.3.1 Study Samples

In this pilot study six female and seven male individuals diagnosed with SCZ according to ICD-10 classification were included with an average age of 42.54 ± 11.38 years. All patients were under medication, details can be obtained from **Table 7**.

Table 7: Characteristics of the included patients

Samples	Age	Gender	Medication	Total proteins (µg / µl)
SCZ 1	32	f	AA	0.35
SCZ 2	33	f	AA, B	0.42
SCZ 3	43	f	B, TCA	0.44
SCZ 4	50	f	AA	0.33
SCZ 5	57	f	TA	0.35
SCZ 6	60	f	AA	0.34
SCZ 7	28	m	AA	0.41
SCZ 8	35	m	TA	0.31
SCZ 9	35	m	SSRI, AA	0.49
SCZ 10	39	m	AA	0.28
SCZ 11	55	m	AA, B	0.42
SCZ 12	55	m	AA	0.57
SCZ 13	31	m	AA	0.51

3.3.2 In-Solution Tryptic Digestion

For in-solution trypsin digestion, 2.5 µl of 0.1 µg/µl trypsin solution in 10 mM ammonium bicarbonate in 10 % acetonitrile was carefully added to the reaction tube holding 5 µl CSF sample each. The individual CSF samples were tryptic digested overnight (minimum 16 hours) at 37 °C. Then, samples were dried in SpeedVac and stored at - 20°C prior to further analysis.

3.3.3 SOLAµ HRP SPE Well Plates

The tryptic digested peptides were purified and concentrated employing SOLAµ™ HRP solid phase extraction well plates. This method comprises of five steps: Firstly, the wells were conditioned/activated with 100 % acetonitrile and then equilibrated using 0.1 % formic acid. In a third step, the wells were loaded with the samples. After washing the wells with 0.1 % formic acid twice, the peptides were eluted, now concentrated and freed from contamination. Details regarding the procedure can be obtained from **Table 8**.

Table 8: Procedure employing SOLAµ™ HRP SPE well plates

	Step	**A/E**	**Volume (µl)**	**Speed (rpm)**	**Duration (min)**
1	Activation	100% ACN	100	3000	1
2	Equilibration	0.1% FA	100	3000	1
3	Sample load 1st	Sample in 0.1 %FA	100	3000	1
3	Sample load 2nd	Sample in 0.1 %FA	100	3000	1
4	Wash 1st	0,1% FA	100	3000	1
4	Wash 2nd	0,1% FA	100	3000	1
5	Elution 1st	50% ACN	75	3000	1
5	Elution 2nd	50% ACN	75	3000	1

3.3.4 Nano-Liquid Chromatography

For nano-LC-MS-analysis, 40 µl of 0.1 % formic acid was carefully pipetted in the reaction tubes to dissolve the peptides. The employed nano-LC system (EASY-nLC 1200 system) was equipped with an Acclaim PepMap RSLC, 75µm x 50 cm, nanoViper analytical column. Two solvents were utilized, referred to as solvent A and B. LC-MS grade water with 0.1 % (v/v) formic acid was used as solvent A and 80% LC-MS grade acetonitrile with 0.1 % (v/v) formic acid as solvent B. The total time for the run of the resulting gradient per sample summed up to 150 min. 0-120 min: 5 % - 30 % B, 120-130 min: 30-100 % B, 130-150 min: 100 % B. A direct connection between the nano-LC system and the employed ESI-LTQ-Orbitrap-XL-MS system was established.

3.3.5 Mass Spectrometry

As described in section 3.2.7 with the following changes: Mass spectra in the range of 300 to 2000 *m/z* were obtained with a resolution set to 60000 at *m/z* 400.

3.3.6 Lable-free Quantification Analysis

As described in section 3.2.8 with the following changes: The resulting mass spectra were processes by MaxQuant software version 1.6.17.0 and a two-samples t-test with p values < 0.05 was employed to determine significantly differently abundant proteins.

4 Results

4.1 Biomarker Discovery Employing 1DE and Capillary-LC-MS

In the first part of the present study the proteome of pooled CSF samples of patients diagnosed with MDD, SCZ and a group of healthy individuals were analysed *via* 1DE and capillary-LC-ESI-MS. Since no prior sample fractionizing *via* depletion was performed, protein separation *via* 1DE was employed and every lane of the resulting gel was sliced into 24 define bands to reduce the high dynamic range of CSF. Each biological replicate (135 µl) was divided into two fractions and the wells were loaded with 67.5 µl sample each to achieve a good protein separation and to minimize the masking effect of the highly abundant proteins, especially serum albumin.

The intensely stained region where albumin was located (approximately 66 kDa, band 10-13) was carefully excised into three tiny bands to reduce the masking effect at its best (**Figure 5**). Furthermore, the two slices of the same replicate were pooled again after excising to ensure a significant amount of proteins yielded to facilitate higher protein detection in MS analysis. **Figure 5** shows the employed cutting scheme in a representative gel for all replicate groups.

In total, 277 (FDR < 1 %) proteins were successfully identified employing the optimized discovery proteomics strategy. Generally, the total number of proteins identified in CTRL, MDD and SCZ were 244 ± 9, 227 ± 12 and 199 ± 11, respectively, as shown in **Figure 6**. Notably, lower numbers of proteins were identified in the SCZ group compared to the CTRL and MDD groups.

A total of 239 proteins were identified in all the three designated groups, as shown in **Figure 7**. An overlap of two proteins was determined exclusively between SCZ and MDD group, four between SCZ and CTRL and twelve between MDD and CTRL. The proteins identified solely for one group amount to two in SCZ, three in MDD and 15 in CTRL, as presented in the venn diagram **Figure 7**. Overall, the optimized discovery proteomics strategy resulted in good reproducibility outcomes between the designated groups.

Functional annotation analysis was performed to investigate biological processes and functions associated with the discovered proteins. Employing IPA software, the log p-values of the three different groups were used to visualize the grade of involvement. Firstly, the affected biological pathways and diseases for all with this discovery approach revealed proteins were compared in **Figure 8**. The top five involved canonical pathways are as followed: acute phase response signaling, LXR/RXR activation, FXR/RXR activation, complement system and coagulation system and the top five associated diseases: amyloidosis, degranulation of blood platelets, tauopathy, dementia and Alzheimer's disease. As mentioned

before, the detected proteins are almost the same for the three groups with 239 of 277 proteins overlapping. Hence, the top functions/diseases and pathways do not differ between the groups and only slight differences in the degree of involvement can be observed.

In order to determine significantly differently expressed proteins, the extracted MaxQuant data was statistically analysed using Perseus software. A Student's T-test was performed comparing the protein intensities of the 277 proteins to explore the number of significantly differently abundant proteins in disease and healthy individuals. Among the identified proteins, 79 were found to be significantly differently abundant between the designated groups of which four proteins, namely collagen alpha-1(VI) chain (COL6A1), SPARC, cathepsin D (CTSD) and complement factor H-related protein 1 (CFHR1), were overlapping in all three groups (**Figure 9**). **Figure 9** presents the pairwise differences in the $\log_2$ differences of the proteome of the comparison analysis and the total number of proteins that were significantly differentially expressed are as follows: MDD vs. CTRL (25), SCZ vs. CTRL (66) and MDD vs. SCZ (22). Among the 79 significantly differently abundant proteins four proteins were shared by all three cohorts. While 13 proteins were overlapping in MDD vs. CTRL and SCZ vs. CRTL as well as in SCZ vs. MDD and SCZ vs. CTRL there were no proteins identified that were altered solely in MDD vs. CTRL and SCZ vs. MDD, as depicted in **Figure 9**. Most changes in protein levels were observed in the SCZ group compared to CTRL; 65 proteins were found to be downregulated. Interestingly, only one protein, namely apolipoprotein A2 (APOA2), appeared to be upregulated in SCZ patients, as depicted in **Figure 10**.

The ten proteins which showed the most significant differences in expression levels were protein TANC2 (p=3.87E-03), neuronal pentraxin-1 (NPTX1) (p=1.2E-03), multiple epidermal growth factor-like domains protein 8 (MEGF8) (p=6.26E-03), complement component C7 (C7) (p=1.98E-02), COL6A1 (p=2.76E-03), inter-alpha-trypsin inhibitor heavy chain H2 (ITIH2) (p=1.35E-02), limbic system-associated membrane protein (LSAMP) (p=1.91E-03), voltage-dependent calcium channel subunit alpha-2/delta-1 (CACNA2D1) (p=9.81E-03), brevican core protein (BCAN) (p=2.58E-02) and contactin-2 (CNTN2) (p=7.55E-03). Notably, all these proteins found to be differently abundant in SCZ vs. CTRL showed lower expression levels in disease. In the MDD cohort the most significantly downregulated proteins were neurexin-3 (NRXN3) (p=1.45E-02), MEGF8 (p=4.18E-02), NPTX1 (p=2.70E-03), seizure 6-like protein (SEZ6L) (p=7.26E-03) and seizure protein 6 homolog (SEZ6) (p=2.58E-02). Further details regarding the intensities can be obtained from **Table 9**.

In summary, this innovative and optimized proteomic strategy allowed the detection of a high number of CSF proteins without performing prior depletion steps. The little variation in total numbers of identified proteins within one cohort proves the workflow to be accurate and reliable. The analysis of the obtained data made it possible to highlight the significantly

differently abundant proteins in disease. The identified proteins were found to be involved in various processes, especially in neurological pathways, which is reasonable since SCZ and MDD are severe mental diseases. Therefore, the employed method provides an avenue for enabling biomarker identification in the designated groups.

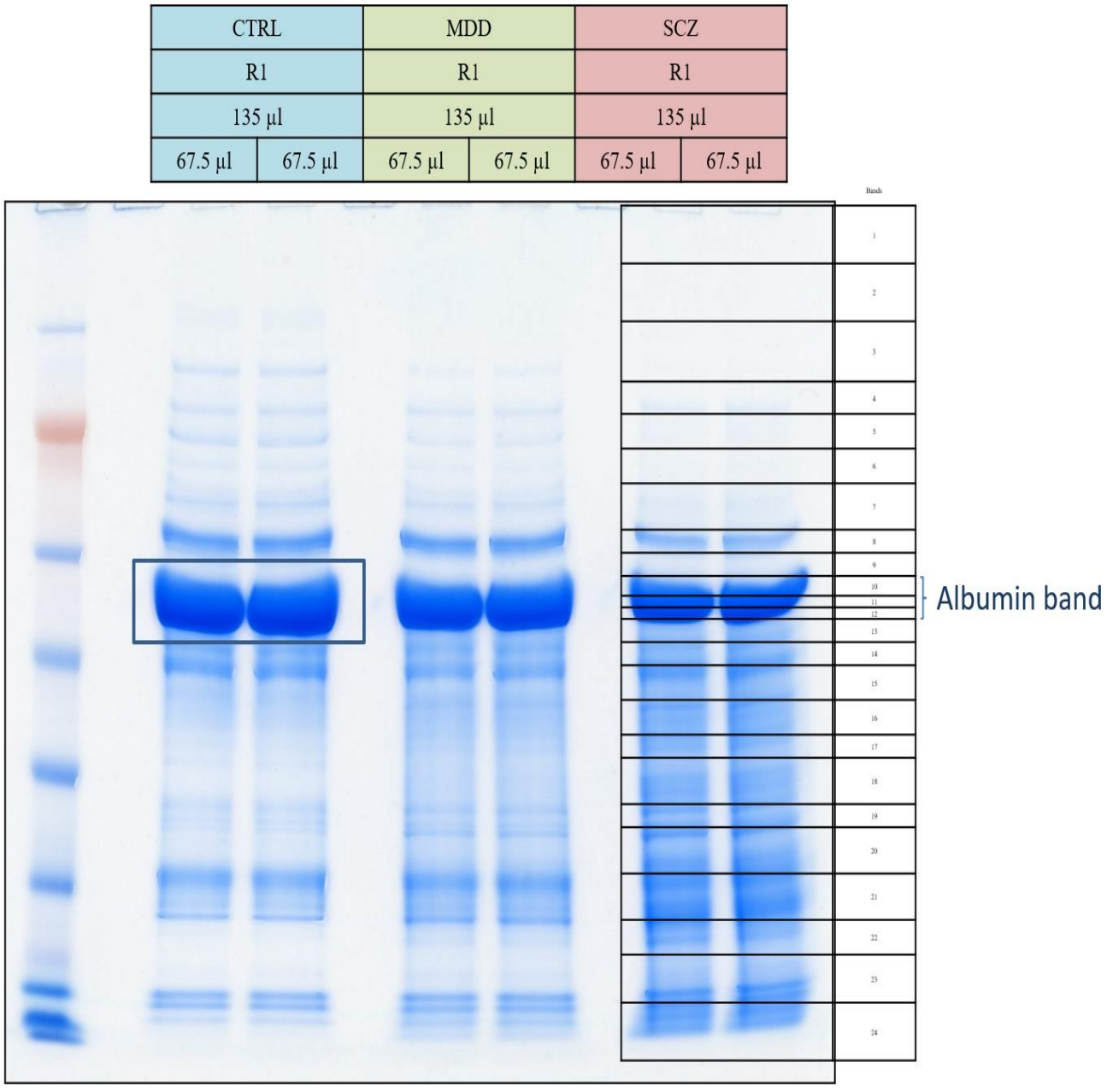

Figure 5: Cutting scheme for 1DE gel

This gel picture shows biological replicate R1 of the three patient cohorts (CTRL, MDD, SCZ). The first row presents the reference marker for the molecular masses. Each biological replicate (135 µl) was divided into two fractions and the wells were loaded with 67.5 µl sample each. On the right side the cutting lines are presented, slicing the gel into 24 bands.

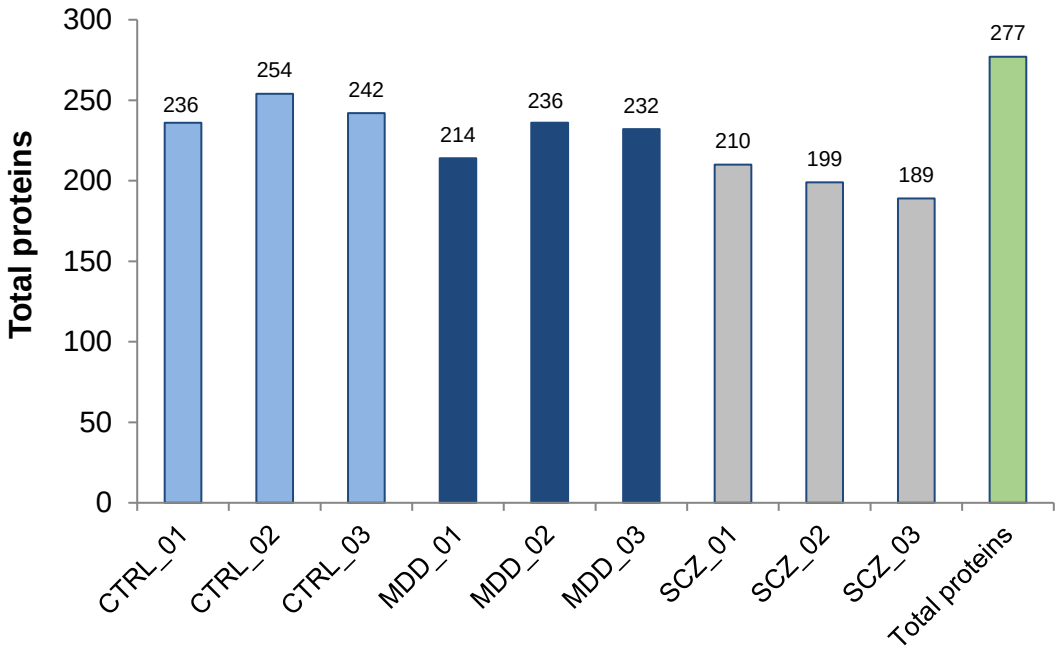

Figure 6: Total number of CSF proteins identified in each biological replicate (FDR < 1 %)
This figure presents the number of identified proteins in the biological replicates R1, R2 and R3 for each patient cohort (CTRL, MDD; SCZ). Identified proteins are overlapping between and within the groups. Numbers of detected proteins in the biological replicates of the same cohort are comparable. However, there is a tendency of less determined proteins in the SCZ group compared to CTRL and MDD. The very right column shows the number of total detected proteins in the present study.

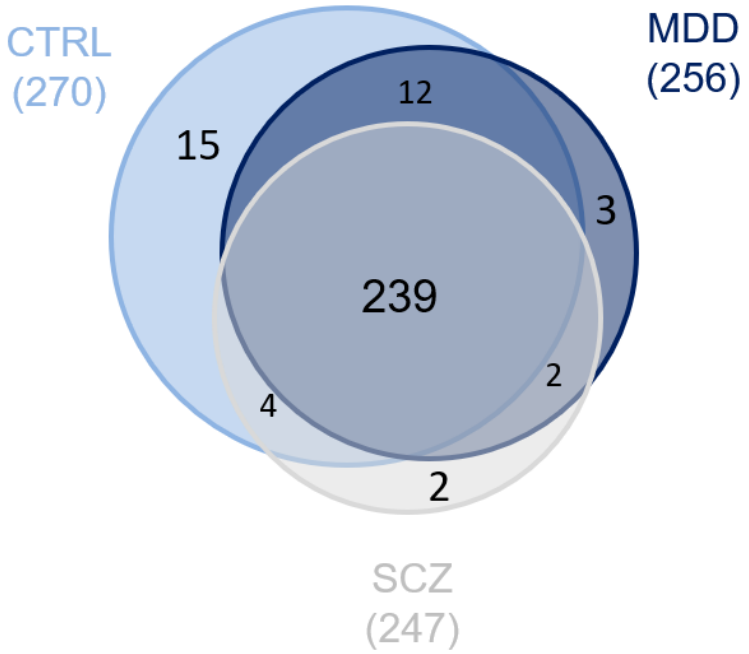

Figure 7: Numbers of detected proteins
This venn diagram presents all identified proteins in the present study in the groups CTRL, MDD and SCZ. A total number of 277 protein was determined with 239 overlapping in all three cohorts.

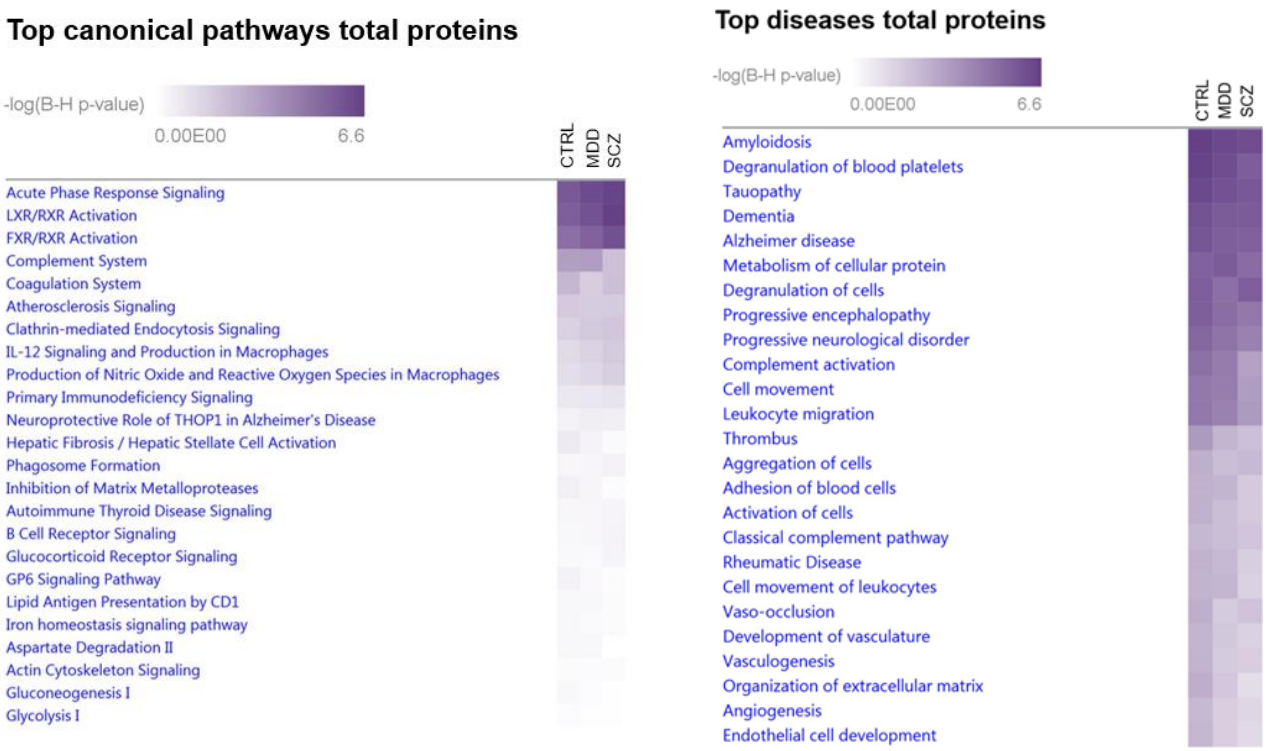

Figure 8: Top canonical pathways and diseases of all identified proteins

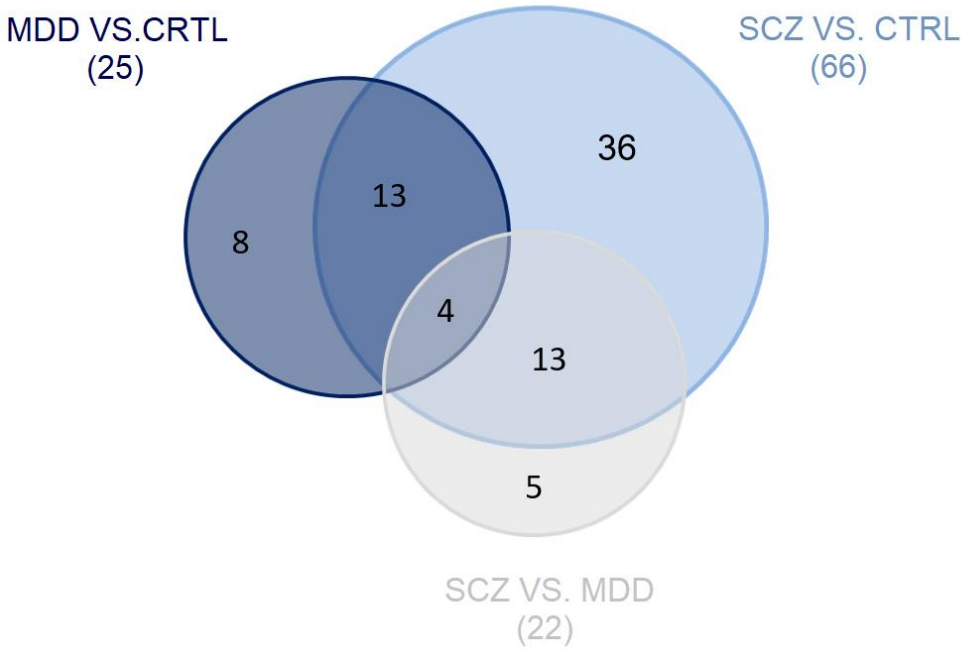

Figure 9: Numbers of differently expressed proteins

This venn diagram presents the numbers of proteins that showed differences in abundance when comparing the patient cohorts MDD and SCZ with healthy individuals and comparing the proteins levels between the two diseases. Most altered proteins were observed in SCZ group.

