universität
wien

MASTERARBEIT

Titel der Masterarbeit

„Investigation of proteome alterations characteristic for tumor-associated cachexia: Combining high resolution MS-based screening with a targeted analysis strategy"

verfasst von

Martin Eisinger BSc.

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2015

Studienkennzahl lt. Studienblatt: A 066 862
Studienrichtung lt. Studienblatt: Masterstudium Chemie
Betreut von: Univ.-Prof. Dr. Christopher Gerner

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Definition
3D	three dimensional
APS	ammoniumperoxodisulfat
ATP	adenosine triphosphate
BAT	brown adipose tissue
CID	collision induced dissociation
CRP	C-reactive protein
CV	coefficient of variation
CVD	cardiovascular disease
DDA	data-depended-acquisition
DDA	data-directed-acquisition
dotp	dot-product
DTT	dithiothreitol
FDR	false discovery rate
GE	gel electrophoreses
GPIbA	platelet glycoprotein Ib alpha chain
GSH-S	glutathione synthetase
HCD	higher-energy collisional dissociation
HDL	high-density lipoprotein
HPLC	high-performance liquid chromatography
HUPO	human proteome organisation
ICAM-1	intercellular adhesion molecule
ITP	isotachophoresis
IEC	ion exchange chromatography
LC	liquid chromatography
LDHA	L-lactate dehydrogenase A chain
LDHB	L-lactate dehydrogenase B chain
LFQ	label-free quantification
LOD	limit of detection
LPS	lipopolysaccharide
LysC	lysyl endopeptidase
m/z	mass-to-charge ratio
MRM	multiple reaction monitoring
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MW	molecular weight
nESI	nanoESI
PAGE	polyacrylamidegel electrophoresis
PEP	posterior error probability
PTLP	Phospholipid transfer protein

Abbreviation	Definition
PTP	proteotypic peptide
QqQ	triple quadrupole
RP	reveresed phase
R-PTP-eta	receptor-type tyrosine-protein phosphatase eta
RT	retention time
SAA1	serum amyloid A-1 protein
SAA2	serum amyloid A-2 protein
SB	sample buffer
sCD163	scavenger receptor cysteine-rich type 1 protein M130
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SIM	selected ion monitoring
SIS	stable isotopically labeled standard
TEMED	tetramethylethylenediamine
TLT-1	trem-like transcript 1 protein
WAT	white adipose tissue
XIC	extracted ion chromatogram

Acknowledgements

I would like to thank the entire AG Gerner staff for the great work atmosphere, Besnik Muqaku for his excellent support throughout the project, and Uni. Prof Christopher Gerner for the opportunity and resources to complete my master-thesis research at his group.

1 Introduction

1.1 Cancer Cachexia – A Multifactorial Syndrome

Cancer associated cachexia, also known as "cancer cachexia" is a multifactorial and mostly irreversible syndrome that affects 50 to 80% of all cancer patients and accounts for approximately 16% of all cancer related deaths. The term cachexia etymological derives from the Greek *kakos* and *hexis*, which translates in "bad condition". However, in modern day diagnostics the term cachexia is used to describe an acute and multifactorial wasting disorder.[1] Currently there is no uniform classification when to speak of cachexia, though all definitions share a common ground. These common characteristics are the massive loss of body fat and skeletal muscle accompanied with appetite loss and deteriorations of the whole nutrition status.[1-2]

Cachexia itself is known for centuries and has been observed in a multitude of chronic and acute inflammatory diseases, like rheumatoid arthritis, tuberculosis and cancer.[3] In cachectic patients a so-called "metabolic switch" takes place, which has very serious outcomes. At first, the metabolic rate of the patient increases dramatically, which leads to higher resting energy consumption. In addition tissue browning is observed resulting in a increased rate of glycolysis. These factors lead to a massive loss in white body fat and the degradation of skeletal muscle as endogenous nitrogen source. This imbalance in the energy household cannot be revoked by consuming more energy. Furthermore, most cachectic patients lose their appetite and show massively decreased water and energy uptake. This tremendously increases patient burden, makes treatment of the causal disease difficult and in most cases leads to a rapid death. [1, 3-4] Even though cachexia in general and cancer cachexia particularly is well described since more than 70 years, the driving mechanisms behind it are poorly understood and proper treatment methods are not yet available.[5]

Though it was long assumed that tumor characteristics are the major contributor to the development of cachexia, recent studies suggest that more importantly host factors and the tumor environment affect the cachectic outcome.[2, 6] A better understanding of cancer cachexia can only be achieved, if tumor and stroma are investigated as one unit. This can be accomplished by several strategies like co-cultivation or inserting tumors to specific body parts in an animal model. Nevertheless, none of these models is comparable to the

complexity observed in human patients. In addition animal models, especially mouse models have shown to only poorly mimic human, especially in inflammatory diseases.[7] This makes patient studies the preferred approach, even though they are the more complex and challenging method.