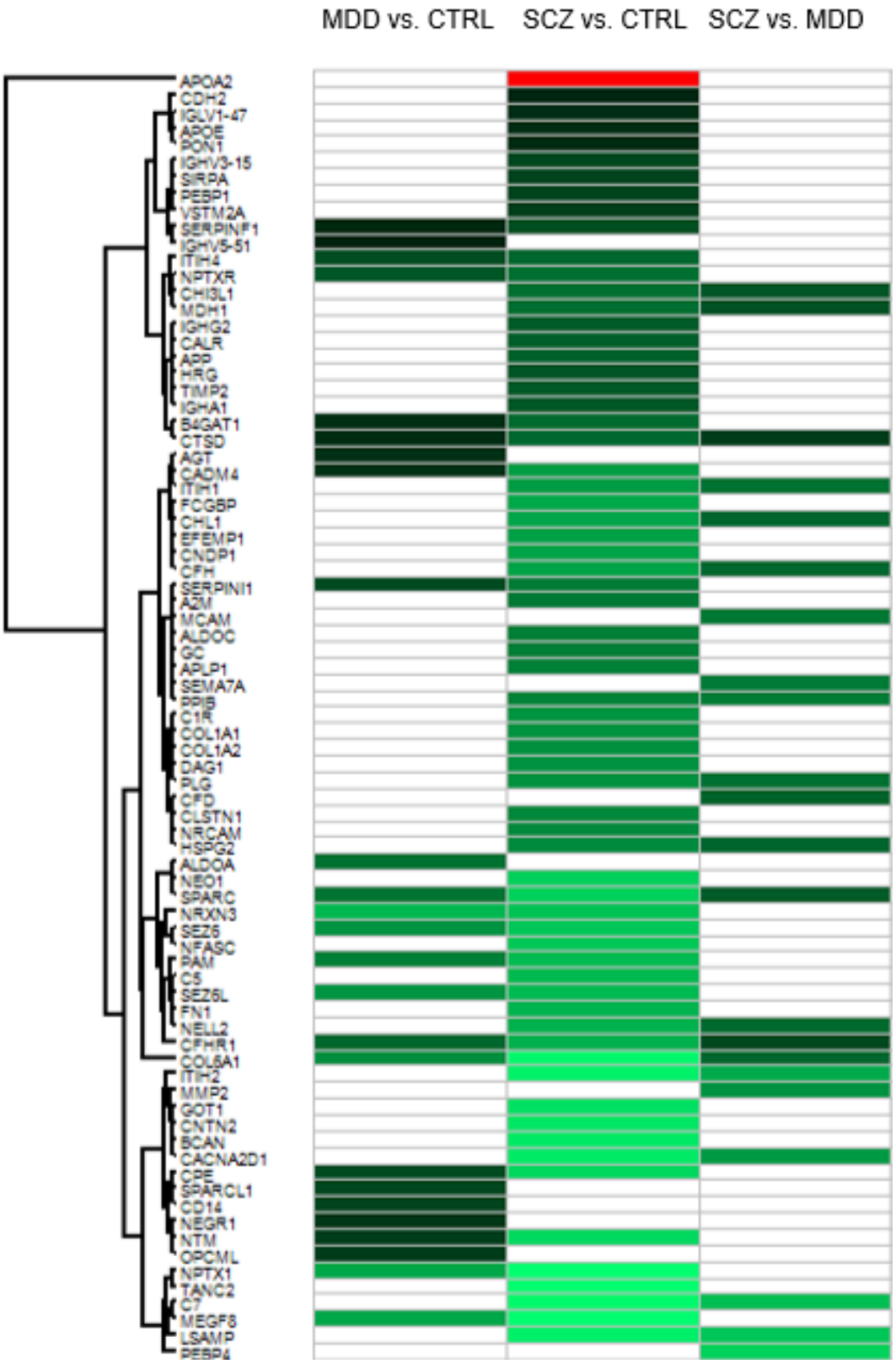

Figure 10: Significantly differently expressed proteins

The heat map depicts the hierarchical clustering of all the 79 differentially abundant proteins based on the log_2 ratio difference related to the designated comparison analysis. The upregulated proteins are shown in red and the downregulated proteins are presented in green.

Table 9: Differently abundant proteins detected in the present study via 1DE and LC-ESI-MS/MS

Protein IDs	Protein names	Gene names	MDD vs. CTRL		SCZ vs. CTRL		SCZ vs. MDD	
			p-value	Log_2 difference	p-value	Log_2 difference	p-value	Log_2 difference
Q96S96	Phosphatidylethanolamine-binding protein 4	PEBP4					6.97E-04	-4.87
Q13449	Limbic system-associated membrane protein	LSAMP			1.91E-03	-5.54	1.84E-03	-4.58
P40925	Malate dehydrogenase, cytoplasmic	MDH1			1.05E-02	-2.58	2.40E-03	-1.90
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2			4.02E-03	-3.20	4.93E-03	-2.33
P23284	Peptidyl-prolyl cis-trans isomerase B	PPIB			7.26E-03	-3.03	8.54E-03	-2.90
P08603	Complement factor H	CFH			7.35E-03	-3.79	9.02E-03	-2.41
O00533	Neural cell adhesion molecule L1-like protein	CHL1			6.77E-03	-3.84	1.20E-02	-2.34
P19827	Inter-alpha-trypsin inhibitor heavy chain H1	ITIH1			3.12E-02	-3.64	1.23E-02	-2.68
P12109	Collagen alpha-1(VI) chain	COL6A1	2.55E-02	-3.31	2.76E-03	-5.72	1.28E-02	-2.41
P00747	Plasminogen	PLG			2.55E-02	-3.39	1.78E-02	-2.61
P43121	Cell surface glycoprotein MUC18	MCAM					1.84E-02	-2.80
P09486	SPARC	SPARC	1.17E-02	-2.71	1.42E-03	-4.84	1.93E-02	-2.13
P54289	Voltage-dependent calcium channel subunit alpha-2/delta-1	CACNA2D1			9.81E-03	-5.41	2.15E-02	-3.58
P07339	Cathepsin D	CTSD	5.15E-03	-1.02	2.63E-03	-2.42	2.24E-02	-1.40
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	ITIH2			1.35E-02	-5.57	2.42E-02	-3.95
P36222	Chitinase-3-like protein 1	CHI3L1			1.20E-02	-2.60	2.68E-02	-2.01
P08253	72 kDa type IV collagenase	MMP2					2.79E-02	-3.37
Q99435	Protein kinase C-binding protein NELL2	NELL2			1.49E-02	-4.12	3.82E-02	-2.46
P10643	Complement component C7	C7			1.98E-02	-5.74	3.86E-02	-4.39
P00746	Complement factor D	CFD					4.28E-02	-2.29
O75326	Semaphorin-7A	SEMA7A					4.28E-02	-2.85
Q03591	Complement factor H-related protein 1	CFHR1	4.36E-02	-2.42	2.56E-03	-4.11	4.62E-02	-1.69

Table 9: Continued

Protein IDs	Protein names	Gene names	MDD vs. CTRL		SCZ vs. CTRL		SCZ vs. MDD	
			p-value	Log_2 difference	p-value	Log_2 difference	p-value	Log_2 difference
P30086	Phosphatidylethanolamine-binding protein 1	PEBP1			1.39E-02	-1.58		
Q02246	Contactin-2	CNTN2			7.55E-03	-5.31		
Q15818	Neuronal pentraxin-1	NPTX1	2.70E-03	-3.86	1.02E-03	-6.01		
P02452	Collagen alpha-1(I) chain	COL1A1			1.49E-02	-3.41		
O94985	Calsyntenin-1	CLSTN1			2.30E-02	-3.17		
P05067	Amyloid beta A4 protein	APP			4.30E-02	-2.25		
Q9Y6R7	IgGFc-binding protein	FCGBP			4.87E-02	-3.91		
Q14118	Dystroglycan	DAG1			3.42E-02	-3.38		
P16035	Metalloproteinase inhibitor 2	TIMP2			4.60E-03	-2.09		
Q8NFZ8	Cell adhesion molecule 4	CADM4	7.15E-03	-1.15	2.99E-02	-3.67		
P17174	Aspartate aminotransferase, cytoplasmic	GOT1			3.03E-02	-5.22		
Q96KN2	Beta-Ala-His dipeptidase	CNDP1			3.70E-02	-3.79		
P08123	Collagen alpha-2(I) chain	COL1A2			1.24E-02	-3.37		
P01023	Alpha-2-macroglobulin	A2M			4.88E-02	-2.82		
P16870	Carboxypeptidase E	CPE	4.66E-03	-1.72	2.50E-02	-4.97		
Q9P121	Neurotrimin	NTM	8.91E-03	-1.41	3.64E-02	-5.01		
P78324	Tyrosine-protein phosphatase non-receptor type substrate 1	SIRPA			3.55E-02	-1.59		
Q9BYH1	Seizure 6-like protein	SEZ6L	7.26E-03	-3.46	1.86E-03	-4.31		
P02751	Fibronectin	FN1			3.40E-02	-4.17		
Q92859	Neogenin	NEO1			4.19E-02	-4.83		
P51693	Amyloid-like protein 1	APLP1			3.60E-02	-2.98		
Q12805	EGF-containing fibulin-like extracellular matrix protein 1	EFEMP1			8.15E-03	-3.74		
P00736	Complement C1r subcomponent	C1R			3.43E-02	-3.45		

Table 9: Continued

Protein IDs	Protein names	Gene names	MDD vs. CTRL		SCZ vs. CTRL		SCZ vs. MDD	
			p-value	Log_2 difference	p-value	Log_2 difference	p-value	Log_2 difference
Q7Z7M0	Multiple epidermal growth factor-like domains protein 8	MEGF8	4.18E-02	-3.87	6.26E-03	-5.84		
Q92823	Neuronal cell adhesion molecule	NRCAM			2.89E-02	-3.20		
O94856	Neurofascin	NFASC			3.71E-03	-4.60		
P08571	Monocyte differentiation antigen CD14	CD14	1.34E-03	-1.60				
P02649	Apolipoprotein E	APOE			3.46E-02	-1.05		
P01859	Ig gamma-2 chain C region	IGHG2			3.97E-02	-2.17		
P01031	Complement C5	C5			4.03E-02	-4.27		
Q99574	Neuroserpin	SERPINI 1	1.91E-04	-1.74	1.54E-02	-2.87		
O43505	Beta-1,4-glucuronyltransferase 1	B4GAT1	3.10E-02	-1.09	5.07E-02	-2.47		
Q53EL9	Seizure protein 6 homolog	SEZ6	2.58E-02	-3.41	8.83E-03	-4.61		
Q8TAG5	V-set and transmembrane domain-containing protein 2A	VSTM2A			2.33E-02	-1.48		
Q96GW7	Brevican core protein	BCAN			2.58E-02	-5.39		
P02652	Apolipoprotein A-II	APOA2			2.01E-02	1.88		
Q9HCD6	Protein TANC2	TANC2			3.87E-03	-6.15		
P02774	Vitamin D-binding protein	GC			1.47E-02	-2.94		
P01876	Ig alpha-1 chain C region	IGHA1			4.39E-02	-2.05		
P01700	Ig lambda chain V-I region HA	IGLV1-47			1.94E-02	-1.04		
P01019	Angiotensinogen	AGT	3.74E-02	-1.12				
P04196	Histidine-rich glycoprotein	HRG			2.05E-02	-1.98		
P36955	Pigment epithelium-derived factor	SERPINF 1	3.65E-02	-0.96	3.35E-02	-1.83		
Q14515	SPARC-like protein 1	SPARCL 1	3.16E-02	-1.63				

Table 9: Continued

Protein IDs	Protein names	Gene names	MDD vs. CTRL		SCZ vs. CTRL		SCZ vs. MDD	
			p-value	Log_2 difference	p-value	Log_2 difference	p-value	Log_2 difference
Q14982	Opioid-binding protein/cell adhesion molecule	OPCML	4.91E-02	-1.38				
P19021	Peptidyl-glycine alpha-amidating monooxygenase	PAM	1.29E-02	-2.96	1.89E-02	-4.32		
Q7Z3B1	Neuronal growth regulator 1	NEGR1	1.55E-02	-1.29				
O95502	Neuronal pentraxin receptor	NPTXR	4.06E-02	-1.99	1.88E-02	-2.61		
A0A0C4DH38	Immunoglobulin heavy variable 5-51	IGHV5-51	2.23E-02	-0.81				
A0A0B4J1V0	Immunoglobulin heavy variable 3-15	IGHV3-15			1.47E-02	-1.68		
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	ITIH4	4.25E-02	-1.80	4.47E-02	-2.43		
P09972	Fructose-bisphosphate aldolase C	ALDOC			1.46E-02	-2.96		
P27169	Serum paraoxonase/arylesterase 1	PON1			1.89E-02	-1.03		
P27797	Calreticulin	CALR			4.83E-03	-2.21		
P04075	Fructose-bisphosphate aldolase A	ALDOA	1.74E-02	-2.61				
P19022	Cadherin-2	CDH2			3.58E-02	-0.86		
Q9Y4C0	Neurexin-3	NRXN3	1.45E-02	-4.24	1.32E-02	-4.50		

4.2 Functional Annotation Analysis

The protein-protein interactions of the differentially expressed proteins were further analysed using Ingenuity Pathways Analysis. The interactions between altered proteins in SCZ vs. CTRL, MDD vs. CTRL and SCZ vs. MDD were depicted to achieve a better understanding of the interplay on molecular level, as shown in **Figure 11-13**. Most proteins found were assigned as extracellular space proteins or as proteins of the plasma membrane. Processes within the cell appear to play a tangential role since very few proteins were allocated to the cytoplasm and nucleus.

Differently expressed proteins in SCZ vs. CTRL added up to 66 proteins of which 35 were assigned to the extracellular space. Notably, most protein-protein interactions were determined in SCZ vs. CTRL, promoted by the fact that the highest number of differently abundant proteins was identified in this designated group, the most striking was amyloid beta A4 (APP) which is known to interact with 20 other differently abundant proteins, as depicted in **Figure 11**. Moreover, functional analysis of IPA networks implies the interaction of APOE with 13, fibronectin (FN1) with 19 and plasminogen (PLG) with ten other altered proteins in SCZ vs. CTRL. Analysing protein functions, most of the proteins in this group were marked as complex/group/others. However, nine proteins were identified as enzymes and five as transporters. Peptidases and transporters were assigned to the cytoplasm and the extracellular space only while enzymes were also allocated to the plasma membrane, as shown in **Figure 11**.

A lower number of proteins was revealed to be differently abundant in MDD group compared to CTRL. Moreover, the known protein interactions are not as complex as they appear in SCZ vs. CTRL. IPA software analysis indicates the interaction of NPTX1 with three other differently expressed proteins, one assigned to the plasma membrane, two to the cytoplasm. Overall, eleven of the 25 altered proteins in MDD were allocated to the extracellular space, ten to the plasma membrane and four to cytoplasm. Again, plasma membrane proteins and extracellular proteins appear to play key roles when analysing the functions of the differently abundant proteins in MDD disease. Details can be obtained from **Figure 12**.

Comparing the identified proteins in SCZ and MDD cohort most striking proteins regarding protein-protein interactions are 72 kDa type IV collagenase (MMP2) and SPARC. 72 kDa type IV collagenase interacts with five, SPARC with further three differently expressed proteins. Thirteen of the 22 differently abundant proteins were assigned to the compartment extracellular space while the plasma membrane and cytoplasm appear to play minor roles. A notable amount of proteins found to be differently expressed in SCZ vs. MDD function as enzymes and peptidases, as presented in **Figure 13**.

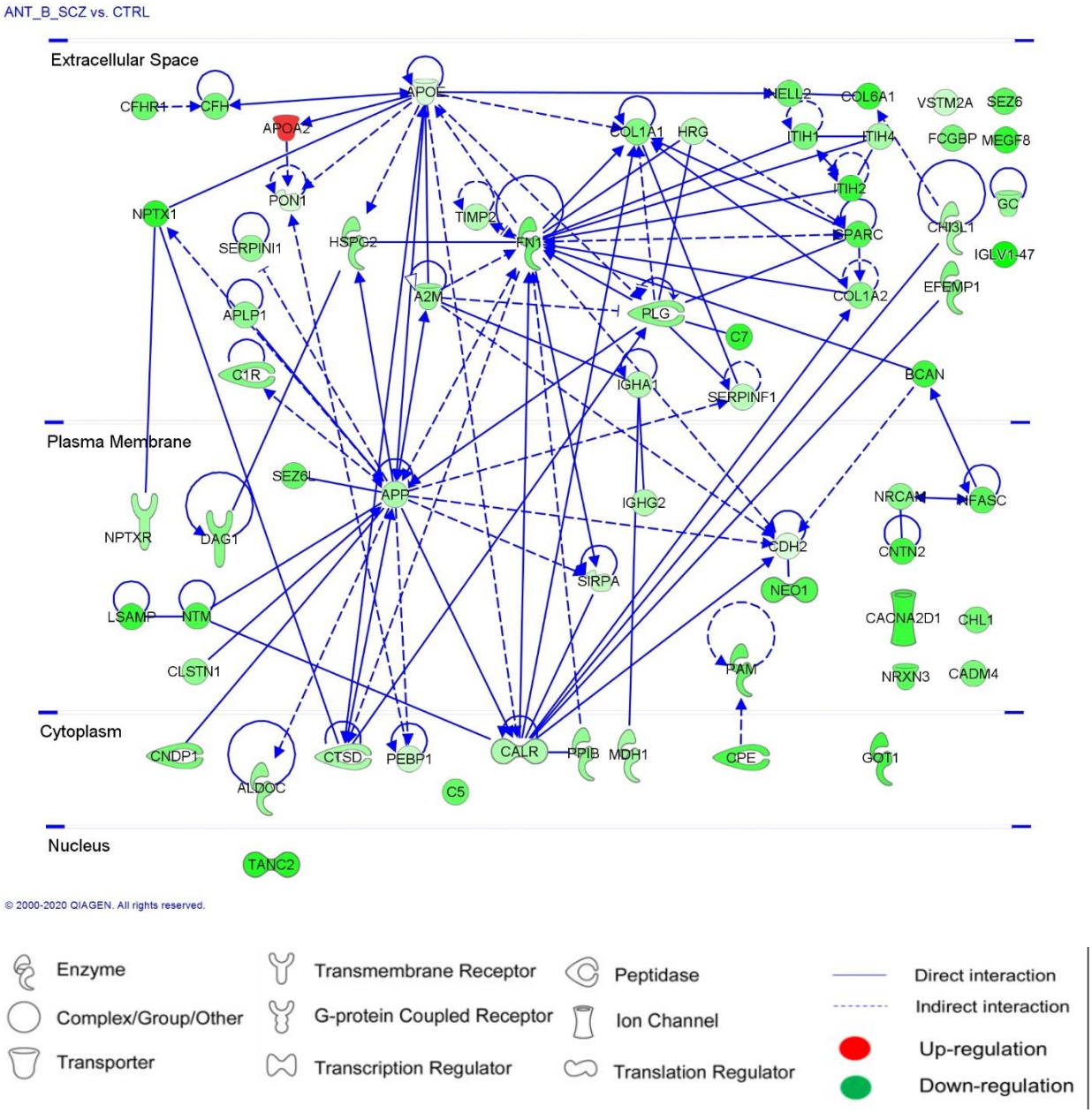

Figure 11: Networks of significantly differently expressed proteins comparing SCZ cohort with healthy individuals analysed with the Ingenuity Pathway Analysis software

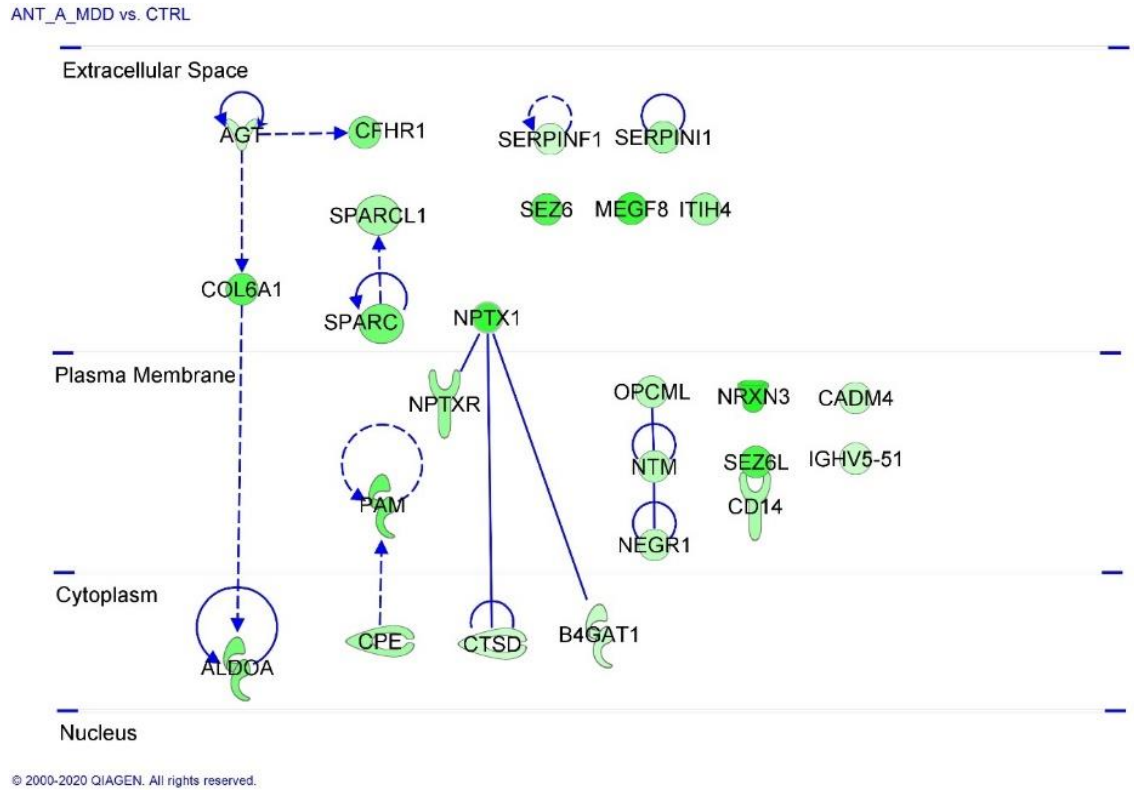

Figure 12: Networks of significantly differently expressed proteins comparing MDD cohort with healthy individuals analysed with the Ingenuity Pathway Analysis software

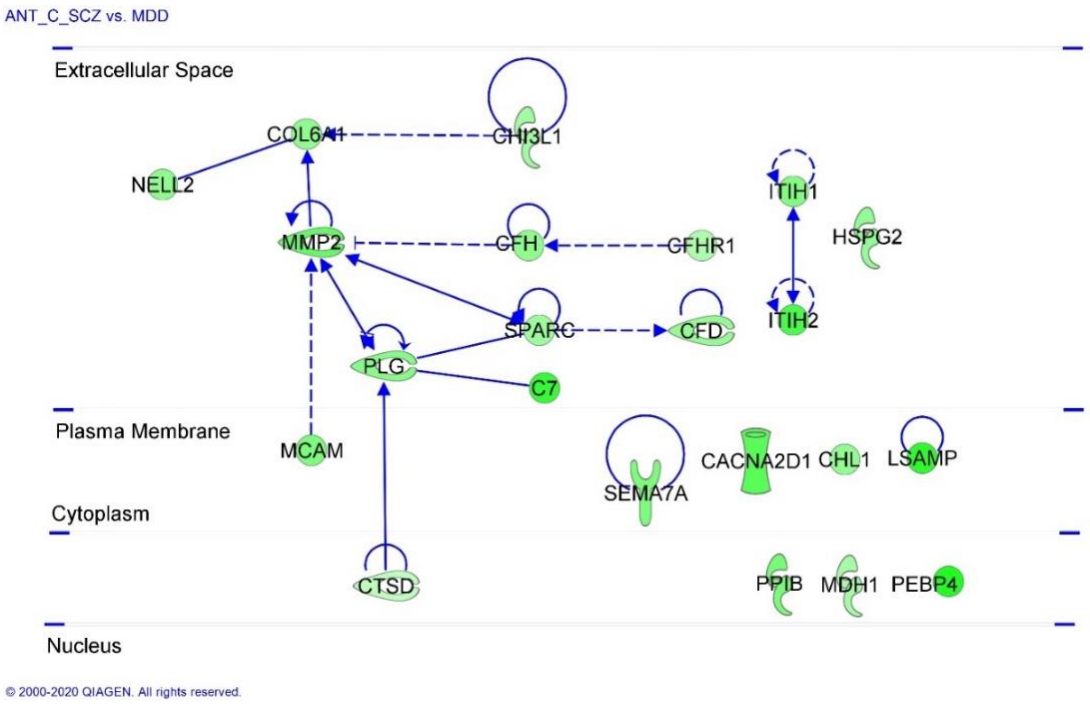

Figure 13: Networks of significantly differently expressed proteins comparing SCZ and MDD cohort analysed with the Ingenuity Pathway Analysis software

The differentially abundant proteins identified in the designated groups underwent a stringent Benjamini-Hochberg multiple testing correction to identify significantly involved canonical pathways and biological functions.

The top five biological functions of the differentially abundant proteins were acute phase response signaling, LXR/RXR activation, hepatic fibrosis/hepatic stellate cell activation, FXR/RXR activation and complement system, as visualized in **Figure 14**. Nine of the proteins found to be differently expressed in SCZ compared to CTRL were highly involved in the acute phase response pathway (p-value = 2.44E-07), namely increased APOA2, decreased histidine-rich glycoprotein (HRG), pigment epithelium-derived factor (SERPINF1), A2M, complement C1r subcomponent (C1R), FN1, Inter-alpha-trypsin inhibitor heavy chain H4 (ITIH4), PLG and ITIH2, as depicted in **Figure 15**.

The altered proteins were related to the following functions and diseases: Organization of extracellular matrix, Alzheimer's disease, progressive encephalopathy, dementia and cell movement. A high p-value (p-value = 7.43E-07) has been determined for Alzheimer's disease. This neurodegenerative disease has been well studied in the past hence many associated proteins have been identified. Sixteen of these proteins also appeared to be altered in the present study when comparing protein intensities of SCZ and CTRL group, namely increased APOA2, decreased APOE, APP, C1R, C7, chitinase-3-like protein 1 (CHI3L1), neural cell adhesion molecule L1-like protein (CHL1), CTSD, vitamin D-binding protein (GC), HRG, basement membrane-specific heparan sulfate proteoglycan core protein (HSPG2), NPTX1, PLG, serum paraoxonase/arylesterase 1 (PON1), SPARC and A2M, as depicted in **Figure 16 A**. **Figure 16 B** presents nine of the differently expressed proteins between SCZ and MDD that were also associated with Alzheimer's disease: CHI3L1, CHL1, CTSD, HSPG2, MMP2, PLG, SPARC, C7 and complement factor D (CFD).

Notably, SCZ also shows up in the list of associated diseases (p-value for SCZ vs. CTRL= 7.61E-05). The differently expressed proteins related to SCZ were APP, CACNA2D1, CHI3L1, aspartate aminotransferase cytoplasmic (GOT1), HSPG2, malate dehydrogenase cytoplasmic (MDH1), protein kinase C-binding protein NELL2 (NELL2), NPTX1, SERPINI1, SEZ6 and fructose-bisphosphate aldolase C (ALDOC), as can be obtained from **Figure 15**. Additionally, many differently abundant proteins in the SCZ group were related to progressive neurological disorder (p-value = 1.95E-06) and severe psychological disorder (p-value= 1.34E-04), details can be obtained from **Figure 17**.

The degree of involvement of the altered proteins in MDD vs. CTRL is relatively low compared to SCZ vs. CTRL which is natural since less proteins were found to be differently abundant in MDD vs. CTRL in the present study. However, three differently expressed proteins in MDD vs. CTRL, namely CD14, ITIH4 and angiotensinogen (AGT), were associated with acute phase reaction (p-value = 2.19E-03), as shown in **Figure 18**.

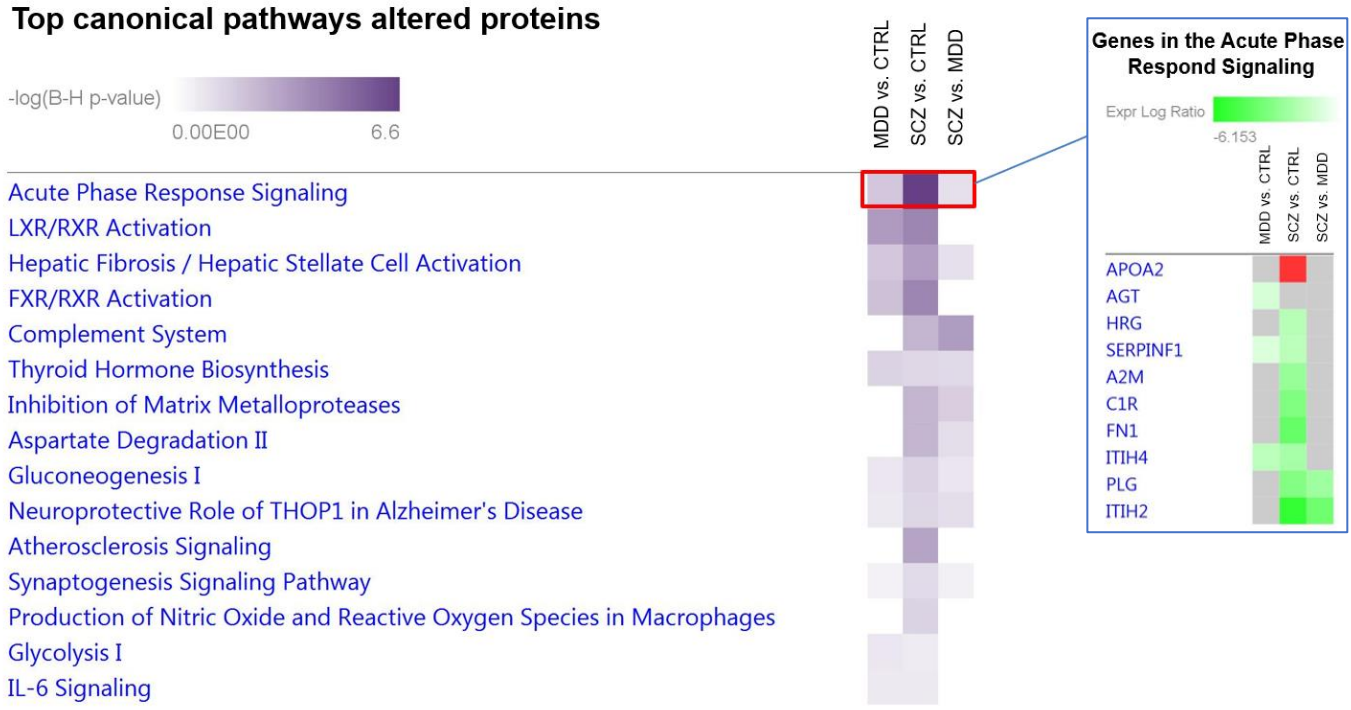

Figure 14: Top canonical pathways of altered proteins

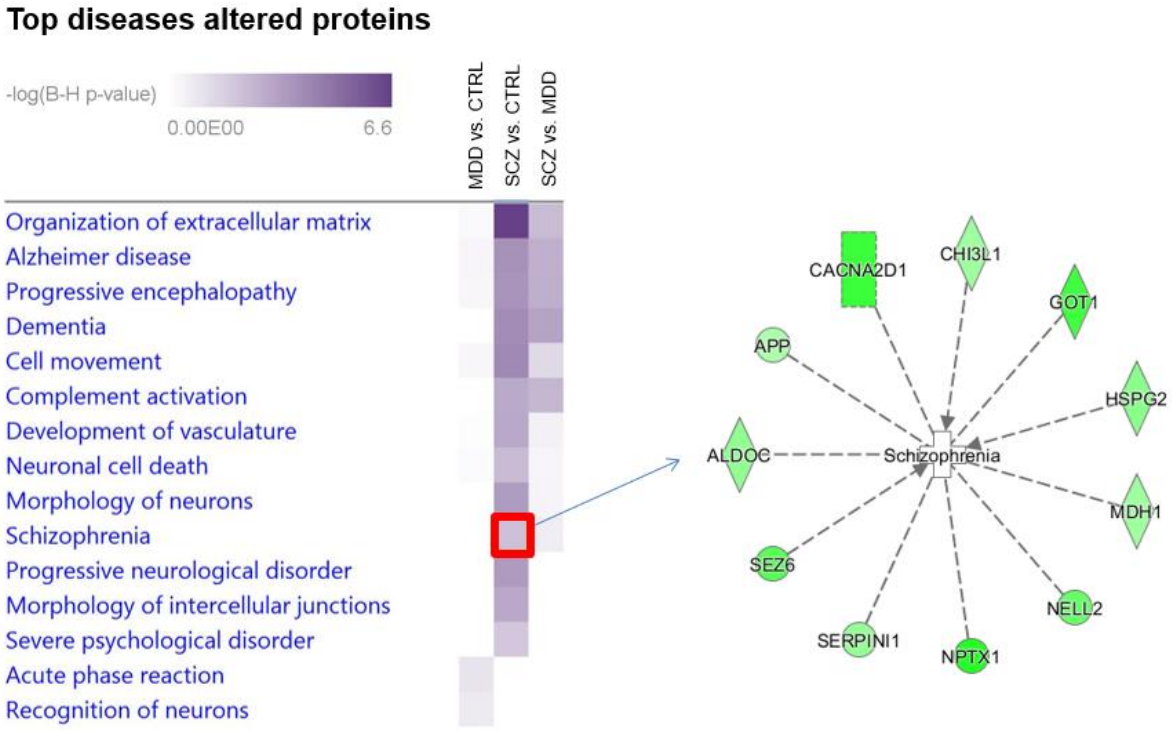

Figure 15: Top diseases of altered proteins

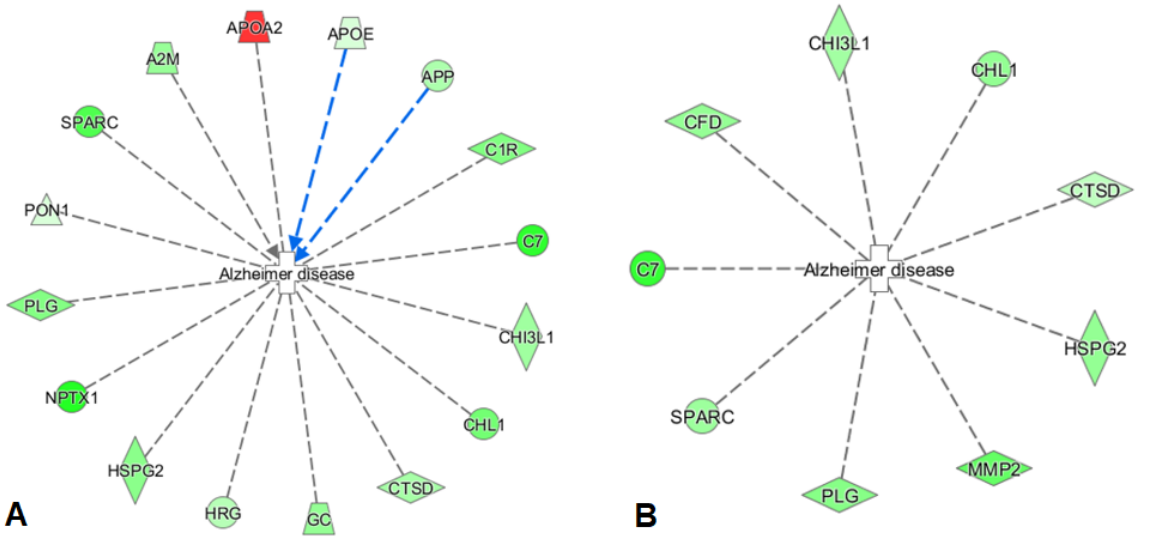

Figure 16: Altered proteins associated with Alzheimer's disease *A: SCZ vs. CTRL, B: SCZ vs. MDD*

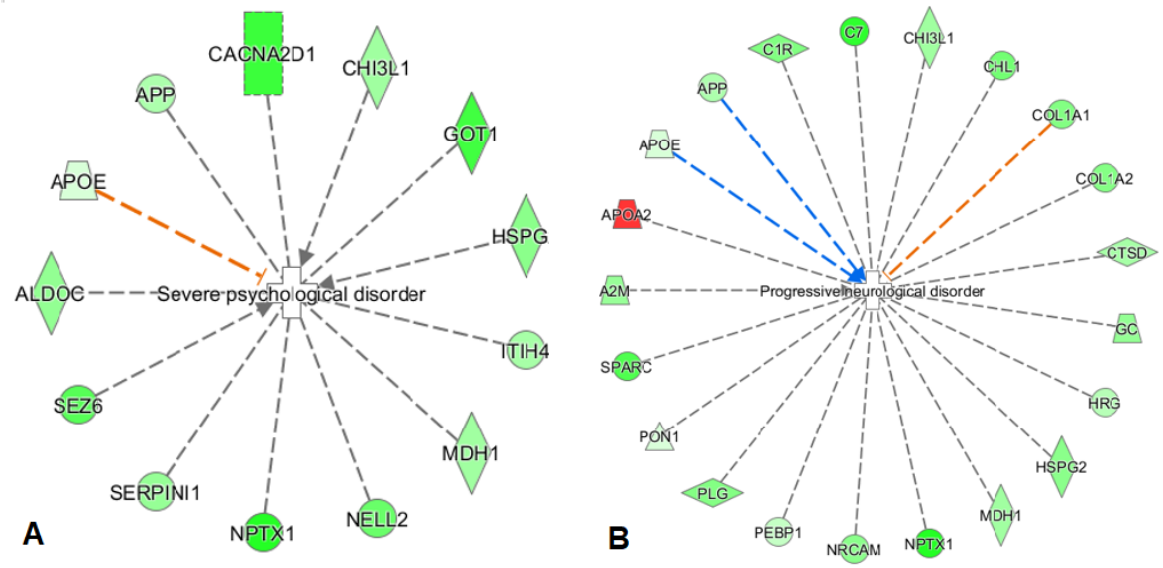

Figure 17: Altered proteins associated with disorders SCZ vs. CTRL *A: Severe psychological disorder, B: Progressive neurological disorder*

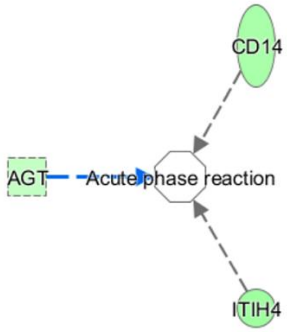

Figure 18: Altered proteins MDD vs. CTRL in acute phase reaction

4.3 Biomarker Discovery Employing In-Solution and Nano-LC-MS/MS

In the second part of the present proteomic discovery analysis, a nano-LC column and ESI-MS/MS system was employed. This technique allows the omission of sample pooling since less total amount of protein is required when employing a nano column enabling the analysis of individual samples. The great advantage of individual sample analysis is that attributes such as gender, race or age can be examined within a cohort. Also, with this method, due to the small diameter of the nano column, a prior protein separation *via* gel may be dispensed consequently reducing the sample preparation to in-solution trypsin digestion. This proteomic discovery approach enabled the identification of a total number of 214 proteins, as presented in **Figure 19**. The total number of proteins detected with the first approach *via* 1DE and capillary-LC-ESI-MS/MS counted 247 proteins. Even without prior protein separation and time-consuming preparation of the samples 86.6% of the total amount of proteins were determined proving the nano-LC column an excellent tool for fast and effective proteomic analysis of CSF. Furthermore, comparing the proteins identified employing the two proteomic workflows, 160 proteins were overlapping, while 87 proteins were unique to the 1DE approach and 54 were exclusively determined employing in-solution and nano-LC column, as depicted in **Figure 19**. Moreover, the nano-LC system analyses each sample individually. This allows a closer look at individual variations in protein intensities enabling the exploration of influences such as age, gender or race.

The comparison of protein intensities in male and female samples revealed seven proteins to be significantly altered. Four proteins, namely SH3 and multiple ankyrin repeat domains protein 2 (SHANK2), complement C5 (C5), Ig alpha-1 chain C region (IGHA1) and insulin-like growth factor-binding protein 7 (IGFBP7), showed higher protein intensities in male patients and three proteins (ankyrin repeat domain-containing protein 36C (ANKRD36C), galectin-3-binding protein (LGALS3BP), pleckstrin homology domain-containing family H member 2 (PLEKHH2)) were higher expressed in female individuals, as shown in **Table 10**. Interestingly the proteins C5 and IGHA1 that both showed higher expression levels in males were also found to be significantly downregulated in CSF of SCZ patients compared to CTRL in the first part of this discovery study. The results are visualized in a heat map (**Figure 20**).

In summary, the advantages of this methods are that the preparation and analysis of samples are less work- and time-consuming implying less potential technical sources of error. In addition, less total amount of proteins is required enabling the analysis of small samples quantities, offering the possibility of individual sample viewing.

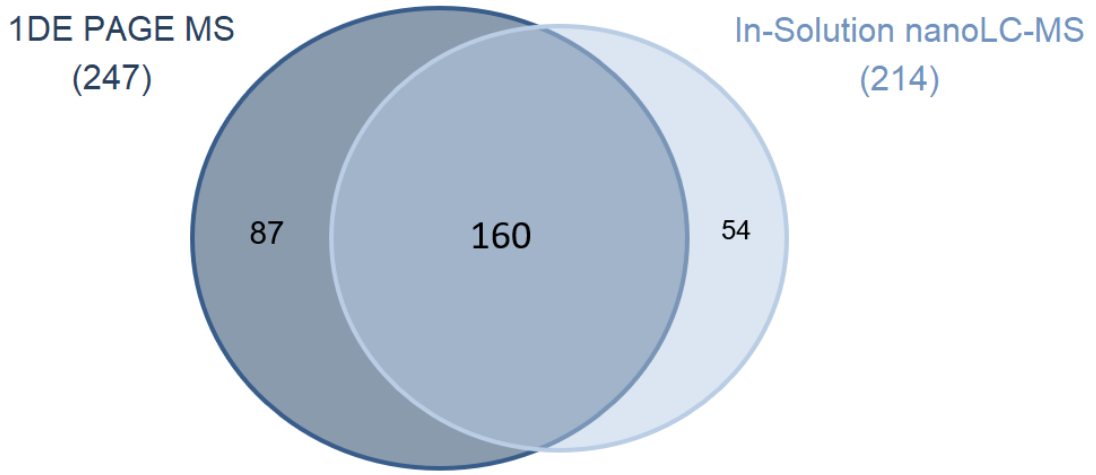

Figure 19: Venn diagram: *Number of detected proteins employing 1DE PAGE MS and in-solution nano-LC-MS.*

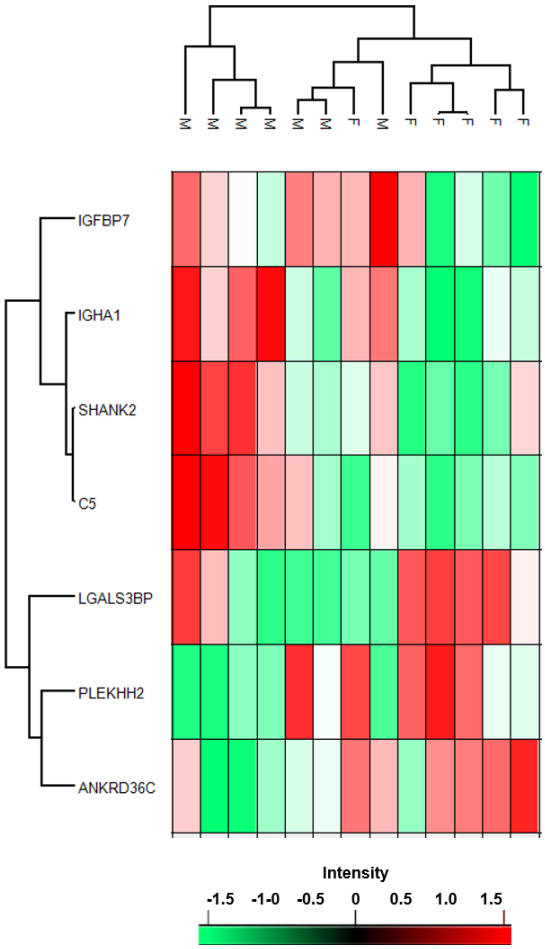

Figure 20: Altered proteins related to gender
Lower intensities are presented in red, higher protein intensities in green. IGFBP7, IGHA1, SHANK2 and C5 and showed higher protein intensities in male patients, LGALS3BP, PLEKHH2 and ANKRD36C were found to be higher expressed in female individuals.