For proteomic human studies, different samples such as blood, urine, or tissue can be used for analysis. If possible, especially when dealing with weakened patients or for long term studies, the least invasive method to take samples should be selected. Here, blood in general and blood serum in particular are preferred over all other human-derived samples. This stems from it being a minimally invasive sample with high stability over various temperatures and best representing the physiological state of an individual due to the blood´s circulatory properties.[8] Despite the inherent challenges (e.g. high complexity), the comparability and comprehensiveness of blood serum in combination with the existing and established medical laboratory infrastructure will ensure that it remains the favoured sample type for years.[8b]

Investigation of serum proteome alterations characteristic for tumor associated cachexia, may lead to a better understanding of the cachectic development. This can be used to develop better treatment strategies and may also help to understand how a tumor influences his micro-environment.

1.2 Proteomics – Current Methods and Limitations

Proteomics is a study that deals with the identification and quantification of all proteins present in a biological sample. The term "proteome" hereby reflects all proteins expressed in the sample under investigation and can be applied to tissue, a whole cell, or simply a body fluid. Proteins play an important role in all biological processes and their expression levels directly reflect the physiological state of a living organism. This is why proteomics has become one of the most powerful tools, not only in system biology, but also in clinical research.[9]

Since biological samples are highly complex, especially on protein level, an analytical technique with high resolving power is needed for proteomic studies. Therefore, liquid chromatography (LC) coupled to mass spectrometry (MS) has become the method of choice. For MS-based protein analysis two main strategies can be distinguished in modern day proteomics. The first is the so called "top-down" approach, where intact proteins get

separated via gel electrophoreses (GE) prior to MS analysis of the intact proteins. The second is the "bottom-up" approach, for which proteins get digested by a selected protease (mostly trypsin) into short peptides. These peptides are typically separated over a reverse phase (RP) LC before MS analysis. For protein identification, peptides with a unique sequence, so called proteotypic peptides (PTPs), are selected using extensive bioinformatics and protein databases. Even though top-down proteomics enables access to the whole protein sequence, it has some serious draw-backs. Intact proteins are prone to side reactions and fragmentation, the separation capacity of proteins via LC is very limited, and not all of them are accessible by mass spectrometry. Tryptic peptides however are mostly stable, easy to separate by RPLC and can be accessed by common MS techniques. This is why bottom-up approaches are mostly preferred over top-down, even though protein digestion radically increases work load and sample complexity.[10] Both strategies can be combined with different sample depletion, enrichments or pre-fractionation methods according to the complexity of the proteome under investigation.

For data acquisition in proteomic bottom-up approaches two main strategies are well established: the non-hypothesis driven untargeted and the hypothesis driven targeted approach. In untargeted proteomics a data acquisition technique called "shotgun" is used for MS analysis. Thereby a MS^1 survey scan is performed on each time-point in chromatographic separation. The most intense precursor ions in this survey scans are selected for fragmentation and measured in product ion scan mode. Peptides get then identified according to their fragment ion spectra and their precursor ion mass using search engines (e.g. proteome discoverer, MaxQuant, etc.) and protein databases (e.g. UniProt, SwissProt). From the screening manner of this approach, combined with the high sample complexity, stems the need for fast state-of-the-art mass spectrometers with high resolving power (e.g. Orbitrap).[11]

A targeted approach is usually conducted on a tandem mass spectrometer with high sensitivity. Therefore a triple quadrupole system (QqQ) is the platform of choice. Here a method design called multiple reaction monitoring (MRM) is used to filter pre-selected peptide masses and fragment them under optimized conditions to reduce matrix effects and enhance sensitivity to a maximum.[12] Both, targeted and untargeted strategies have their advantages and drawbacks. Untargeted approaches show a lack of sensitivity and accuracy, but on the other side benefit from the ability to identify and quantify hundreds of proteins within a single

runs and hence open a wide view into biological processes. Targeted approaches in comparison show high accuracy and sensitivity, but their pre-selective manner makes them blind for unforeseen biological events. The fundamental working principal of both methods is displayed in Figure 1.
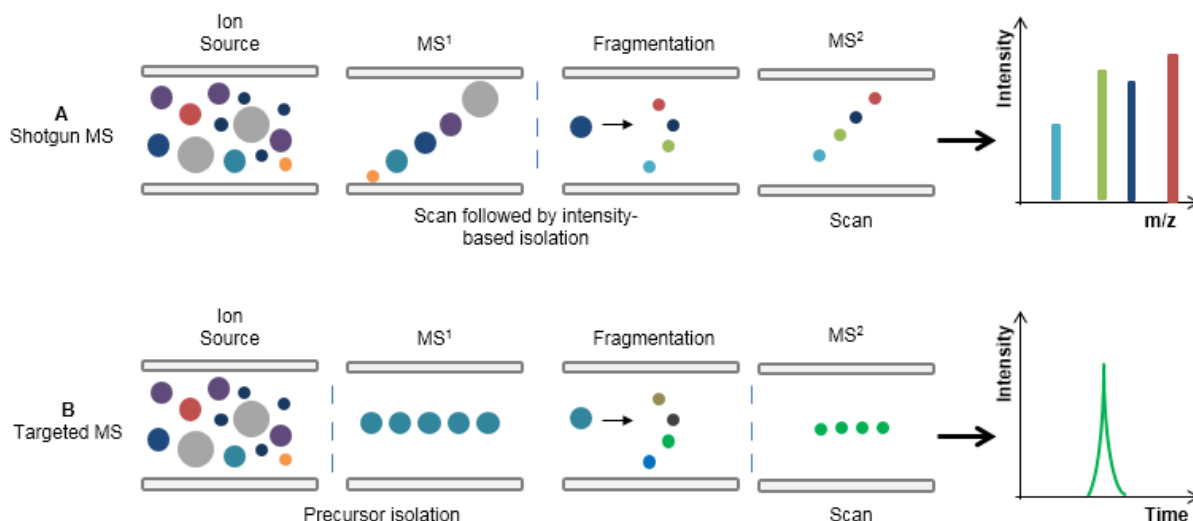
Figure 1: Fundamental working process of untargeted (A) and targeted (B) proteomics[9]
In untargeted shotgun measurements (A) a high resolution MS survey scan is performed and the 8-12 most abundant masses are then selected for MS^2 fragmentation and product ion scan. This is performed in a dynamic cyclic manner to assemble as much data as possible for post-acquisition identification. In targeted proteomics (B) pre-selected precursors for each peptide are isolated and fragmented. Selected fragments of these precursors are then further isolated before measuring to reduce background effects. This is done in a more static manner leading to shorter cycle times and higher throughput.