Table 10: Proteins that showed significantly altered expression levels related to gender

Gene names	Student's T-test p-value Female vs. Male	Student's T-test Log_2 fold Difference Female vs. Male
SHANK2	1.21E-2	-2.02
C5	1.7E-3	-1.77
IGHA1	3.68E-2	-1.65
IGFBP7	3.21E-2	-1.64
ANKRD36C	2.55E-2	1.23
LGALS3BP	4.55E-2	2.03
PLEKHH2	1.94E-2	3.30

5 Discussion

5.1 Biomarker Discovery Employing 1DE and Capillary-LC-ESI-MS

In the present study, as many as 79 CSF proteins were found to be differently abundant in the SCZ vs. CTRL, MDD vs. CTRL and SCZ vs. MDD cohort proving excellent sensitivity of the employed proteomic method. The greatest amount of significantly altered proteins was observed in statistical analysis between SCZ cohort and healthy individuals. Moreover, the proteins which showed the most significant differences in expression levels were all observed in the comparison SCZ vs. CRTL while the comparison MDD vs. CTRL revealed less alterations in protein numbers. These results could be explained by the severity of the SCZ disease with deviations in thoughts and behaviour and the complexity of the disorder expressed in a variety of symptoms (12). The least differences in abundance were determined comparing SCZ and MDD samples. MDD and SCZ are both classified as mental disorders (73). Within the mental disorders symptoms can be overlapping (74), it thus seems reasonable that fewer protein alterations were determined comparing the two designated groups.

The comparison of the differently expressed proteins in this study with those discussed in the CSF studies presented in chapter 2.6.1.1 and 2.6.2.1. revealed overlapping alterations in SCZ cohort for APOE and A2M as well as for some APP fragments (**Table 1**). In MDD group altered SERPINF1, COL6A1 and AGT were observed which had been perceived with differences in expression levels in disease in the past, as shown in **Table 2**. Most correlations occurred when comparing our results with the recently published paper by Al Shweiki *et al.* that presented proteomic findings of 161 altered proteins in MDD vs. CTRL and 153 proteins in SCZ vs. CTRL. This study is of particular interest since the group was working with the same study medium (CSF), the same cohort (SCZ and MDD; deviating from the present study they also included BPD patients) and all included patients were under medication. In addition, the group employed comparable methods, except for the prior protein labelling by iTRAQ. Ten correlating differently abundant proteins for MDD were identified comparing the results of Al Shweiki *et al*. and this study, namely AGT, fructose-bisphosphate aldolase A (ALDOA), COL6A1, carboxypeptidase E (CPE), NPTX1, NPTXR, NRXN3, opioid-binding protein/cell adhesion molecule (OPCML), peptidyl-glycine alpha-amidating monooxygenase (PAM) and SERPINI1 and an overlap of 17 differently expressed proteins for SCZ were determined (APOA2, APP, C5, complement factor H (CFH), calsyntenin-1 (CLSTN1), beta-Ala-His dipeptidase (CNDP1), COL6A1, GC, IGHA1, inter-alpha-trypsin inhibitor heavy chain H1 (ITIH1), ITIH2, ITIH4, LSAMP, neurofascin (NFASC), PLG, SERPINI1, tyrosine-protein phosphatase non-receptor type substrate 1 (SIRPA)).

In the SCZ vs. CTRL group APOE appeared to be downregulated. Apolipoprotein E is a core component of a number of lipoproteins and plays an important role in the transport of lipids in CSF (123, 124). The 34 kDa lipoprotein consists of 299 amino acids. There are three alleles (polymorphic variants) existing for the gene located on chromosome 19q13.2 coding for different isoforms, namely ε 2, ε 3 and ε 4. A human being carries two chromosome sets allowing six possible combinations of the alleles resulting in six feasible genotypes (125-127). The occurrence of the different isoforms varies with race, ethnicity and geographical region (128, 129). Apolipoprotein E functions in the transport and metabolism regulation of lipids and is believed to be involved in the reaction of the nervous system to nerve injury (123, 130, 131). As a reaction to inflammation and oxidative stress the turnover of synaptic proteins can change. This process is indirectly related to APOE through its impact on the regulation of cholesterol (106, 132). Altered APOE has been associated with SCZ in the past and has been the subject of multiple studies (106, 127, 128, 133-136). Our finding of decreased APOE corroborates with the observations presented by Wan *et al.* in 2006 comparing CSF of 35 SCZ patients with 36 controls as shown before in **Table 1** (69). Notably, in both cases patients diagnosed with SCZ were under medication. Yet, it is not clear which impact medication has on the expression of APOE. Interestingly, Martins-De-Souza *et al.* detected increased APOE levels in first-onset schizophrenic patients (**Table 1**) (83). None of the included individuals had been treated with antipsychotics for longer than six weeks. Taking this observation into account it is possible that APOE levels are affected by the treatment time with the medication administered. Yet, this assumption conflicts with the findings of Gupta *et al.* who presented lower APOE levels in drug naïve SCZ patients compared to treated individuals (106). Nevertheless, it might not be enough to look at APOE as a general term for all isoforms. Several studies have focused on evaluating the impact of APOE's polymorphic character, seeking to connect frequencies of different alleles with schizophrenia disease. Al-Asmary *et al.* found higher frequencies of allele ε2 and its genotypes and higher frequencies for ε4 allele in schizophrenic patients who showed positive symptoms. For APOE ε3 they concluded a possible protective effect (127). Conflicting results were presented by Akanji *et al*. who found ε2 to be less frequent in SCZ patients. Also, they observed low frequencies of allele ε4 in patients older than 31 at disease onset (128). Jonas *et al*. stated that APOE ε4 was a risk factor for worsening symptoms of hallucinations and delusion that is related to age (133). Kecmanović *et al.* were not able to find any significant alterations in the frequency of APOE alleles or genotype comparing SCZ patients and healthy control (137). This indicates that a further investigation of the different isoforms might be reasonable since APOE variants emerge to be promising biomarkers for future SCZ diagnosis.

Furthermore, decreased SERPINF1 levels were observed when comparing protein intensities of MDD patients versus healthy controls. Pigment epithelium-derived factor is an endogenous 50 kDa glycoprotein which belongs to the group of serine protease inhibitors and consists of 418 amino acids (94, 138-140). The glycoprotein is known to exhibit antitumorigenic, neurotrophic, neuroprotective and antiangiogenetic effects and contributes to anti-inflammation (138, 139, 141). Pigment epithelium-derived factor has been discussed in the context of depression before (94, 138, 142). Ditzen *et al.* found increased levels of SERPINF1 in CSF of twelve depressive patients compared to healthy controls employing 2DE and MALDI-TOF-MS, as shown in **Table 12**. However, it was found to be downregulated in the following targeted LC-MS-based approach. The authors explained this discrepancy by stating that the peptides for the LC-MS approach and 2D-PAGE method are part of different protein isoforms (94). Moreover, a study performed in 2017 found elevated SERPINF1 levels in plasma of depressed patients which were increased further following electroconvulsive therapy (139). Nevertheless, a recently published study found decreased SERPINF1 levels in plasma of drug-naïve MDD patients. However, the same study reported elevated levels in combination with antidepressants (142).

It is not surprising to see the anti-inflammatory protein SERPINF1 to be altered in MDD patients because inflammation is a striking feature in depression (143). Since SERPINF1 is known to be involved in anti-inflammatory activity (138), downregulated SERPINF1 does go in line with the pathophysiology of the disease. However, this accounts for drug-naïve patients; in which way different antipsychotics influence the protein levels is still debatable.

The results of the present study indicate an involvement of SERPINF1 in depression. The glycoprotein may present a potential therapeutic target. Yet, more research regarding the impact of antidepressants on SERPINF1 levels is inevitable. Also, the different isoforms need to be considered.

Next, decreased COL6A1 levels were observed in MDD and SCZ cohort, a protein of the extracellular matrix that provides structural support (144). Similar observations have been previously reported by Ditzen *et al.* in 2012 (94). However, COL6A1 has not been the subject of many studies in the context of depression. Nonetheless, it has been associated with Parkinson's and Alzheimer's disease in the past (145, 146). Since it has been found to be altered before, it might yet be an interesting target for future research.

In the present study, AGT appeared to be downregulated in MDD cohort. Yet, opposite findings have been determined in human CSF before (94). Nevertheless, experiments in mouse model revealed depressive-like behaviour in mice with low AGT level supporting the findings of this study. Angiotensinogen plays a major role in the renin-angiotensin system (RAS) thereby regulating the blood pressure as well as the hydroelectrolyte balance. Apart from the

cardiovascular regulation AGT is also believed to be involved in higher brain activities, e.g. depression, anxiety, memory, adaptive reaction to stressful impulses, cognition and pain perception (147-151). To date, there is not enough data available to unambiguously evaluate the impact of AGT on MDD.

The greatest consensus regarding differently expressed proteins in SCZ and MDD was determined between this study and the study by Al Shweiki *et al.* published in 2020 (16). In SCZ group 17 shared differently expressed proteins were identified. The present study verified downregulated APP, CLSTN1, CNDP1, COL6A1, LSAMP, NFASC, SERPINI1, SIRPA and APOA2. However, differences regarding up- and downregulation of C5, CFH, GC, IGHA1, ITIH1, ITIH2, ITIH4 and PLG were observed. These differences may occur due to the different methods and the sensitivity of the employed MS system. Also, as mentioned before, the diagnosis of SCZ and MDD are subjective and symptoms within the diagnoses can vary. Also, no standardized medication was described. Moreover, the selected individuals for CTRL in this study did not undergo psychiatric assessment in contrary to those of the CTRL cohort of Al Shweiki *et al.* However, the selected age group was very similar in both studies (present study: 37.76 ± 12.35, Al Shweiki *et al*.: 36.3 ± 12.7).
Nevertheless, all ten correlating altered proteins in MDD versus CTRL appeared to be downregulated in their study perfectly matching our results, namely AGT, ALDOA, COL6A1, CPE, NPTX1, NPTXR, NRXN3, OPCML, PAM and SERPINI1.

Overall, there are multiple factors that might limit the correlation between our results and previously detected altered proteins in CSF. Firstly, there are not many study groups that have analysed CSF of SCZ and MDD patients in the past compared to other diseases, e.g. Alzheimer's disease. Secondly, the variety in age and symptoms might contribute to differences in results as well as the administered medication.
Interestingly, in the present study altered proteins have been identified in MDD that have been discussed in the context of SCZ in the past, for example CD14. This clarifies the complexity of mental disorders on protein level and the need for considerable efforts that must be directed to the biomarker research in SCZ and MDD.

5.2 Functional Annotation Analysis

In the illustration of protein interactions of differently abundant proteins in SCZ versus CTRL (**Figure 11**) APP stands out with a central position interacting with 20 other differently expressed proteins. As an integral membrane protein APP is a localized on the cell's surface and is involved in neuronal adhesion, the growth of neurite, axonogenesis and synaptogenesis (152, 153). Located on chromosome 21 (21q21.2-3) the APP gene is coding for three major isoforms of which the 695 amino acid isoform appears to be primarily produced in the brain (154-156). In the present study, APP was found to be downregulated in SCZ group verifying the results of Al Shweiki *et al.* who observed the same phenomenon in SCZ CSF (16). Amyloid-beta A4 protein has been reported to be a molecule of high complexity which is involved in biological processes in full-length form while at the same time presenting a source for fragments which differently effect neurological function (154). Two of these fragments, namely sAPPα and Aβ1-42 were found to be differently abundant in a proteomic CSF study in 2012 (86). Since APP is associated with the regulation of presynaptic structure and function, further investigations regarding the involvement of APP in its full-length confirmation as well as its fragments might be promising to evaluate the probable relation to mental disorders on neuronal level (154). Amyloid-beta A4 protein can be of particular interest due to its interaction with many other proteins and due to the fact that is has been observed to be altered in other SCZ CSF studies before.

Another differently abundant protein in SCZ that was highly involved in protein-protein interactions was FN1. Fibronectin has been discussed in the context of SCZ before, lower FN1 content has been observed in SCZ patient's fibroblasts (157). However, no study reporting altered FN1 levels in CSF was found to date. Fibronectin is believed to be involved in cell shape maintenance, cell adhesion, opsonization, cell motility and wound healing (158-162). It has been suggested that a dysfunction of adhesion molecules may lead to neurodevelopmental abnormalities (163). Therefore, FN1 is an interesting target of investigation in the context of mental diseases.

In the MDD cohort NPTX1 was the altered protein exhibiting most interactions with other differently abundant proteins. It was also among the top three proteins showing the most significant changes in intensities in SCZ as well as in MDD group. Neuronal pentraxin-1 is believed to be involved in the modulation of neuronal excitability thereby negatively regulating the number of excitatory synapses and to limit the plasticity of excitatory synapses (164). Neuronal pentraxin-1 has been suggested as a "plasma biomarker of excitatory synaptic dysfunction" in the past and has been found to be increased in the context of Alzheimer's disease when analysing plasma and brain in mouse model (165). Analysing CSF of MDD

patients NPTX1 was found to be decreased in the study recently published by Al Shweiki *et al.* presenting similar observations than the present study in CSF (16).

Moreover, TANC2 was one of the three most significantly altered proteins in the comparison SCZ vs. CTRL. Protein TANC2 is a scaffold protein that can predominantly be depicted in the developing brain where it is particularly expressed in glia cells and excitatory neurons. Interestingly, it is less pronounced in Drosophila glial cells in which it was reported to have an impact on behavioural outcomes (166). Protein TANC2 expression positively regulates together with TANC1 the denseness of synaptic spines and excitatory synapses. Also, TANC2 plays a role in embryonic development (167); it was reported to be associated to Hippo signalling which is involved in the growing and branching of neurites (168). However, to date, the specific molecular functions of the TANC proteins are still unknown (168). Mutations in the TANC2 gene have been associated with neuropsychiatric disorders in previous publications (166, 168, 169). Nevertheless, to qualify TANC2 protein as a marker for SCZ further investigations are inevitable.

Next, MEGF8 was significantly downregulated in both SCZ and MDD cohort compared to CTRL. Multiple epidermal growth factor-like domains protein 8 is involved in the negative regulation of hedgehog signalling (170, 171). To our knowledge, this is the first study that presents altered MEGF8 in CSF of SCZ and MDD patients. However, MEGF8 has been discussed in the context of Parkinson's disease (PD) before and an involvement in astrocytes' phagocytic function has been hypothesised (172, 173). The association of MEGF8 with PD is of interest since dopamine is believed to play a role in the pathophysiology of both PD and SCZ. While in SCZ a higher expression of dopamine or an oversensitivity of dopamine receptors have been discussed as the causes of the disorder, in PD a decrease of dopamine in the brain, more precisely in the substantia nigra, is believed to lead to the disease (174). Overall, very little is known about the involvement of MEGF8 in neurological disorders (173) but our results might impulse future research in this field.

Neurexin-3 was determined to be significantly downregulated in the comparison MDD vs. CTRL. Correlating with these findings Al Shweiki *et al.* also presented significantly decreased NRXN3 (and CNTNAP4) levels. Neurexin-3 is located on the surface of neuronal cells and may play a role in cell recognition and cell adhesion as well as trans-synaptic signalling (175-177). Al Shweiki *et al.* concluded an impairment of the synaptic signalling in MDD due to the involvement of NRXN3 and CNTNAP4 in synaptic transmission and intercellular property exchange (16). This repeatedly observed alteration in NRXN3 levels provides significant grounds to further investigate NRXN3 as a potential biomarker for MDD in future.

The only protein that appeared to be unregulated in disease compared to control was APOA2 which showed higher expression levels in SCZ cohort. Similar observations in SCZ CSF were described by Al Shweiki *et al.* (16). Deviating findings in serum APOA2 levels were published in 1984 by Sasaki *et al.* (178). Another serum study recently associated APOA2 and other apolipoproteins with metabolic imbalance (179). More data is highly desirable to evaluate the role of apolipoproteins in SCZ.

Performing functional annotation analysis, the differently expressed proteins were associated with abnormalities on neuronal level like neuronal cell death and morphology of neurons. Moreover, the altered proteins were related to multiple diseases of the central and peripheral nervous system, namely dementia, including its most common form Alzheimer disease (180), schizophrenia, progressive neurological und severe psychological disorder. Most of the associated diseases are mental disorders with a shared molecular neuropathology (181). Our results support the assumption of an overlap in pathophysiology within the different neurological disorders and prove the involvement of overlapping altered proteins in diseases. The assumption of a shared pathophysiology stresses the need for biomarker panels rather than single biomarker proteins.

The top canonical pathways related to the proteins found to be altered in the present study included acute phase response (APR) signalling. The APR is a defensive systemic reaction of the innate immune system to stress, infection, trauma, neoplasia and inflammation in order to foster healing (182, 183). Stress and inflammation have been associated with schizophrenia in the past, pro-inflammatory markers, e.g. cytokines, were found to be elevated in blood and CSF of schizophrenic patients (184). Moreover, the influence of stress on depression has been discussed in several publications (185, 186). Stress can induce a hypothalamic-pituitary-adrenal axis activation, a known neurobiological change that has been observed in patients with MDD (185, 187, 188). Stress and inflammation promote a systemic response, it does thus not come as a surprise that alterations in the proteome occur.
Also, in functional annotation analysis the complement system was implicated with the differently expressed proteins. Recently, Ishii *et al.* presented significantly high C5 levels in MDD and SCZ indicating that the activation of the complement system plays a role in the neurological pathophysiology (189).
Moreover, IL-6 signalling was on the list of the top canonical pathways. This result is not surprising since alterations in interleukin-6 levels have been observed in depression (**Table 1**) (97, 190). The pro-inflammatory cytokines are involved in the pathogenesis of depression and the blockade of Il-6 receptor resulted in antidepressant effects in mice (191).

Next, thyroid hormone biosynthesis was among the top associated canonical pathways. Hyper- and hypothyroidism, resulting in concentration changes of thyroid hormones, are believed to be risk factors for depression and lead to psychotic and depressive symptoms in mental diseases (78, 185, 192). The relation between MDD and SCZ decreased levels of thyroid stimulating hormone (TSH) has been a subject of earlier studies (193). A review published by Kirkegaard *et al.* investigated thyroid hormones' role in depression describing decreased thyroxine (T4) levels in total serum, free serum and CSF and no changes in serum triiodothyronine (T3) levels (192). Hence, a correlation appears to exist between changes in thyroid hormone levels and mental diseases.

The present study presents 25 differently abundant proteins in the comparison MDD vs. CTRL and 66 differently expressed proteins in the comparison SCZ vs. CTRL. While some proteins were identified to be differently expressed in these mental diseases for the very first time, 19 of the 66 differently abundant proteins have been presented in literature in the context of SCZ before (**Table 11**). Eleven of the 25 altered proteins in MDD have been discussed as potential biomarkers for MDD in the past (**Table 12**). This offers a new perspective on the mental disorders MDD and SCZ on proteomic level. 1DE gel electrophoresis and capillary-LC-ESI-MS analysis were successfully employed and excellent results were achieved in the present study without prior sample depletion thereby preventing unintentional co-depletion of proteins of interest. The results prove a great sensitivity of the proteomic approach and an objective, reproducible, complex way of protein processing. Even so, further research in the field of psychiatric disorders is inevitable to validate the presented data.

Table 11: Altered proteins SCZ vs. CTRL in this study compared to previous studies

Altered proteins SCZ	**This study**	**Previous studies**
A2M	↓	↓ Gupta *et al*. 2019
APOA2	↑	↑ Al Shweiki *et al.* 2020
APOE	↓	↓ Wan *et al*. 2006, ↑ Martins-De-Souza *et al.* 2010, Tendency: ↓ drug-naive, ↑ treated Gupta *et al.* 2019
APP	↓	↓ Al Shweiki *et al.* 2020, (sAPPα↑; Aβ1-42↓ Albertini *et al.* 2012)
C5	↓	↑ Al Shweiki *et al*. 2020
CFH	↓	↑ Al Shweiki *et al*. 2020
CLSTN1	↓	↓ Al Shweiki *et al.* 2020
CNDP1	↓	↓ Al Shweiki *et al*. 2020
COL6A1	↓	↓ Al Shweiki *et al.* 2020
GC	↓	↑ Al Shweiki *et al.* 2020
IGHA1	↓	↑ Al Shweiki *et al.* 2020
ITIH1	↓	↑ Al Shweiki *et al.* 2020
ITIH2	↓	↑ Al Shweiki *et al.* 2020
ITIH4	↓	↑ Al Shweiki *et al.* 2020
LSAMP	↓	↓ Al Shweiki *et al.* 2020
NFASC	↓	↓ Al Shweiki *et al.* 2020
PLG	↓	↑ Al Shweiki *et al*. 2020
SERPINI1	↓	↓ Al Shweiki *et al.* 2020
SIRPA	↓	↓ Al Shweiki *et al.* 2020
TOTAL: 19		Out of 66 altered proteins in the comparison SCZ versus CTRL 19 proteins have been presented in proteomic CSF studies before, 47 first presented altered proteins in CSF

Table 12: Altered proteins MDD vs. CTRL in this study compared to previous studies

Altered proteins MDD	**This study**	**Previous studies**
AGT	↓	↓ Al Shweiki *et al.* 2020, ↑ Ditzen *et al*. 2011
ALDOA	↓	↓ Al Shweiki *et al.* 2020
COL6A1	↓	↓ Al Shweiki *et al.* 2020, ↓ Ditzen *et al.* 2011
CPE	↓	↓ Al Shweiki *et al.* 2020
NPTX1	↓	↓ Al Shweiki *et al.* 2020
NPTXR	↓	↓ Al Shweiki *et al.* 2020
NRXN3	↓	↓ Al Shweiki *et al.* 2020
OPCML	↓	↓ Al Shweiki *et al.* 2020
PAM	↓	↓ Al Shweiki *et al.* 2020
SERPINF1	↓	↓/↑ Ditzen *et al.* 2011
SERPINI1	↓	↓ Al Shweiki *et al*. 2020
TOTAL: 11		Out of 25 altered proteins in the comparison MDD versus CTRL 11 proteins have been presented in proteomic CSF studies before, 14 first presented altered proteins in CSF

5.3 Biomarker Discovery Employing In-Solution and Nano-LC-MS/MS

In order to optimize the methodology for CSF analysis, another proteomic discovery study was performed employing the in-solution and state-of-art nano-LC-MS strategy.

In this pilot study, CSF from 13 patients diagnosed with SCZ was analysed resulting in the total identification of 214 proteins. This accounts for 86,6% of the number of total proteins that had been determined employing 1DE and capillary-LC-MS/MS. The great number of detected proteins is an excellent achievement since it proves the nano-LC column to be a powerful tool for future proteomic analysis which requires less time for preparation and measurement. Generally, the measurement of one sample employing nano-LC takes 150 min (2,5h), while utilizing 1DE and capillary-MS the total measuring time per sample sums up to 24 hours. This is because each band of one sample (total 24 bands) needs to be measured separately, each measurement taking about 60 min. Therefore, analysis employing nano-LC is ten times faster than 1DE capillary-LC. In addition, less amount of sample volume is required for nano-LC analysis. Twenty-eight times more sample volume is necessary for 1DE capillary-MS analysis (140µl per pool) in comparison to nano-LC (5 µl per sample).

Furthermore, this innovative technique also offers a new perspective on individual variations since samples no longer need to be pooled to reach an adequate protein concentration. This advantage allows the comparison of protein intensities in individuals rather than in cohorts. In summary, nano-LC is a faster tool for MS analysis that requires less sample volume and enables individual sample analysis. The comparison of male and female patients becomes possible and the role of gender in schizophrenia can be examined further. In statistical analysis four proteins (SHANK2, C5, IGHA1, IGFBP7) were found to be upregulated in male individuals in this study while three proteins (ANKRD36C, LGALS3BP, PLEKHH2) were found to be higher expressed in female patients.

Even though men and women show identical symptoms in SCZ during the acute phases of the illness and an equivalent distribution of schizophrenia exists between the sexes, several known gender-specific differences have been described in literature (194, 195). Women appear to be of higher age at disease onset, to have a better disease progression, a better premorbid adjustment and a clearer symptom profile. Also, cognitive deficits as well as structural brain abnormalities vary between the sexes (194). Moreover, women show better outcome of disease compared to men under medication as well as without drug treatment. Several factors may contribute to the advantage: Firstly, drug metabolism results in relatively high blood levels in women compared to men. Secondly, women appear to follow drug prescription more likely. Thirdly, while men are more likely to show cognitive and negative symptoms, women tend to suffer from hallucinations and delusions. These symptoms appear to respond better to

treatment with antipsychotics (196-198). Another factor of interest is the influence of gonadal hormones in schizophrenia since the onset of illness is concurring with a drop in oestrogen levels and oestrogen is believed to positively influence the disease outcome by enhancing dopamine blockade (196, 199, 200). Since there are known sex differences in SCZ disease these differences may also be reflected on proteomic level.

In statistical analysis seven proteins were found to be altered in male versus female in this study (SHANK2, C5, IGHA1, IGFBP7, ANKRD36C, LGALS3BP and PLEKHH2). Two of the proteins with higher protein levels in males, namely C5 and IGHA1, had also shown significantly altered intensities in SCZ compared to CTRL.

SH3 and multiple ankyrin repeat domains protein 2 belongs to the SHANK gene family which code for postsynaptic scaffold proteins in excitatory neurons. SHANK proteins form a matrix structure functioning as a “net” to organize other proteins in the postsynapse. SH3 and multiple ankyrin repeat domains protein 2 has been associated with neuropsychiatric disorders, primarily with autism spectrum disorder, but also with schizophrenia in the past (201, 202). An interesting paper published by Berkel *et al*. investigated the influence of sex hormones on the expression of shank proteins. The group concluded that there is a difference in expression of SHANK between the sexes since they found SHANK expression to be related to hormone expression. Moreover, they state that estrogen and androgen receptor antagonists can block the effect Dihydrotestosterone (DHT) and 17β-Estradiol have on the expression of the SHANK genes. Their findings let Berkel *et al.* suggest a “direct regulatory influence of DHT on SHANK2” (203). This leads to the questions to which extent increased SHANK levels in males are physiological and if hormonal expression levels beyond physiological ranges may contribute to the development and progress of SCZ.

Complement C5 expression has also been associated with sex hormone levels. Baba *et al.* stated that testosterone promotes C5 expression in mice (204). Complement C5 is a circulating innate immunity protein that has also been observed to be differently expressed between the sexes in mouse model in a study of liver metastasis caused by pancreatic neuroendocrine tumours (205). The role of C5 as part of an immune response in SCZ is not well studied, yet, it has been determined to be differently abundant in this study using two different proteomic approaches and may thus be an interesting target for future research.

No publications regarding gender/sex differences and differences in IGHA1 expression levels were found. Immunoglobulins are produced by B lymphocytes. As part of the humoral immunity, they function as membrane-based receptors that trigger the differentiation of B

lymphocytes into plasma cells that secrete immunoglobulins when binding a specific antigen (206-208).

Insulin-like growth factor-binding protein 7, the name itself explaining its main molecular function, is involved in a variety of biological processes like cell growth regulation, cell adhesion and metabolic processes as well as in the regulation of cell proliferation (209-211). Insulin-like growth factor-binding protein 7 (and IGFBP3) have been observed to be associated with different forms of cancer, including breast, gastroesophageal, prostate and colon cancer (212-215). Insulin-like growth factor-binding protein 7 appears to be upregulated in some cancer types (e.g. oesophageal cancer) while in others it was observed to be downregulated (e.g. Melanoma) (212). Gender and age differences in IGFBP7 expression as a marker in patients with acute kidney injury have been described by Boddu *et al.* (216). Insulin-like growth factor-binding protein 7 gene was also examined in the context of sex effects on the risk for opioid dependence (217).

Ankyrin repeat domain-containing protein 36C is part of the ANKDRD36 family located in chromosome 2 which have been associated with inflammation (218, 219). Ankyrin repeat domain-containing protein 36C is believed to function as an inhibitor on ion channels (220, 221).

Galectin-3-binding protein's main molecular function is its scavenger receptor activity (222). Galectin-3-binding protein is believed to be involved in several biological processes such as cell adhesion, signal transduction and cellular defence reaction (222-224). A study by Melin *et al.* in 2019 demonstrated the association between high LGALS3BP levels and elevated total cholesterol and sCD163 levels, reduced HDL-cholesterol levels and, of interest regarding sexual dysmorphism, the female sex in patients diagnosed with type 1 diabetes. According to the study, women showed a prevalence of LGALS3BP levels twice as high as men (225).

Pleckstrin homology domain-containing family H member 2, an intracellular protein in podocytes which appears to play a role in actin dynamics and matrix adhesion, has been associated with schizophrenia in previous publications (226, 227). However, no study analysing PLEKHH2 expression in the aspect of variations in gender was found.

This study successfully identified seven proteins associated with gender in SCZ cohort. The method could be potentially targeted for future clinical studies regarding diagnostic and drug development. However, in the present study out of 271 detected proteins, only seven proteins showed significant changes in protein levels, suggesting that gender does not considerably influence the CSF proteome.

5.4 Limitations

The major limitation of this study is that all patients were drug treated and medication varied between the individuals. Since one cannot foresee the impact of drugs and their combinations on proteomic level it cannot be ruled out which protein alterations are manifestations of the disease and which have adapted as a biological response to medication. Also, as mentioned before, mental disorders are highly complex and clinical expressions can vary. No correlation between symptoms and altered proteins can be concluded from the pooled samples. Hence, there is a need to critically scrutinise the selection of patients regarding age and symptoms. It might be necessary to analyse subgroups of patients in order to answer fundamental yet elusive questions regarding the pathophysiology of SCZ and MDD.
Overall, a larger number of samples is needed to verify the results of the present study and to obtain representative alterations in protein concentrations.
In the present study the CTRL group consisted of patients that underwent lumbar puncture for diagnostic reasons that had no relation to mental disorders. However, patients of the control group were not examined by psychologists to exclude patients with psychiatric diseases.

To conclude, the research on the pathophysiology of MDD and SCZ is still in an early stage and this study offers a great input of potentially altered proteins in disease which may contribute to the identification of biomarker sets in the future.

6 Conclusion

Herein, we report a proteomic investigation of CSF samples taken from depressed and schizophrenic patients and healthy controls resulting in the identification of a total of 277 proteins. Since the proteome is a dynamic system that reflects the current state and environment of an organism, alterations on protein level were explored in order to achieve a better understanding of the pathophysiology behind the mental disorders. Sixty-six proteins were found to be differently abundant in the comparison SCZ versus healthy individuals and 25 proteins in the comparison MDD versus CTRL. All altered proteins except for APO2 appeared to be downregulated in disease.

This study verifies altered expression levels of 11 proteins that have been presented in previous studies. These 11 proteins, namely AGT, ALDOA, COL6A1, CPE, NPTX1, NPTXR, NRXN3, OPCML, PAM, SERPINF1 and SERPINI1, may be potential biomarkers for MDD disease. Moreover, 14 proteins were observed to be differently expressed in MDD for the first time. These potential biomarker candidates need to be investigated further in the future.

In SCZ cohort, 19 differently abundant proteins were identified that have been discussed as potential biomarkers for SCZ in previous studies (A2M, APOA2, APOE, APP, C5, CFH, CLSTN1, CNDP1, COL6A1, GC, IGHA1, ITIH1, ITIH2, ITIH4, LSAMP, NFASC, PLG, SERPINI1 and SIRPA). However, the findings regarding up and downregulations are partly conflicting suggesting that the molecular diversity and the expression on proteomic level is even more complex than the variation in clinical symptoms. Extensive research under consideration of the administered medication is inevitable for progress in this field. Furthermore, 47 differently abundant proteins were determined that may be involved in SCZ pathophysiology and which to our knowledge have not been observed to be altered in previous studies.

Mental disorders share some clinical manifestations indicating that the pathophysiology behind the diseases is partly alike. This assumption is supported by the shared alterations in expression levels of some proteins in SCZ and MDD. It thus becomes clear that the complexity of mental disorders may require a set of biomarkers (biosignature) rather than single altered proteins as markers for future diagnostic and treatment. Further validation of the proteins found to be altered in the present study is of great interest. To avoid technical errors a larger patient cohort and a division of patients according to their symptoms and age is advisable. Also, the analysis of drug naïve patients enables a better understanding of the manifestation of mental disorders on proteomic level. The findings prove the employment of 1DE gel and capillary-LC-ESI-MS to be a reliable and solid method for the analysis of CSF.

At the same time, this study presents an alternative method for proteomic analysis which enabled the detection of about 86.6% of the proteins detected with 1DE gel and LC-ESI-MS.

This method, in-solution technique combined with nano-LC-ESI-MS/MS, is less time consuming and fewer preparation steps minimize the sources of technical errors. Moreover, samples can be analysed individually which opens new perspectives; for example, a comparison of differences in expression levels between the sexes becomes possible. In the SCZ cohort seven proteins were identified that showed differences in expression correlating with gender (SHANK2, C5, IGHA1, IGFBP7, ANKRD36C, LGALS3BP, PLEKHH2) of which two had also been observed to be downregulated in SCZ cohort compared to control. Even though only a small number of samples were analysed employing in-solution and nano-LC MS we can state that this method is an excellent alternative to the 1DE gel approach. In addition, it proves that the field of proteomics is rapidly improving in view of technology and that highly sensitive tools are available these days for the investigation of dynamic body fluids. This study provides numerous proteins that are possibly related to the mental disorders SCZ and MDD and which have a great potential to contribute to the diagnostic and treatment of these diseases in the future.

6.1 Zusammenfassung

In dieser Studie ist es gelungen, mittels proteomischer Analyse von Liquorproben depressiver und schizophrener Patienten sowie einer gesunden Kontrollgruppe, insgesamt 277 Liquorproteine zu identifizieren. Um ein besseres Verständnis der Pathophysiologie hinter den psychischen Erkrankungen Schizophrenie (SCZ) und Depression (MDD) zu erlangen, wurden Veränderungen auf der Proteinebene untersucht, da das Proteom ein dynamisches System ist, das den aktuellen Zustand und die Umgebung eines Organismus widerspiegelt. Im direkten Vergleich der Proteinintensitäten der Kohorten SCZ und CTRL konnten 66 signifikant unterschiedlich exprimierte Proteine identifiziert werden, 25 unterschiedlich abundante Proteine im Vergleich MDD und CTRL. Alle veränderten Proteine mit Ausnahme von APO2 erschienen im Krankheitsfall herunterreguliert.

Diese Studie verifiziert die veränderte Expression von 11 Proteinen, die in der Vergangenheit für MDD beschrieben wurden. Diese 11 Proteine, AGT, ALDOA, COL6A1, CPE, NPTX1, NPTXR, NRXN3, OPCML, PAM, SERPINF1 und SERPINI1, könnten potenzielle Biomarker für MDD sein. Darüber hinaus wurden 14 weitere Proteine identifiziert, die zum ersten Mal im Zusammenhang mit MDD als unterschiedlich exprimiert beschrieben wurden. Diese potenziellen Biomarker-Kandidaten sollten in der Zukunft weiter untersucht werden. In der SCZ-Kohorte wurden 19 unterschiedlich abundante Proteine identifiziert, die in früheren Studien als potenzielle Biomarker für SCZ diskutiert wurden (A2M, APOA2, APOE, APP, C5, CFH, CLSTN1, CNDP1, COL6A1, GC, IGHA1, ITIH1, ITIH2, ITIH4, LSAMP, NFASC, PLG, SERPINI1 und SIRPA). Die Ergebnisse hinsichtlich der Hoch- und Herunterregulierung sind jedoch teilweise widersprüchlich, was darauf hindeutet, dass die molekulare Vielfalt und die Expression auf proteomischer Ebene noch komplexer ist als die Variation der klinischen Symptome. Umfangreiche Untersuchungen unter Berücksichtigung der verabreichten Medikamente sind für Fortschritte auf diesem Gebiet unumgänglich. Darüber hinaus wurden 47 unterschiedlich exprimierte Proteine bestimmt, die möglicherweise in die Pathophysiologie die SCZ Erkrankung involviert sind und die unseres Wissens zufolge in noch keiner anderen früheren Studie im Zusammenhang mit SCZ diskutiert wurden.

Psychische Erkrankungen haben einige klinische Manifestationen gemeinsam, was darauf hindeutet, dass sich die Pathophysiologie hinter den Krankheitsbildern teilweise ähnelt. Diese Annahme wird durch Proteine, die sowohl in der SCZ- als auch in der MDD-Kohorte als unterschiedlich exprimiert identifiziert wurden, gestützt. Damit wird deutlich, dass die Komplexität psychischer Störungen möglicherweise eine Reihe von Biomarkern (Biosignatur) und nicht nur einzelne veränderte Proteine als Marker für die künftige Diagnose und Behandlung erfordert. Eine weitere Validierung der in der vorliegenden Studie festgestellten veränderten Proteine ist von großem Interesse. Um potenzielle Fehlerquellen zu vermeiden, ist eine größere Patientenkohorte und eine Kategorisierung der Patienten nach Symptomen

und Alter ratsam. Außerdem ermöglicht die Analyse von Patienten, die keine Medikamente einnehmen, ein besseres Verständnis der Manifestation psychischer Erkrankungen auf proteomischer Ebene.

Die Ergebnisse belegen, dass der Einsatz von 1DE-Gel und LC-ESI-MS eine zuverlässige und solide Methode für die Analyse von Liquor cerebrospinalis darstellt.

Gleichzeitig wird in dieser Studie eine alternative Methode zur Proteomanalyse vorgestellt, mit der etwa 86,6 % der mit 1DE-Gel und LC-ESI-MS nachgewiesenen Proteine erfasst werden konnten. Diese Methode, In-Solution-Technik in Kombination mit nano -LC-ESI-MS/MS, ist weniger zeitaufwendig und weniger Vorbereitungsschritte minimieren Fehlerquellen. Darüber hinaus können die Proben individuell analysiert werden, was neue Perspektiven eröffnet; so wird beispielsweise ein Vergleich der Geschlechter möglich. In der SCZ-Kohorte wurden in dieser Studie sieben Proteine identifiziert, die geschlechtsspezifische Unterschiede in der Expression aufwiesen (SHANK2, C5, IGHA1, IGFBP7, ANKRD36C, LGALS3BP, PLEKHH2), von denen zwei auch in der SCZ-Kohorte im Vergleich zur Kontrollgruppe herunterreguliert waren. Obwohl nur eine kleine Anzahl von Proben mit In-Solution und nano-LC MS analysiert wurde, können wir feststellen, dass diese Methode eine hervorragende Alternative zum 1DE-Gel-Ansatz darstellt. Darüber hinaus beweist sie, dass sich der Bereich der Proteomik technologisch rasant weiterentwickelt und heute hochempfindliche Instrumente für die Untersuchung dynamischer Körperflüssigkeiten zur Verfügung stehen.

Diese Studie liefert zahlreiche Proteine, die möglicherweise mit den psychischen Erkrankungen SCZ und MDD in Verbindung stehen und die ein großes Potenzial haben, in Zukunft zur Diagnose und Behandlung dieser Krankheiten bei.

Reference List

1. Severson M, Strecker-McGraw MK. Cerebrospinal Fluid Leak. StatPearls. Treasure Island (FL): StatPearls Publishing; 2022: 1
2. Waybright TJ. Preparation of human cerebrospinal fluid for proteomics biomarker analysis. Methods Mol Biol. 2013;1002:61-70.
3. Johnson MD, Yu LR, Conrads TP, Kinoshita Y, Uo T, McBee JK, et al. The proteomics of neurodegeneration. Am J Pharmacogenomics. 2005;5(4):259-70.
4. Fernández-Irigoyen J, Santamaría E, editors. Cerebrospinal Fluid (CSF) Proteomics Methods and Protocols. New York, NY: Springer Science+Business Media, LLC, part of Springer Nature 2019; 2019: 4.
5. Kitchen RR, Rozowsky JS, Gerstein MB, Nairn AC. Decoding neuroproteomics: integrating the genome, translatome and functional anatomy. Nat Neurosci. 2014;17(11):1491-9.
6. Graves PR, Haystead TA. Molecular biologist's guide to proteomics. Microbiol Mol Biol Rev. 2002;66(1):39-63.
7. Huang JT, McKenna T, Hughes C, Leweke FM, Schwarz E, Bahn S. CSF biomarker discovery using label-free nano-LC-MS based proteomic profiling: technical aspects. J Sep Sci. 2007;30(2):214-25.
8. Mesri M. Advances in Proteomic Technologies and Its Contribution to the Field of Cancer. Adv Med. 2014;2014:238045.
9. Fountoulakis M. Application of proteomics technologies in the investigation of the brain. Mass Spectrom Rev. 2004;23(4):231-58.
10. Huang JT, Leweke FM, Oxley D, Wang L, Harris N, Koethe D, et al. Disease biomarkers in cerebrospinal fluid of patients with first-onset psychosis. PLoS Med. 2006;3(11):e428.
11. Pan American Health Organization. Mental health problems are the leading cause of disability worldwide [Internet]. Washington: Regional Office for the Americas of the World Health Organization; 2019 [cited 2022 May 08]. Available from: https://www3.paho.org/hq/index.php?option=com_content&view=article&id=15481:mental-health-problems-are-the-leading-cause-of-disability-worldwide-say-experts-at-paho-directing-council-side-event&Itemid=72565&lang=en.
12. Vasic N, Connemann BJ, Wolf RC, Tumani H, Brettschneider J. Cerebrospinal fluid biomarker candidates of schizophrenia: where do we stand? Eur Arch Psychiatry Clin Neurosci. 2012;262(5):375-91.
13. Tapiola T, Alafuzoff I, Herukka SK, Parkkinen L, Hartikainen P, Soininen H, et al. Cerebrospinal fluid {beta}-amyloid 42 and tau proteins as biomarkers of Alzheimer-type pathologic changes in the brain. Arch Neurol. 2009;66(3):382-9.
14. Zhang J, Goodlett DR, Quinn JF, Peskind E, Kaye JA, Zhou Y, et al. Quantitative proteomics of cerebrospinal fluid from patients with Alzheimer disease. J Alzheimers Dis. 2005;7(2):125-33; discussion 73-80.
15. Finehout EJ, Franck Z, Choe LH, Relkin N, Lee KH. Cerebrospinal fluid proteomic biomarkers for Alzheimer's disease. Ann Neurol. 2007;61(2):120-9.
16. Al Shweiki MR, Oeckl P, Steinacker P, Barschke P, Dorner-Ciossek C, Hengerer B, et al. Proteomic analysis reveals a biosignature of decreased synaptic protein in cerebrospinal fluid of major depressive disorder. Transl Psychiatry. 2020;10(1):144.
17. Suzuki M, Nohara S, Hagino H, Takahashi T, Kawasaki Y, Yamashita I, et al. Prefrontal abnormalities in patients with simple schizophrenia: structural and functional brain-imaging studies in five cases. Psychiatry Res. 2005;140(2):157-71.
18. Amboss. Meningen, Liquorräume und Blut-Hirn-Schranke [Internet]. Berlin: amboss; 2020 [cited 2019 December 17]. Available from: https://www.amboss.com/de/wissen/Meningen%2C_Liquorr%C3%A4ume_und_Blut-Hirn-Schranke#xid=ZJ0ZsS&anker=Zeb38a456b50015636e8a5b4b07bfbbe0.

19. Zougman A, Pilch B, Podtelejnikov A, Kiehntopf M, Schnabel C, Kumar C, et al. Integrated analysis of the cerebrospinal fluid peptidome and proteome. J Proteome Res. 2008;7(1):386-99.
20. Conly JM, Ronald AR. Cerebrospinal fluid as a diagnostic body fluid. Am J Med. 1983;75(1b):102-8.
21. Bailey CS, Vernau W. Chapter 27 - Cerebrospinal Fluid. In: Kaneko JJ, Harvey JW, Bruss ML, editors. Clinical Biochemistry of Domestic Animals (Fifth Edition). San Diego: Academic Press; 1997. p. 785-827.
22. Brown PD, Davies SL, Speake T, Millar ID. Molecular mechanisms of cerebrospinal fluid production. Neuroscience. 2004;129(4):957-70.
23. Huhmer AF, Biringer RG, Amato H, Fonteh AN, Harrington MG. Protein analysis in human cerebrospinal fluid: Physiological aspects, current progress and future challenges. Dis Markers. 2006;22(1-2):3-26.
24. Lun MP, Monuki ES, Lehtinen MK. Development and functions of the choroid plexus-cerebrospinal fluid system. Nat Rev Neurosci. 2015;16(8):445-57.
25. Louveau A, Smirnov I, Keyes TJ, Eccles JD, Rouhani SJ, Peske JD, et al. Structural and functional features of central nervous system lymphatic vessels. Nature. 2015;523(7560):337-41.
26. Pardridge WM. CSF, blood-brain barrier, and brain drug delivery. Expert opinion on drug delivery. 2016;13(7):963-75.
27. Reiber H. Proteins in cerebrospinal fluid and blood: barriers, CSF flow rate and source-related dynamics. Restor Neurol Neurosci. 2003;21(3-4):79-96.
28. Uhr M, Tumani H, Lange P. Strategies for cerebrospinal fluid analysis - Integrated results report. Nervenarzt. 2016;87(12):1271-5.
29. Gluck T, Reng M, Linde HJ, Scholmerich J. Cerebrospinal fluid analysis. Internist (Berl). 2000;41(3):293-303.
30. Jiang L, Lindpaintner K, Li HF, Gu NF, Langen H, He L, et al. Proteomic analysis of the cerebrospinal fluid of patients with schizophrenia. Amino Acids. 2003;25(1):49-57.
31. Schünke M, Schulte E, Schumacher U, Voll M, Wesker K. Prometheus LernAtlas der Anatomie, Kopf, Hals und Neuroanatomie. 5. Auflage ed. Stuttgart: Thieme Verlag; 2005: 314, 419, 462.
32. Santa C, Coelho J, Madeira N, Manadas B. Circulating biomarkers in schizophrenia: a proteomics perspective. International Journal of Clinical Neurosciences and Mental Health. 2017.
33. Harrison PJ. Using our brains: the findings, flaws, and future of postmortem studies of psychiatric disorders. Biol Psychiatry. 2011;69(2):102-3.
34. Kim SI, Voshol H, van Oostrum J, Hastings TG, Cascio M, Glucksman MJ. Neuroproteomics: expression profiling of the brain's proteomes in health and disease. Neurochem Res. 2004;29(6):1317-31.
35. Waybright T, Avellino AM, Ellenbogen RG, Hollinger BJ, Veenstra TD, Morrison RS. Characterization of the human ventricular cerebrospinal fluid proteome obtained from hydrocephalic patients. J Proteomics. 2010;73(6):1156-62.
36. Macron C, Lane L, Nunez Galindo A, Dayon L. Deep Dive on the Proteome of Human Cerebrospinal Fluid: A Valuable Data Resource for Biomarker Discovery and Missing Protein Identification. J Proteome Res. 2018;17(12):4113-26.
37. Hrishi AP, Sethuraman M. Cerebrospinal Fluid (CSF) Analysis and Interpretation in Neurocritical Care for Acute Neurological Conditions. Indian J Crit Care Med. 2019;23(Suppl 2):S115-s9.
38. Deutsche Gesellschaft für Liquordiagnostik und Klinische Neurochemie e.V. Ausgewählte Methoden der Liquordiagnostik und Klinischen Neurochemie [Internet]. 3. Auflage. 2014 [cited 2022 May 08] Available from: https://07525720-0688-4380-840d-0a4af942fef7.filesusr.com/ugd/92c932_454e4d6908d94f64b3623b621179eade.pdf.
39. Seehusen DA, Reeves MM, Fomin DA. Cerebrospinal fluid analysis. Am Fam Physician. 2003;68(6):1103-8.

40. Pedrero-Prieto CM, García-Carpintero S, Frontiñán-Rubio J, Llanos-González E, Aguilera García C, Alcaín FJ, et al. A comprehensive systematic review of CSF proteins and peptides that define Alzheimer's disease. Clin Proteomics. 2020;17:21.
41. National Research Council (US). Defining the Mandate of Proteomics in the Post-Genomics Era: Workshop Report: National Academies Press Washington (DC); 2002: 4.
42. Cox J, Mann M. Quantitative, high-resolution proteomics for data-driven systems biology. Annu Rev Biochem. 2011;80:273-99.
43. Guest PC, Guest FL, Martins-de Souza D. Making Sense of Blood-Based Proteomics and Metabolomics in Psychiatric Research. Int J Neuropsychopharmacol. 2016;19(6).
44. Jankovska E, Svitek M, Holada K, Petrak J. Affinity depletion versus relative protein enrichment: a side-by-side comparison of two major strategies for increasing human cerebrospinal fluid proteome coverage. Clin Proteomics. 2019;16:9.
45. Yuan X, Desiderio DM. Proteomics analysis of human cerebrospinal fluid. J Chromatogr B Analyt Technol Biomed Life Sci. 2005;815(1-2):179-89.
46. Hühmer AF, Biringer RG, Amato H, Fonteh AN, Harrington MG. Protein analysis in human cerebrospinal fluid: Physiological aspects, current progress and future challenges. Dis Markers. 2006;22(1-2):3-26.
47. Zhang Y, Guo Z, Zou L, Yang Y, Zhang L, Ji N, et al. A comprehensive map and functional annotation of the normal human cerebrospinal fluid proteome. J Proteomics. 2015;119:90-9.
48. Teunissen CE, Tumani H, Engelborghs S, Mollenhauer B. Biobanking of CSF: international standardization to optimize biomarker development. Clin Biochem. 2014;47(4-5):288-92.
49. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, et al. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science. 2001;292(5518):929-34.
50. Freeman WM, Hemby SE. Proteomics for protein expression profiling in neuroscience. Neurochem Res. 2004;29(6):1065-81.
51. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003;422(6928):198-207.
52. Shores KS, Knapp DR. Assessment approach for evaluating high abundance protein depletion methods for cerebrospinal fluid (CSF) proteomic analysis. J Proteome Res. 2007;6(9):3739-51.
53. Yuan X, Desiderio DM. Proteomics analysis of prefractionated human lumbar cerebrospinal fluid. Proteomics. 2005;5(2):541-50.
54. Gallagher SR. One-dimensional SDS gel electrophoresis of proteins. Curr Protoc Mol Biol. 2006;Chapter 10:Unit 10.2A: 2.
55. Smith BJ. SDS Polyacrylamide Gel Electrophoresis of Proteins. Methods in Molecular Biology, Volume 1 Proteins: Humana Press; 1984: 41-43.
56. Olsen JV, Ong SE, Mann M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Mol Cell Proteomics. 2004;3(6):608-14.
57. Romijn EP, Krijgsveld J, Heck AJ. Recent liquid chromatographic-(tandem) mass spectrometric applications in proteomics. J Chromatogr A. 2003;1000(1-2):589-608.
58. Lindsay S. High Performance Liquid Chromatography: John Wiley & Sons Inc; 1992: 1-3,28.
59. Perumal N. Characterization of human tear proteome in dry eye syndrome [Thesis]. Mainz: Johannes Gutenberg-Universität Mainz; 2015.
60. Ong SE, Foster LJ, Mann M. Mass spectrometric-based approaches in quantitative proteomics. Methods. 2003;29(2):124-30.
61. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. Science. 1989;246(4926):64-71.
62. Gross J. Massenspektrometrie: Ein Lehrbuch. Heidelberg: Springer-Verlag; 2011: 209-215.

63. Spahr CS, Susin SA, Bures EJ, Robinson JH, Davis MT, McGinley MD, et al. Simplification of complex peptide mixtures for proteomic analysis: reversible biotinylation of cysteinyl peptides. Electrophoresis. 2000;21(9):1635-50.
64. Ma S, Chowdhury SK, Alton KB. Application of mass spectrometry for metabolite identification. Curr Drug Metab. 2006;7(5):503-23.
65. Choudhary J, Grant SG. Proteomics in postgenomic neuroscience: the end of the beginning. Nat Neurosci. 2004;7(5):440-5.
66. Nahnsen S, Bielow C, Reinert K, Kohlbacher O. Tools for label-free peptide quantification. Mol Cell Proteomics. 2013;12(3):549-56.
67. Marx V. Targeted proteomics. Nat Methods. 2013;10(1):19-22.
68. Borras E, Sabido E. What is targeted proteomics? A concise revision of targeted acquisition and targeted data analysis in mass spectrometry. Proteomics. 2017;17(17-18).
69. Wan C, Yang Y, Li H, La Y, Zhu H, Jiang L, et al. Dysregulation of retinoid transporters expression in body fluids of schizophrenia patients. J Proteome Res. 2006;5(11):3213-6.
70. Manes NP, Nita-Lazar A. Application of targeted mass spectrometry in bottom-up proteomics for systems biology research. J Proteomics. 2018;189:75-90.
71. Aebersold R, Bensimon A, Collins BC, Ludwig C, Sabido E. Applications and Developments in Targeted Proteomics: From SRM to DIA/SWATH. Proteomics. 2016;16(15-16):2065-7.
72. Zhang Z, Wu S, Stenoien DL, Paša-Tolić L. High-throughput proteomics. Annu Rev Anal Chem (Palo Alto Calif). 2014;7:427-54.
73. World Health Organization. Mental disorders [Internet]. Geneva: WHO; 2019 [cited 2020 July 12]. Available from: https://www.who.int/news-room/fact-sheets/detail/mental-disorders.
74. Giusti L, Ciregia F, Mazzoni MR, Lucacchini A. Proteomics insight into psychiatric disorders: an update on biological fluid biomarkers. Expert Rev Proteomics. 2016;13(10):941-50.
75. Europe WHO. Mental health [Internet]. Copenhagen: WHO Regional Office for Europe; 2018 [cited 2020 July 12]. Available from: https://www.euro.who.int/en/health-topics/noncommunicable-diseases/mental-health/mental-health.
76. Comes AL, Papiol S, Mueller T, Geyer PE, Mann M, Schulze TG. Proteomics for blood biomarker exploration of severe mental illness: pitfalls of the past and potential for the future. Transl Psychiatry. 2018;8(1):160.
77. Macron C, Lane L, Nunez Galindo A, Dayon L. Identification of Missing Proteins in Normal Human Cerebrospinal Fluid. J Proteome Res. 2018;17(12):4315-9.
78. Huang JT, Wang L, Prabakaran S, Wengenroth M, Lockstone HE, Koethe D, et al. Independent protein-profiling studies show a decrease in apolipoprotein A1 levels in schizophrenia CSF, brain and peripheral tissues. Mol Psychiatry. 2008;13(12):1118-28.
79. Amboss. Schizophrenie [Internet]. Berlin: Amboss; 2019, [cited 2020 May 04]. Available from: https://www.amboss.com/de/wissen/Schizophrenie.
80. National Institute of Mental Health. Schizophrenia [Internet]. Bethesda: NIMH; 2016 [cited 2020 January 04]. Available from: https://www.nimh.nih.gov/health/topics/schizophrenia/index.shtml.
81. Owen MJ, Williams NM, O'Donovan MC. The molecular genetics of schizophrenia: new findings promise new insights. Mol Psychiatry. 2004;9(1):14-27.
82. Holmes E, Tsang TM, Huang JT, Leweke FM, Koethe D, Gerth CW, et al. Metabolic profiling of CSF: evidence that early intervention may impact on disease progression and outcome in schizophrenia. PLoS Med. 2006;3(8):e327.
83. Martins-De-Souza D, Wobrock T, Zerr I, Schmitt A, Gawinecka J, Schneider-Axmann T, et al. Different apolipoprotein E, apolipoprotein A1 and prostaglandin-H2 D-isomerase levels in cerebrospinal fluid of schizophrenia patients and healthy controls. World J Biol Psychiatry. 2010;11(5):719-28.
84. Semmes OJ, Feng Z, Adam BL, Banez LL, Bigbee WL, Campos D, et al. Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass

spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. Clin Chem. 2005;51(1):102-12.
85. Huang JT, Leweke FM, Tsang TM, Koethe D, Kranaster L, Gerth CW, et al. CSF metabolic and proteomic profiles in patients prodromal for psychosis. PLoS One. 2007;2(8):e756.
86. Albertini V, Benussi L, Paterlini A, Glionna M, Prestia A, Bocchio-Chiavetto L, et al. Distinct cerebrospinal fluid amyloid-beta peptide signatures in cognitive decline associated with Alzheimer's disease and schizophrenia. Electrophoresis. 2012;33(24):3738-44.
87. Johansson V, Jakobsson J, Fortgang RG, Zetterberg H, Blennow K, Cannon TD, et al. Cerebrospinal fluid microglia and neurodegenerative markers in twins concordant and discordant for psychotic disorders. Eur Arch Psychiatry Clin Neurosci. 2017;267(5):391-402.
88. Gupta AK, Pokhriyal R, Khan MI, Kumar DR, Gupta R, Chadda RK, et al. Cerebrospinal Fluid Proteomics For Identification Of alpha2-Macroglobulin As A Potential Biomarker To Monitor Pharmacological Therapeutic Efficacy In Dopamine Dictated Disease States Of Parkinson's Disease And Schizophrenia. Neuropsychiatr Dis Treat. 2019;15:2853-67.
89. Amboss. Depression [Internet]. Berlin: Amboss; 2020 [cited 2020 May 04]. Available from: https://www.amboss.com/de/wissen/Depression.
90. World Health Organization. Depression Key facts [Internet]. Geneva: WHO; 2020 [cited 2020 February 12]. Available from: https://www.who.int/news-room/fact-sheets/detail/depression.
91. Greenberg PE, Fournier A-A, Sisitsky T, Simes M, Berman R, Koenigsberg SH, et al. The Economic Burden of Adults with Major Depressive Disorder in the United States (2010 and 2018). Pharmacoeconomics. 2021;39(6):653-65.
92. Al Shweiki MR, Oeckl P, Steinacker P, Hengerer B, Schonfeldt-Lecuona C, Otto M. Major depressive disorder: insight into candidate cerebrospinal fluid protein biomarkers from proteomics studies. Expert Rev Proteomics. 2017;14(6):499-514.
93. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry. 2006;163(11):1905-17.
94. Ditzen C, Tang N, Jastorff AM, Teplytska L, Yassouridis A, Maccarrone G, et al. Cerebrospinal fluid biomarkers for major depression confirm relevance of associated pathophysiology. Neuropsychopharmacology. 2012;37(4):1013-25.
95. Belmaker RH, Agam G. Major depressive disorder. N Engl J Med. 2008;358(1):55-68.
96. Diniz BS, Teixeira AL, Machado-Vieira R, Talib LL, Radanovic M, Gattaz WF, et al. Reduced cerebrospinal fluid levels of brain-derived neurotrophic factor is associated with cognitive impairment in late-life major depression. J Gerontol B Psychol Sci Soc Sci. 2014;69(6):845-51.
97. Kern S, Skoog I, Borjesson-Hanson A, Blennow K, Zetterberg H, Ostling S, et al. Higher CSF interleukin-6 and CSF interleukin-8 in current depression in older women. Results from a population-based sample. Brain Behav Immun. 2014;41:55-8.
98. Gudmundsson P, Skoog I, Waern M, Blennow K, Zetterberg H, Rosengren L, et al. Is there a CSF biomarker profile related to depression in elderly women? Psychiatry Res. 2010;176(2-3):174-8.
99. Pomara N, Bruno D, Sarreal AS, Hernando RT, Nierenberg J, Petkova E, et al. Lower CSF amyloid beta peptides and higher F2-isoprostanes in cognitively intact elderly individuals with major depressive disorder. Am J Psychiatry. 2012;169(5):523-30.
100. Sakamoto S, Putalun W, Vimolmangkang S, Phoolcharoen W, Shoyama Y, Tanaka H, et al. Enzyme-linked immunosorbent assay for the quantitative/qualitative analysis of plant secondary metabolites. J Nat Med. 2018;72(1):32-42.
101. Grabe HJ, Ahrens N, Rose HJ, Kessler C, Freyberger HJ. Neurotrophic factor S100 beta in major depression. Neuropsychobiology. 2001;44(2):88-90.
102. Yoon HS, Hattori K, Sasayama D, Kunugi H. Low cocaine- and amphetamine-regulated transcript (CART) peptide levels in human cerebrospinal fluid of major depressive disorder (MDD) patients. J Affect Disord. 2018;232:134-8.

103. Maccarrone G, Ditzen C, Yassouridis A, Rewerts C, Uhr M, Uhlen M, et al. Psychiatric patient stratification using biosignatures based on cerebrospinal fluid protein expression clusters. J Psychiatr Res. 2013;47(11):1572-80.
104. Schonknecht P, Pantel J, Kaiser E, Thomann P, Schroder J. Increased tau protein differentiates mild cognitive impairment from geriatric depression and predicts conversion to dementia. Neurosci Lett. 2007;416(1):39-42.
105. Reis T, Brandao CO, Freire Coutinho ES, Engelhardt E, Laks J. Cerebrospinal fluid biomarkers in Alzheimer's disease and geriatric depression: preliminary findings from Brazil. CNS Neurosci Ther. 2012;18(7):524-9.
106. Gupta AK, Pokhriyal R, Das U, Khan MI, Ratna Kumar D, Gupta R, et al. Evaluation of α-synuclein and apolipoprotein E as potential biomarkers in cerebrospinal fluid to monitor pharmacotherapeutic efficacy in dopamine dictated disease states of Parkinson's disease and schizophrenia. Neuropsychiatr Dis Treat. 2019;15:2073-85.
107. Brunner J, Bronisch T, Uhr M, Ising M, Binder E, Holsboer F, et al. Proteomic analysis of the CSF in unmedicated patients with major depressive disorder reveals alterations in suicide attempters. Eur Arch Psychiatry Clin Neurosci. 2005;255(6):438-40.
108. Deuschle M, Schilling C, Leweke FM, Enning F, Pollmächer T, Esselmann H, et al. Hypocretin in cerebrospinal fluid is positively correlated with Tau and pTau. Neurosci Lett. 2014;561:41-5.
109. Kern S, Skoog I, Börjesson-Hanson A, Blennow K, Zetterberg H, Ostling S, et al. Higher CSF interleukin-6 and CSF interleukin-8 in current depression in older women. Results from a population-based sample. Brain Behav Immun. 2014;41:55-8.
110. Wu L, Han DK. Overcoming the dynamic range problem in mass spectrometry-based shotgun proteomics. Expert Rev Proteomics. 2006;3(6):611-9.
111. Michalski A. Overcoming Challenges of Shotgun Proteomics [Thesis]. München: Ludwig-Maximilians-Universität München; 2012.
112. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics. 2002;1(11):845-67.
113. Thermo Fisher Scientific. Overview of Protein Electrophoresis [Internet]. Waltham: Thermo Fisher Scientific; 2021 [cited 2022 May 05]. Available from: https://www.thermofisher.com/de/de/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-electrophoresis.html.
114. Schagger H. Tricine-SDS-PAGE. Nat Protoc. 2006;1(1):16-22.
115. Perumal N, Straßburger L, Schmelter C, Gericke A, Pfeiffer N, Grus FH, et al. Sample Preparation for Mass-spectrometry-based Proteomics Analysis of Ocular Microvessels. J Vis Exp. 2019(144).
116. Herzog DP, Perumal N, Manicam C, Treccani G, Nadig J, Rossmanith M, et al. Longitudinal CSF proteome profiling in mice to uncover the acute and sustained mechanisms of action of rapid acting antidepressant (2R,6R)-hydroxynorketamine (HNK). Neurobiol Stress. 2021;15:100404.
117. Wong SS. Chemistry of protein conjugation and cross-linking: CRC press; 1991: 10, 27.
118. Millipore Corporation. User Guide for Reversed-Phase ZipTip [Internet]. Darmstadt: Merck Millipore Corporation; 2005 [cited 2021 December 12]. Available from: http://personal.rhul.ac.uk/upba/211/Zip-tip.pdf.
119. Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, Pesch R, et al. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. Mol Cell Proteomics. 2005;4(12):2012.
120. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008;26(12):1367-72.
121. Krämer A, Green J, Pollard Jr J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. Bioinformatics. 2013;30(4):523-30.

122. Tschäbunin A. Identification of clinical biomarkers in dry eye disease using a targeted proteomics approach [Thesis]. Mainz: Johannes Gutenberg-Universität Mainz; 2020.
123. Mahley RW. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. Science. 1988;240(4852):622-30.
124. Gibbons AS, Udawela M, Jeon WJ, Seo MS, Brooks L, Dean B. The neurobiology of APOE in schizophrenia and mood disorders. Front Biosci (Landmark Ed). 2011;16:962-79.
125. Das HK, McPherson J, Bruns GA, Karathanasis SK, Breslow JL. Isolation, characterization, and mapping to chromosome 19 of the human apolipoprotein E gene. J Biol Chem. 1985;260(10):6240-7.
126. Artiga MJ, Bullido MJ, Sastre I, Recuero M, García MA, Aldudo J, et al. Allelic polymorphisms in the transcriptional regulatory region of apolipoprotein E gene. FEBS Lett. 1998;421(2):105-8.
127. Al-Asmary SM, Kadasah S, Arfin M, Tariq M, Al-Asmari A. Apolipoprotein E polymorphism is associated with susceptibility to schizophrenia among Saudis. Arch Med Sci. 2015;11(4):869-76.
128. Akanji AO, Ohaeri JU, Al-Shammri SN, Fatania HR. Apolipoprotein E polymorphism and clinical disease phenotypes in Arab patients with schizophrenia. Neuropsychobiology. 2009;60(2):67-72.
129. Farrer LA, Mitchell F, Brin MF, Elsas L. Statement on use of apolipoprotein E testing for Alzheimer disease. American College of Medical Genetics/American Society of Human Genetics Working Group on ApoE and Alzheimer disease. JAMA. 1995;274(20):1627-9.
130. Carlin C, Murray L, Graham D, Doyle D, Nicoll J. Involvement of apolipoprotein E in multiple sclerosis: absence of remyelination associated with possession of the APOE epsilon2 allele. J Neuropathol Exp Neurol. 2000;59(5):361-7.
131. Boyles JK, Zoellner CD, Anderson LJ, Kosik LM, Pitas RE, Weisgraber KH, et al. A role for apolipoprotein E, apolipoprotein A-I, and low density lipoprotein receptors in cholesterol transport during regeneration and remyelination of the rat sciatic nerve. J Clin Invest. 1989;83(3):1015-31.
132. de Chaves EP, Narayanaswami V. Apolipoprotein E and cholesterol in aging and disease in the brain. Future Lipidol. 2008;3(5):505-30.
133. Jonas K, Clouston S, Li K, Fochtmann LJ, Lencz T, Malhotra AK, et al. Apolipoprotein E-ε4 allele predicts escalation of psychotic symptoms in late adulthood. Schizophr Res. 2019;206:82-8.
134. Gupta AK, Rani K, Swarnkar S, Kumar GK, Khan MI, Pokhriyal R, et al. Evaluation of Serum Apolipoprotein E as a Potential Biomarker for Pharmacological Therapeutic Efficacy Monitoring in Dopamine Dictated Disease Spectrum of Schizophrenia and Parkinson's disease: A Preliminary Study. J Cent Nerv Syst Dis. 2018;10:1179573518803585.
135. Kimura T, Yokota S, Igata-Yi R, Shono M, Takamatsu J, Miyakawa T. Apolipoprotein E epsilon2 allele and early onset schizophrenia. Neurosci Lett. 1997;231(1):53-5.
136. Ban C, Zhang Q, Feng J, Li H, Qiu Q, Tian Y, et al. Low prevalence of lipid metabolism abnormalities in APOE ε2-genotype and male patients 60 years or older with schizophrenia. BMC Psychiatry. 2017;17(1):399.
137. Kecmanović M, Dobrićić V, Dimitrijević R, Keckarević D, Savić-Pavićević D, Keckarević-Marković M, et al. Schizophrenia and apolipoprotein E gene polymorphism in Serbian population. Int J Neurosci. 2010;120(7):502-6.
138. Wen H, Liu M, Liu Z, Yang X, Liu X, Ni M, et al. PEDF improves atherosclerotic plaque stability by inhibiting macrophage inflammation response. Int J Cardiol. 2017;235:37-41.
139. Ryan KM, Glaviano A, O'Donovan SM, Kolshus E, Dunne R, Kavanagh A, et al. Electroconvulsive therapy modulates plasma pigment epithelium-derived factor in depression: a proteomics study. Transl Psychiatry. 2017;7(3):e1073.
140. Tombran-Tink J, Chader GG, Johnson LV. PEDF: a pigment epithelium-derived factor with potent neuronal differentiative activity. Exp Eye Res. 1991;53(3):411-4.
141. Broadhead ML, Becerra SP, Choong PF, Dass CR. The applied biochemistry of PEDF and implications for tissue homeostasis. Growth Factors. 2010;28(4):280-5.

142. Tian T, Yang Y, Xu B, Qin Y, Zang G, Zhou C, et al. Pigment epithelium-derived factor alleviates depressive-like behaviors in mice by modulating adult hippocampal synaptic growth and Wnt pathway. Prog Neuropsychopharmacol Biol Psychiatry. 2020;98:109792.
143. Kiecolt-Glaser JK, Derry HM, Fagundes CP. Inflammation: depression fans the flames and feasts on the heat. Am J Psychiatry. 2015;172(11):1075-91.
144. Zhu S, He M, Liu Z, Qin Z, Wang Z, Duan L. Shared genetic susceptibilities for irritable bowel syndrome and depressive disorder in Chinese patients uncovered by pooled whole-exome sequencing. J Adv Res. 2020;23:113-21.
145. Cheng JS, Dubal DB, Kim DH, Legleiter J, Cheng IH, Yu GQ, et al. Collagen VI protects neurons against Abeta toxicity. Nat Neurosci. 2009;12(2):119-21.
146. Grozdanov V, Bliederhaeuser C, Ruf WP, Roth V, Fundel-Clemens K, Zondler L, et al. Inflammatory dysregulation of blood monocytes in Parkinson's disease patients. Acta Neuropathol. 2014;128(5):651-63.
147. Kangussu LM, Almeida-Santos AF, Bader M, Alenina N, Fontes MA, Santos RA, et al. Angiotensin-(1-7) attenuates the anxiety and depression-like behaviors in transgenic rats with low brain angiotensinogen. Behav Brain Res. 2013;257:25-30.
148. Costa AC, Becker LK, Moraes ER, Romero TR, Guzzo L, Santos RA, et al. Angiotensin-(1-7) induces peripheral antinociception through mas receptor activation in an opioid-independent pathway. Pharmacology. 2012;89(3-4):137-44.
149. Lazaroni TL, Raslan AC, Fontes WR, de Oliveira ML, Bader M, Alenina N, et al. Angiotensin-(1-7)/Mas axis integrity is required for the expression of object recognition memory. Neurobiol Learn Mem. 2012;97(1):113-23.
150. Voigt JP, Hörtnagl H, Rex A, van Hove L, Bader M, Fink H. Brain angiotensin and anxiety-related behavior: the transgenic rat TGR(ASrAOGEN)680. Brain Res. 2005;1046(1-2):145-56.
151. Xing J, Kong J, Lu J, Li J. Angiotensin-(1-7) inhibits neuronal activity of dorsolateral periaqueductal gray via a nitric oxide pathway. Neurosci Lett. 2012;522(2):156-61.
152. UniProt. UniProtKB - P05067 (A4_HUMAN) [Internet]. Hinxton: European Bioinformatics Institut; 2020 [cited 2020 June 08]. Available from: https://www.uniprot.org/uniprot/P05067.
153. Baumkötter F, Schmidt N, Vargas C, Schilling S, Weber R, Wagner K, et al. Amyloid precursor protein dimerization and synaptogenic function depend on copper binding to the growth factor-like domain. J Neurosci. 2014;34(33):11159-72.
154. Turner PR, O'Connor K, Tate WP, Abraham WC. Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory. Prog Neurobiol. 2003;70(1):1-32.
155. Schmechel DE, Goldgaber D, Burkhart DS, Gilbert JR, Gajdusek DC, Roses AD. Cellular localization of messenger RNA encoding amyloid-beta-protein in normal tissue and in Alzheimer disease. Alzheimer Dis Assoc Disord. 1988;2(2):96-111.
156. Tanaka S, Shiojiri S, Takahashi Y, Kitaguchi N, Ito H, Kameyama M, et al. Tissue-specific expression of three types of beta-protein precursor mRNA: enhancement of protease inhibitor-harboring types in Alzheimer's disease brain. Biochem Biophys Res Commun. 1989;165(3):1406-14.
157. Miyamae Y, Nakamura Y, Kashiwagi Y, Tanaka T, Kudo T, Takeda M. Altered adhesion efficiency and fibronectin content in fibroblasts from schizophrenic patients. Psychiatry Clin Neurosci. 1998;52(3):345-52.
158. UniProt. UniProtKB - P02751 (FINC_HUMAN) [Internet]. Hinxton: European Bioinformatics Institut; 2020 [cited 2020 June 17]. Available from: https://www.uniprot.org/uniprot/P02751.
159. Owens RJ, Baralle FE. Mapping the collagen-binding site of human fibronectin by expression in Escherichia coli. EMBO J. 1986;5(11):2825-30.
160. Calaycay J, Pande H, Lee T, Borsi L, Siri A, Shively JE, et al. Primary structure of a DNA- and heparin-binding domain (Domain III) in human plasma fibronectin. J Biol Chem. 1985;260(22):12136-41.

161. Garcia-Pardo A, Rostagno A, Frangione B. Primary structure of human plasma fibronectin. Characterization of a 38 kDa domain containing the C-terminal heparin-binding site (Hep III site) and a region of molecular heterogeneity. Biochem J. 1987;241(3):923-8.
162. Rostagno A, Williams MJ, Baron M, Campbell ID, Gold LI. Further characterization of the NH2-terminal fibrin-binding site on fibronectin. J Biol Chem. 1994;269(50):31938-45.
163. Nakata K, Ujike H, Sakai A, Takaki M, Imamura T, Tanaka Y, et al. Association study between the fibronectin gene and schizophrenia. Am J Med Genet B Neuropsychiatr Genet. 2003;116b(1):41-4.
164. Figueiro-Silva J, Gruart A, Clayton KB, Podlesniy P, Abad MA, Gasull X, et al. Neuronal pentraxin 1 negatively regulates excitatory synapse density and synaptic plasticity. J Neurosci. 2015;35(14):5504-21.
165. Ma QL, Teng E, Zuo X, Jones M, Teter B, Zhao EY, et al. Neuronal pentraxin 1: A synaptic-derived plasma biomarker in Alzheimer's disease. Neurobiol Dis. 2018;114:120-8.
166. Guo H, Bettella E, Marcogliese PC, Zhao R, Andrews JC, Nowakowski TJ, et al. Disruptive mutations in TANC2 define a neurodevelopmental syndrome associated with psychiatric disorders. Nat Commun. 2019;10(1):4679.
167. Han S, Nam J, Li Y, Kim S, Cho SH, Cho YS, et al. Regulation of dendritic spines, spatial memory, and embryonic development by the TANC family of PSD-95-interacting proteins. J Neurosci. 2010;30(45):15102-12.
168. Gasparini A, Tosatto SCE, Murgia A, Leonardi E. Dynamic scaffolds for neuronal signaling: in silico analysis of the TANC protein family. Sci Rep. 2017;7(1):6829.
169. Stucchi R, Plucińska G, Hummel JJA, Zahavi EE, Guerra San Juan I, Klykov O, et al. Regulation of KIF1A-Driven Dense Core Vesicle Transport: Ca(2+)/CaM Controls DCV Binding and Liprin-α/TANC2 Recruits DCVs to Postsynaptic Sites. Cell Rep. 2018;24(3):685-700.
170. UniProt. UniProtKB - Q7Z7M0 (MEGF8_HUMAN) [Internet]. Hinxton: European Bioinformatics Institut; 2020 [cited 2020 June 17]. Available from: https://www.uniprot.org/uniprot/Q7Z7M0.
171. Pusapati GV, Kong JH, Patel BB, Krishnan A, Sagner A, Kinnebrew M, et al. CRISPR Screens Uncover Genes that Regulate Target Cell Sensitivity to the Morphogen Sonic Hedgehog. Dev Cell. 2018;44(1):113-29.e8.
172. Lööv C, Hillered L, Ebendal T, Erlandsson A. Engulfing astrocytes protect neurons from contact-induced apoptosis following injury. PLoS One. 2012;7(3):e33090.
173. Pan C, Zhou Y, Dator R, Ginghina C, Zhao Y, Movius J, et al. Targeted discovery and validation of plasma biomarkers of Parkinson's disease. J Proteome Res. 2014;13(11):4535-45.
174. Birtwistle J, Baldwin D. Role of dopamine in schizophrenia and Parkinson's disease. Br J Nurs. 1998;7(14):832-4, 6, 8-41.
175. UniProt. UniProtKB - Q9Y4C0 (NRX3A_HUMAN) [Internet]. Hinxton: European Bioinformatics Institut; 2020 [cited 2020 June 19]. Available from: https://www.uniprot.org/uniprot/Q9Y4C0.
176. Südhof TC. Neuroligins and neurexins link synaptic function to cognitive disease. Nature. 2008;455(7215):903-11.
177. Ushkaryov YA, Petrenko AG, Geppert M, Südhof TC. Neurexins: synaptic cell surface proteins related to the alpha-latrotoxin receptor and laminin. Science. 1992;257(5066):50-6.
178. Sasaki J, Kumagae G, Sata T, Kuramitsu M, Arakawa K. Decreased concentration of high density lipoprotein cholesterol in schizophrenic patients treated with phenothiazines. Atherosclerosis. 1984;51(2-3):163-9.
179. Boiko AS, Mednova IA, Kornetova EG, Semke AV, Bokhan NA, Loonen AJM, et al. Apolipoprotein serum levels related to metabolic syndrome in patients with schizophrenia. Heliyon. 2019;5(7):e02033.
180. Scheltens P, Blennow K, Breteler MM, de Strooper B, Frisoni GB, Salloway S, et al. Alzheimer's disease. Lancet. 2016;388(10043):505-17.
181. Gandal MJ, Haney JR, Parikshak NN, Leppa V, Ramaswami G, Hartl C, et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. Science. 2018;359(6376):693-7.

182. Cray C, Zaias J, Altman NH. Acute phase response in animals: a review. Comp Med. 2009;59(6):517-26.
183. Murata H, Shimada N, Yoshioka M. Current research on acute phase proteins in veterinary diagnosis: an overview. Vet J. 2004;168(1):28-40.
184. Müller N. Inflammation in Schizophrenia: Pathogenetic Aspects and Therapeutic Considerations. Schizophr Bull. 2018;44(5):973-82.
185. Yang L, Zhao Y, Wang Y, Liu L, Zhang X, Li B, et al. The Effects of Psychological Stress on Depression. Curr Neuropharmacol. 2015;13(4):494-504.
186. Park C, Rosenblat JD, Brietzke E, Pan Z, Lee Y, Cao B, et al. Stress, epigenetics and depression: A systematic review. Neurosci Biobehav Rev. 2019;102:139-52.
187. Pariante CM, Lightman SL. The HPA axis in major depression: classical theories and new developments. Trends Neurosci. 2008;31(9):464-8.
188. Holsboer F. The corticosteroid receptor hypothesis of depression. Neuropsychopharmacology. 2000;23(5):477-501.
189. Ishii T, Hattori K, Miyakawa T, Watanabe K, Hidese S, Sasayama D, et al. Increased cerebrospinal fluid complement C5 levels in major depressive disorder and schizophrenia. Biochem Biophys Res Commun. 2018;497(2):683-8.
190. Liu Y, Ho RC, Mak A. Interleukin (IL)-6, tumour necrosis factor alpha (TNF-α) and soluble interleukin-2 receptors (sIL-2R) are elevated in patients with major depressive disorder: a meta-analysis and meta-regression. J Affect Disord. 2012;139(3):230-9.
191. Zhang JC, Yao W, Dong C, Yang C, Ren Q, Ma M, et al. Blockade of interleukin-6 receptor in the periphery promotes rapid and sustained antidepressant actions: a possible role of gut-microbiota-brain axis. Transl Psychiatry. 2017;7(5):e1138.
192. Kirkegaard C, Faber J. The role of thyroid hormones in depression. Eur J Endocrinol. 1998;138(1):1-9.
193. Baumgartner A, Gräf KJ, Kürten I, Meinhold H. The hypothalamic-pituitary-thyroid axis in psychiatric patients and healthy subjects: Parts 1-4. Psychiatry Res. 1988;24(3):271-332.
194. Canuso CM, Pandina G. Gender and schizophrenia. Psychopharmacol Bull. 2007;40(4):178-90.
195. Tamminga CA. Gender and schizophrenia. J Clin Psychiatry. 1997;58 Suppl 15:33-7.
196. Seeman MV. Does Gender Influence Outcome in Schizophrenia? Psychiatr Q. 2019;90(1):173-84.
197. Smith S. Gender differences in antipsychotic prescribing. Int Rev Psychiatry. 2010;22(5):472-84.
198. Wils RS, Gotfredsen DR, Hjorthøj C, Austin SF, Albert N, Secher RG, et al. Antipsychotic medication and remission of psychotic symptoms 10years after a first-episode psychosis. Schizophr Res. 2017;182:42-8.
199. de Bie J, Lim CK, Guillemin GJ. Kynurenines, Gender and Neuroinflammation; Showcase Schizophrenia. Neurotox Res. 2016;30(3):285-94.
200. da Silva TL, Ravindran AV. Contribution of sex hormones to gender differences in schizophrenia: A review. Asian J Psychiatr. 2015;18:2-14.
201. Eltokhi A, Rappold G, Sprengel R. Distinct Phenotypes of Shank2 Mouse Models Reflect Neuropsychiatric Spectrum Disorders of Human Patients With SHANK2 Variants. Front Mol Neurosci. 2018;11:240.
202. Peykov S, Berkel S, Schoen M, Weiss K, Degenhardt F, Strohmaier J, et al. Identification and functional characterization of rare SHANK2 variants in schizophrenia. Mol Psychiatry. 2015;20(12):1489-98.
203. Berkel S, Eltokhi A, Fröhlich H, Porras-Gonzalez D, Rafiullah R, Sprengel R, et al. Sex Hormones Regulate SHANK Expression. Front Mol Neurosci. 2018;11:337.
204. Baba A, Fujita T, Tamura N. Sexual dimorphism of the fifth component of mouse complement. J Exp Med. 1984;160(2):411-9.
205. Contractor T, Kobayashi S, da Silva E, Clausen R, Chan C, Vosburgh E, et al. Sexual dimorphism of liver metastasis by murine pancreatic neuroendocrine tumors is affected by expression of complement C5. Oncotarget. 2016;7(21):30585-96.

206. UniProtKB. UniProtKB - P01876 (IGHA1_HUMAN) [Internet]. Hinxton: European Bioinformatics Institut; 2021 [cited 2021 April 10]. Available from: https://www.uniprot.org/uniprot/P01876.
207. McHeyzer-Williams M, Okitsu S, Wang N, McHeyzer-Williams L. Molecular programming of B cell memory. Nat Rev Immunol. 2011;12(1):24-34.
208. Schroeder HW, Jr, Cavacini L. Structure and function of immunoglobulins. J Allergy Clin Immunol. 2010;125(2 Suppl 2):S41-52.
209. UniProtKB. UniProtKB - Q16270 (IBP7_HUMAN) [Internet]. Hinxton: European Bioinformatics Institut; 2021 [cited 2021, April 10]. Available from: https://www.uniprot.org/uniprot/Q16270.
210. Akaogi K, Okabe Y, Funahashi K, Yoshitake Y, Nishikawa K, Yasumitsu H, et al. Cell adhesion activity of a 30-kDa major secreted protein from human bladder carcinoma cells. Biochem Biophys Res Commun. 1994;198(3):1046-53.
211. Oh Y, Nagalla SR, Yamanaka Y, Kim HS, Wilson E, Rosenfeld RG. Synthesis and characterization of insulin-like growth factor-binding protein (IGFBP)-7. Recombinant human mac25 protein specifically binds IGF-I and -II. J Biol Chem. 1996;271(48):30322-5.
212. Jin L, Shen F, Weinfeld M, Sergi C. Insulin Growth Factor Binding Protein 7 (IGFBP7)-Related Cancer and IGFBP3 and IGFBP7 Crosstalk. Front Oncol. 2020;10:727.
213. Hwa V, Tomasini-Sprenger C, Bermejo AL, Rosenfeld RG, Plymate SR. Characterization of insulin-like growth factor-binding protein-related protein-1 in prostate cells. J Clin Endocrinol Metab. 1998;83(12):4355-62.
214. Tomimaru Y, Eguchi H, Wada H, Noda T, Murakami M, Kobayashi S, et al. Insulin-like growth factor-binding protein 7 alters the sensitivity to interferon-based anticancer therapy in hepatocellular carcinoma cells. Br J Cancer. 2010;102(10):1483-90.
215. Ahmed S, Jin X, Yagi M, Yasuda C, Sato Y, Higashi S, et al. Identification of membrane-bound serine proteinase matriptase as processing enzyme of insulin-like growth factor binding protein-related protein-1 (IGFBP-rP1/angiomodulin/mac25). The FEBS journal. 2006;273(3):615-27.
216. Boddu R, Fan C, Rangarajan S, Sunil B, Bolisetty S, Curtis LM. Unique sex- and age-dependent effects in protective pathways in acute kidney injury. Am J Physiol Renal Physiol. 2017;313(3):F740-f55.
217. Yang BZ, Han S, Kranzler HR, Palmer AA, Gelernter J. Sex-specific linkage scans in opioid dependence. Am J Med Genet B Neuropsychiatr Genet. 2017;174(3):261-8.
218. Basu MK, Massicano F, Yu L, Halkidis K, Pillai V, Cao W, et al. Exome Sequencing Identifies Abnormalities in Glycosylation and ANKRD36C in Patients with Immune-Mediated Thrombotic Thrombocytopenic Purpura. Thromb Haemost. 2021;121(4):506-17.
219. Fang Y, Wang X, Li W, Han J, Jin J, Su F, et al. Screening of circular RNAs and validation of circANKRD36 associated with inflammation in patients with type 2 diabetes mellitus. Int J Mol Med. 2018;42(4):1865-74.
220. Kumar R, Haugen JD, Wieben ED, Londowski JM, Cai Q. Inhibitors of renal epithelial phosphate transport in tumor-induced osteomalacia and uremia. Proc Assoc Am Physicians. 1995;107(3):296-305.
221. UniProt. UniProtKB - Q5JPF3 (AN36C_HUMAN) [Internet]. Hinxton: European Bioinformatics Institut; 2022 [cited 2022 February 23]. Available from: https://www.uniprot.org/uniprot/Q5JPF3.
222. Koths K, Taylor E, Halenbeck R, Casipit C, Wang A. Cloning and characterization of a human Mac-2-binding protein, a new member of the superfamily defined by the macrophage scavenger receptor cysteine-rich domain. J Biol Chem. 1993;268(19):14245-9.
223. UniProtKB. UniProtKB - Q08380 (LG3BP_HUMAN) [Internet]. Hinxton: European Bioinformatics Institut; 2021 [cited 2021 April 10]. Available from: https://www.uniprot.org/uniprot/Q08380.
224. Ullrich A, Sures I, D'Egidio M, Jallal B, Powell TJ, Herbst R, et al. The secreted tumor-associated antigen 90K is a potent immune stimulator. J Biol Chem. 1994;269(28):18401-7.

225. Melin EO, Dereke J, Hillman M. Female sex, high soluble CD163, and low HDL-cholesterol were associated with high galectin-3 binding protein in type 1 diabetes. Biol Sex Differ. 2019;10(1):51.
226. Etemadikhah M, Niazi A, Wetterberg L, Feuk L. Transcriptome analysis of fibroblasts from schizophrenia patients reveals differential expression of schizophrenia-related genes. Sci Rep. 2020;10(1):630.
227. Perisic L, Lal M, Hulkko J, Hultenby K, Önfelt B, Sun Y, et al. Plekhh2, a novel podocyte protein downregulated in human focal segmental glomerulosclerosis, is involved in matrix adhesion and actin dynamics. Kidney Int. 2012;82(10):1071-83.

Appendix

Table 13: Total proteins detected employing in-solution and nano-LC MS

Protein IDs	Protein names	Gene names	Intensity
P04217	Alpha-1B-glycoprotein	A1BG	49788000
P01023	Alpha-2-macroglobulin	A2M	71986000
P43652	Afamin	AFM	9682500
O00468	Agrin	AGRN	1407500
P01019	Angiotensinogen	AGT	120060000
P02765	Alpha-2-HS-glycoprotein	AHSG	520610000
P02768	Serum albumin	ALB	4688800000
Q13740	CD166 antigen	ALCAM	1064300
P04075	Fructose-bisphosphate aldolase A	ALDOA	4327600
P09972	Fructose-bisphosphate aldolase C	ALDOC	3184200
P02760	Protein AMBP	AMBP	156480
Q5JPF3	Ankyrin repeat domain-containing protein 36C	ANKRD36C	1925200
P51693	Amyloid-like protein 1	APLP1	242060000
Q06481	Amyloid-like protein 2	APLP2	3974000
P02647	Apolipoprotein A-I	APOA1	1955900000
P02652	Apolipoprotein A-II	APOA2	390820000
P06727	Apolipoprotein A-IV	APOA4	136570000
P02656	Apolipoprotein C-III	APOC3	997690
P05090	Apolipoprotein D	APOD	34678000
P02649	Apolipoprotein E	APOE	1269200000
P02749	Beta-2-glycoprotein 1	APOH	34587000
O14791	Apolipoprotein L1	APOL1	1382900
P05067	Amyloid beta A4 protein	APP	105360000
P25311	Zinc-alpha-2-glycoprotein	AZGP1	4086300
P61769	Beta-2-microglobulin	B2M	1177200000
O43505	Beta-1,4-glucuronyltransferase 1	B4GAT1	42188000
Q96GW7	Brevican core protein	BCAN	51688000
P02746	Complement C1q subcomponent subunit B	C1QB	9046800
P02747	Complement C1q subcomponent subunit C	C1QC	6691000
P00736	Complement C1r subcomponent	C1R	1325800
P09871	Complement C1s subcomponent	C1S	26754000
P06681	Complement C2	C2	6340800
P01024	Complement C3	C3	625600000
P0C0L5	Complement C4-B	C4B	354800000
P01031	Complement C5	C5	3841300
P13671	Complement component C6	C6	3036500
P10643	Complement component C7	C7	3697100
P02748	Complement component C9	C9	9174100
P54289	Voltage-dependent calcium channel subunit alpha-2/delta-1	CACNA2D1	1179600
Q9BY67	Cell adhesion molecule 1	CADM1	2524000
P27797	Calreticulin	CALR	994410

Appendix

P13987	CD59 glycoprotein	CD59	14250000
Q8TCZ2	CD99 antigen-like protein 2	CD99L2	3832800
P55290	Cadherin-13	CDH13	2757000
P19022	Cadherin-2	CDH2	4756700
P00751	Complement factor B	CFB	178810000
P00746	Complement factor D	CFD	1032900
P08603	Complement factor H	CFH	52888000
Q03591	Complement factor H-related protein 1	CFHR1	2388700
P05156	Complement factor I	CFI	9440500
Q99674	Cell growth regulator with EF hand domain protein 1	CGREF1	1652700
P10645	Chromogranin-A	CHGA	129290000
P05060	Secretogranin-1	CHGB	776370000
P36222	Chitinase-3-like protein 1	CHI3L1	99511
O00533	Neural cell adhesion molecule L1-like protein	CHL1	46466000
P05452	Tetranectin	CLEC3B	15925000
O94985	Calsyntenin-1	CLSTN1	18538000
P10909	Clusterin	CLU	950770000
Q15846	Clusterin-like protein 1	CLUL1	208140
Q96KN2	Beta-Ala-His dipeptidase	CNDP1	89520000
P26992	Ciliary neurotrophic factor receptor subunit alpha	CNTFR	3186500
Q12860	Contactin-1	CNTN1	55602000
Q02246	Contactin-2	CNTN2	1697800
Q9C0A0	Contactin-associated protein-like 4	CNTNAP4	849870
P39060	Collagen alpha-1(XVIII) chain	COL18A1	2877200
P02452	Collagen alpha-1(I) chain	COL1A1	1508300
P08123	Collagen alpha-2(I) chain	COL1A2	1276800
P12109	Collagen alpha-1(VI) chain	COL6A1	12730000
P00450	Ceruloplasmin	CP	363630000
P16870	Carboxypeptidase E	CPE	7662500
O95196	Chondroitin sulfate proteoglycan 5	CSPG5	1954900
P01034	Cystatin-C	CST3	508630000
P07339	Cathepsin D	CTSD	8368100
Q14118	Dystroglycan	DAG1	5895000
P07108	Acyl-CoA-binding protein	DBI	524430
Q9UBP4	Dickkopf-related protein 3	DKK3	165020000
Q16610	Extracellular matrix protein 1	ECM1	5473800
O94769	Extracellular matrix protein 2	ECM2	1514800
Q13822	Ectonucleotide pyrophosphatase/phosphodiesterase family member 2	ENPP2	65722000
P00734	Prothrombin	F2	19376000
P12259	Coagulation factor V	F5	498270
Q92520	Protein FAM3C	FAM3C	95168000
P23142	Fibulin-1	FBLN1	16377000

Appendix

Q9Y6R7	IgGFc-binding protein	FCGBP	782100
P02671	Fibrinogen alpha chain	FGA	39002000
P02675	Fibrinogen beta chain	FGB	20538000
P02679	Fibrinogen gamma chain	FGG	39693000
P02751	Fibronectin	FN1	124520000
Q14393	Growth arrest-specific protein 6	GAS6	674600
P02774	Vitamin D-binding protein	GC	212590000
P14136	Glial fibrillary acidic protein	GFAP	658380
P22352	Glutathione peroxidase 3	GPX3	1430300
P06396	Gelsolin	GSN	445990000
P69905	Hemoglobin subunit alpha	HBA1	11084000
P68871	Hemoglobin subunit beta	HBB	4507600
P02042	Hemoglobin subunit delta	HBD	133310
P00738	Haptoglobin	HP	1654100000
P02790	Hemopexin	HPX	150600000
P04196	Histidine-rich glycoprotein	HRG	5014000
P11021	78 kDa glucose-regulated protein	HSPA5	1100200
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2	470750
P24592	Insulin-like growth factor-binding protein 6	IGFBP6	84946000
Q16270	Insulin-like growth factor-binding protein 7	IGFBP7	3263500
P01876	Ig alpha-1 chain C region	IGHA1	67980000
P01857	Ig gamma-1 chain C region	IGHG1	262670000
P01859	Ig gamma-2 chain C region	IGHG2	5263600
P01860	Ig gamma-3 chain C region	IGHG3	11405000
P01861	Ig gamma-4 chain C region	IGHG4	1486900
P01871	Ig mu chain C region	IGHM	4565000
P0DOX7	Immunoglobulin kappa light chain	IGK	69047000
P01614	Ig kappa chain V-II region Cum	IGKV A18	4583400
P0DOY3	Ig lambda-6 chain C region	IGLC6	93006000
B9A064	Immunoglobulin lambda-like polypeptide 5	IGLL5	8484100
A0A075B6K5	Ig lambda chain V-III region LOI	IGLV3-9	2042900
O14498	Immunoglobulin superfamily containing leucine-rich repeat protein	ISLR	2907000
P19827	Inter-alpha-trypsin inhibitor heavy chain H1	ITIH1	10358000
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	ITIH2	24535000
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	ITIH4	15955000
O43896	Kinesin-like protein KIF1C	KIF1C	155240000
Q92876	Kallikrein-6	KLK6	86469000
P01042	Kininogen-1	KNG1	12465000
P04264	Keratin, type II cytoskeletal 1	KRT1	56388000
P13645	Keratin, type I cytoskeletal 10	KRT10	17707000
P13646	Keratin, type I cytoskeletal 13	KRT13	42505
P08779	Keratin, type I cytoskeletal 16	KRT16	180880
P35908	Keratin, type II cytoskeletal 2 epidermal	KRT2	5665300

Appendix

P19013	Keratin, type II cytoskeletal 4	KRT4	298050
P35527	Keratin, type I cytoskeletal 9	KRT9	25212000
P04180	Phosphatidylcholine-sterol acyltransferase	LCAT	1631700
P07195	L-lactate dehydrogenase B chain	LDHB	3394400
Q8NES3	Beta-1,3-N-acetylglucosaminyltransferase lunatic fringe	LFNG	54531000
Q08380	Galectin-3-binding protein	LGALS3BP	33575000
P02750	Leucine-rich alpha-2-glycoprotein	LRG1	15379000
Q9NT99	Leucine-rich repeat-containing protein 4B	LRRC4B	673290
Q13449	Limbic system-associated membrane protein	LSAMP	20216000
P51884	Lumican	LUM	11112000
P43121	Cell surface glycoprotein MUC18	MCAM	5815700
P40925	Malate dehydrogenase, cytoplasmic	MDH1	696040
Q7Z7M0	Multiple epidermal growth factor-like domains protein 8	MEGF8	1853900
P08493	Matrix Gla protein	MGP	1819800
Q14CX7	N-alpha-acetyltransferase 25, NatB auxiliary subunit	NAA25	3886100
P13591	Neural cell adhesion molecule 1	NCAM1	64965000
O15394	Neural cell adhesion molecule 2	NCAM2	2910900
O14594	Neurocan core protein	NCAN	1401100
Q7Z3B1	Neuronal growth regulator 1	NEGR1	4353100
Q99435	Protein kinase C-binding protein NELL2	NELL2	19148000
Q92859	Neogenin	NEO1	298030
P23582	C-type natriuretic peptide	NPPC	656110
O95502	Neuronal pentraxin receptor	NPTXR	5587600
Q92823	Neuronal cell adhesion molecule	NRCAM	57814000
Q9NPD7	Neuritin	NRN1	783780
Q9Y4C0	Neurexin-3	NRXN3	4563300
Q9P121	Neurotrimin	NTM	22630000
P20774	Mimecan	OGN	12115000
Q99983	Osteomodulin	OMD	1094300
P02763	Alpha-1-acid glycoprotein 1	ORM1	8088300
P19652	Alpha-1-acid glycoprotein 2	ORM2	574150
Q15113	Procollagen C-endopeptidase enhancer 1	PCOLCE	5026800
Q9UHG2	ProSAAS	PCSK1N	14240000
P30086	Phosphatidylethanolamine-binding protein 1	PEBP1	10368000
P01210	Proenkephalin-A	PENK	28200000
Q96PD5	N-acetylmuramoyl-L-alanine amidase	PGLYRP2	744180
Q8IVE3	Pleckstrin homology domain-containing family H member 2	PLEKHH2	13631000
P00747	Plasminogen	PLG	16739000
P55058	Phospholipid transfer protein	PLTP	10700000
P27169	Serum paraoxonase/arylesterase 1	PON1	3569700
Q6S8J3	POTE ankyrin domain family member E	POTEE	3812700
P14314	Glucosidase 2 subunit beta	PRKCSH	736250

Appendix

P04156	Major prion protein	PRNP	17835000
P07225	Vitamin K-dependent protein S	PROS1	8565600
Q5FWE3	Proline-rich transmembrane protein 3	PRRT3	545120
P41222	Prostaglandin-H2 D-isomerase	PTGDS	1242300000
P23471	Receptor-type tyrosine-protein phosphatase zeta	PTPRZ1	10632000
Q16769	Glutaminyl-peptide cyclotransferase	QPCT	428320
P78509	Reelin	RELN	1256400
Q9BRS2	Serine/threonine-protein kinase RIO1	RIOK1	114870000
Q9P1V8	Sterile alpha motif domain-containing protein 15	SAMD15	6670000
P13521	Secretogranin-2	SCG2	46411000
Q8WXD2	Secretogranin-3	SCG3	60277000
P05408	Neuroendocrine protein 7B2	SCG5	17903000
Q13228	Selenium-binding protein 1	SELENBP1	159210
O75326	Semaphorin-7A	SEMA7A	9681400
P49908	Selenoprotein P	SEPP1	2894300
P01009	Alpha-1-antitrypsin	SERPINA1	804370000
P01011	Alpha-1-antichymotrypsin	SERPINA3	155600000
P08185	Corticosteroid-binding globulin	SERPINA6	791520
P05543	Thyroxine-binding globulin	SERPINA7	667430
P01008	Antithrombin-III	SERPINC1	2827000
P05546	Heparin cofactor 2	SERPIND1	32866000
P36955	Pigment epithelium-derived factor	SERPINF1	390220000
P08697	Alpha-2-antiplasmin	SERPINF2	9544700
Q9UPX8	SH3 and multiple ankyrin repeat domains protein 2	SHANK2	2029900
Q6ZSJ9	Protein shisa-6 homolog	SHISA6	1115200
P78324	Tyrosine-protein phosphatase non-receptor type substrate 1	SIRPA	11356000
P00441	Superoxide dismutase [Cu-Zn]	SOD1	3597100
P08294	Extracellular superoxide dismutase [Cu-Zn]	SOD3	8158900
P09486	SPARC	SPARC	1766500
Q14515	SPARC-like protein 1	SPARCL1	207690000
Q92563	Testican-2	SPOCK2	431610
P10451	Osteopontin	SPP1	238490000
P61278	Somatostatin	SST	11440000
P02787	Serotransferrin	TF	1377800000
O43493	Trans-Golgi network integral membrane protein 2	TGOLN2	5972700
P04216	Thy-1 membrane glycoprotein	THY1	2786300
P01033	Metalloproteinase inhibitor 1	TIMP1	6676000
P16035	Metalloproteinase inhibitor 2	TIMP2	1260000
P63313	Thymosin beta-10	TMSB10	433150
P62328	Thymosin beta-4	TMSB4X	7615500
P02766	Transthyretin	TTR	645740000
P10599	Thioredoxin	TXN	728400

Appendix

O15240	Neurosecretory protein VGF	VGF	39494000
P04004	Vitronectin	VTN	2878700
Q8TEU8	WAP, Kazal, immunoglobulin, Kunitz and NTR domain-containing protein 2	WFIKKN2	1363200
P01619	Ig kappa chain V-III region B6	IGKV3-20	91057

Appendix

Table 14: All identified proteins with one or the other method

	Gene names	1DE_PAGE MS	In-solution – nano-LC MS
1	COL18A1	√	√
2	MEGF8	√	√
3	SPARC	√	√
4	COL1A2	√	√
5	KRT16	√	√
6	LSAMP	√	√
7	C6	√	√
8	NRXN3	√	√
9	LUM	√	√
10	AGRN	√	√
11	F5	√	√
12	MDH1	√	√
13	NEO1	√	√
14	CFD	√	√
15	HSPG2	√	√
16	C5	√	√
17	PCSK1N	√	√
18	CALR	√	√
19	PENK	√	√
20	SEMA7A	√	√
21	CACNA2D1	√	√
22	MCAM	√	√
23	ALDOC	√	√
24	TIMP2	√	√
25	COL1A1	√	√
26	ECM1	√	√
27	HBD	√	√
28	C2	√	√
29	CNTN2	√	√
30	ITIH2	√	√
31	CHGA	√	√
32	APOC3	√	√
33	NCAN	√	√
34	ALDOA	√	√
35	FCGBP	√	√
36	ITIH1	√	√
37	NCAM2	√	√
38	GPX3	√	√
39	SERPINF2	√	√
40	DAG1	√	√

41	BCAN	√	√
42	PRNP	√	√
43	TIMP1	√	√
44	C7	√	√
45	LDHB	√	√
46	SERPINA7	√	√
47	C9	√	√
48	NTM	√	√
49	CPE	√	√
50	HRG	√	√
51	C1QB	√	√
52	OGN	√	√
53	CDH13	√	√
54	CD59	√	√
55	NELL2	√	√
56	CHGB	√	√
57	C1R	√	√
58	AMBP	√	√
59	NPTXR	√	√
60	NEGR1	√	√
61	IGHM	√	√
62	PROS1	√	√
63	SIRPA	√	√
64	FGA	√	√
65	CFI	√	√
66	C1S	√	√
67	PEBP1	√	√
68	CDH2	√	√
69	CHI3L1	√	√
70	SERPINA6	√	√
71	SPP1	√	√
72	PLG	√	√
73	PLTP	√	√
74	PON1	√	√
75	IGLV3-9	√	√
76	CFH	√	√
77	FGB	√	√
78	CTSD	√	√
79	ITIH4	√	√
80	AFM	√	√
81	PCOLCE	√	√
82	SPARCL1	√	√
83	CLSTN1	√	√

84	APP	√	√
85	SERPIND1	√	√
86	NCAM1	√	√
87	CHL1	√	√
88	SOD3	√	√
89	FGG	√	√
90	THY1	√	√
91	SOD1	√	√
92	FAM3C	√	√
93	LRG1	√	√
94	VTN	√	√
95	IGFBP6	√	√
96	APLP1	√	√
97	C1QC	√	√
98	IGFBP7	√	√
99	FN1	√	√
100	CNTN1	√	√
101	NRCAM	√	√
102	FBLN1	√	√
103	KNG1	√	√
104	LGALS3BP	√	√
105	IGHG3	√	√
106	CFB	√	√
107	CLEC3B	√	√
108	RELN	√	√
109	CNDP1	√	√
110	APOA2	√	√
111	ENPP2	√	√
112	F2	√	√
113	HBA1	√	√
114	B4GAT1	√	√
115	IGKV A18	√	√
116	APOH	√	√
117	KLK6	√	√
118	DKK3	√	√
119	A1BG	√	√
120	AZGP1	√	√
121	IGHG4	√	√
122	IGLL5	√	√
123	APOA4	√	√
124	KRT2	√	√
125	HBB	√	√
126	ORM2	√	√

127	IGKV3-20	√	√
128	GSN	√	√
129	SERPINC1	√	√
130	A2M	√	√
131	AHSG	√	√
132	IGHA1	√	√
133	CP	√	√
134	KRT9	√	√
135	SERPINA3	√	√
136	GC	√	√
137	B2M	√	√
138	C4B	√	√
139	KRT10	√	√
140	SERPINF1	√	√
141	HP	√	√
142	AGT	√	√
143	IGLC6	√	√
144	HPX	√	√
145	APOD	√	√
146	CLU	√	√
147	IGHG2	√	√
148	KRT1	√	√
149	ORM1	√	√
150	C3	√	√
151	APOE	√	√
152	CST3	√	√
153	APOA1	√	√
154	IGK	√	√
155	SERPINA1	√	√
156	PTGDS	√	√
157	IGHG1	√	√
158	TF	√	√
159	TTR	√	√
160	ALB	√	√
161	FSTL4	√	
162	NPTX1	√	
163	PEBP4	√	
164	NFASC	√	
165	SEZ6L	√	
166	IGLV1-36	√	
167	KRT72	√	
168	MMP2	√	
169	CST4	√	

170	LYNX1	√	
171	F12	√	
172	ATRN	√	
173	IGFBP2	√	
174	PIK3IP1	√	
175	KRT6B	√	
176	PAM	√	
177	GOT1	√	
178	CTSL	√	
179	RNASET2	√	
180	IGHV3-43D	√	
181	RARRES2	√	
182	C4A	√	
183	OMG	√	
184	BTD	√	
185	C4orf48	√	
186	DCD	√	
187	FRMPD1	√	
188	IGHV3-64D	√	
189	ACTG1	√	
190	CADM3	√	
191	C1QA	√	
192	IGHV3-49	√	
193	ENDOD1	√	
194	CRTAC1	√	
195	SAA4	√	
196	ABHD12B	√	
197	IGHV3-23	√	
198	OPCML	√	
199	HIST1H4A	√	
200	IGKV1-8	√	
201	SERPINI1	√	
202	IGKV1-6	√	
203	C8G	√	
204	PIBF1	√	
205	CADM4	√	
206	CSF1R	√	
207	VSTM2A	√	
208	CUTA	√	
209	IGHV4-61	√	
210	IGHV5-51	√	
211	LYZ	√	
212	DMD	√	

213	IGKV2-24	√	
214	ITGA11	√	
215	NPC2	√	
216	ANKRD50	√	
217	IGLV1-47	√	
218	KRT5	√	
219	IGHV3-15	√	
220	IGKV1-33	√	
221	TMEM198	√	
222	CD14	√	
223	ABCF1	√	
224	GM2A	√	
225	IGHV3-21	√	
226	RNASE1	√	
227	PSAP	√	
228	HIST2H2AC	√	
229	KRT14	√	
230	EFEMP1	√	
231	IGHA2	√	
232	UBA52	√	
233	CD22	√	
234	IGKV3D-11	√	
235	NCKAP5	√	
236	IGHV3-74	√	
237	IGKV4-1	√	
238	RBP4	√	
239	IGKV3D-7	√	
240	DNAH2	√	
241	GULP1	√	
242	HIST1H2AJ	√	
243	ELMSAN1	√	
244	MROH6	√	
245	SERPING1	√	
246	HABP2	√	
247	IGKC	√	
248	KIF1C		√
249	RIOK1		√
250	SCG3		√
251	LFNG		√
252	SCG2		√
253	VGF		√
254	SCG5		√
255	PLEKHH2		√

256	COL6A1		√
257	SST		√
258	PTPRZ1		√
259	TMSB4X		√
260	SAMD15		√
261	TGOLN2		√
262	APLP2		√
263	NAA25		√
264	CD99L2		√
265	POTEE		√
266	CNTFR		√
267	ISLR		√
268	SEPP1		√
269	CADM1		√
270	CFHR1		√
271	SHANK2		√
272	CSPG5		√
273	ANKRD36C		√
274	MGP		√
275	CGREF1		√
276	LCAT		√
277	ECM2		√
278	APOL1		√
279	WFIKKN2		√
280	SHISA6		√
281	HSPA5		√
282	OMD		√
283	ALCAM		√
284	CNTNAP4		√
285	NRN1		√
286	PGLYRP2		√
287	PRKCSH		√
288	TXN		√
289	GAS6		√
290	LRRC4B		√
291	GFAP		√
292	NPPC		√
293	PRRT3		√
294	DBI		√
295	TMSB10		√
296	SPOCK2		√
297	QPCT		√
298	KRT4		√

299	CLUL1		√
300	SELENBP1		√
301	KRT13		√

Declaration

I declare that this thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Johannes Gutenberg-Universität Mainz or at any other institution.