The quantification of the proteins under investigation can be done in an absolute and a semi-quantitative manner. The absolute quantification needs standards for each analyte under investigation. These peptide or protein standards are usually isotopically labeled in order to perform internal calibration and compensate for measurement variations. Semi-quantification is performed by simply comparing signal intensities or peak areas of selected peptides in two or more different groups. As this approach does not require isotopically labeling, it is often referred as "label-free quantification". Especially in untargeted approaches label-free quantification is used, since the synthesis and incorporation of labeled standards for the multitude of identified proteins is not applicable. The drawback of label-free quantification is that it is limited in accuracy, which is why it mainly focuses on a minimum 2-fold change of protein concentration. For most biological questions a LFQ is more than satisfying, however when a method should be moved towards clinical application the need of stable isotopically labeled standards (SIS; be it peptides or proteins) becomes imperative.[13]

1.3 Proposed Plan

Aim of this project was the investigation of serum proteome alterations characteristic for cancer associated cachexia. Therefor two different sample pretreatment methods should be tested, one using a sodium dodecyl sulfate (SDS)-GE for pre-fractionation and another which utilizes serum depletion for serum complexity reduction. The more appropriate strategy should then be applied on individual plasma samples obtained from final-stage cachectic and non-cachectic cancer patients as well as from healthy donors. A bottom-up shotgun analysis conducted on a Q Exactive Orbitrap followed by a label-free quantification using MaxQuant should then be used to compare the groups (healthy, non-cachectic, cachectic) against each other. Thereby proteins that are significantly altered in cachectic patients should be identified. For these candidate proteins a spectral reference library should be built in Skyline, considering possible interferences based on the high resolution data.

This library should then be used for the establishment of a targeted MRM strategy conducted on one of Agilent`s latest triple quadrupole mass spectrometer (Agilent 6490). To reduce sample complexity and expand the dynamic range of detection, a chromatographic separation should be performed prior to mass spectrometric analysis. Therefore a nanoChip-LC system interfered to the mass spectrometer via an electron spray ionization (ESI) source combined with an ion funnel should be used to ensure maximum sensitivity and reproducibility. Rigorous method development will be performed to select only interference free transitions and ensure optimal MRM parameters. Stemming from these results, a dynamic nanoChip-LC MRM method will be developed for the rapid, multiplexed, sensitive, and accurate label-free protein quantification in patient's serum samples.

This method should undergo thoroughly validation before be applied for the analysis of a multitude of patient samples: first, to compare cachectic cancer patients against non-cachectic and healthy patients and to identify characteristic proteome alterations; second, to measure different time points per cancer patients to screen for intra-patients variations and may identify possible progression markers. All data gathered should undergo thorough evaluation using sophisticated statistical methods and recent peer-reviewed literature. In the end, protein regulations significant for cancer cachexia will be outlined. This proteins may lead to a better understanding of the driving factors and mechanisms behind cancer cachexia and can possible be used as biomarker candidates.

2 Theoretical Background

2.1 Cancer Cachexia – A Multifactorial Syndrome

Cancer associated or tumor induced cachexia, short cancer cachexia, has been recognized since more than 70 years. It is observed in 50% to 80% of all cancer patients[1] and accounts for 10% to 22% of all cancer related deaths.[3] Cachexia is only observed in late-stage or final-stage tumor patients, though the frequency and progression of the cachectic outcome strongly differs between different tumor types.[3] Cachectic patients show a massive loss of body weight together with a loss of appetite and a dysfunctional energy uptake.[1] Even though these symptoms apply to a variety of eating disorders, cachexia is clearly to distinguish from anorexia and their like.[3] Cachectic patients show a strongly increased metabolic rate, which leads to massive loss of body fat and skeletal muscle, even if enough energy is consumed by the patient.[1, 3, 14] The final-stage of cancer cachexia is mostly starvation, although high caloric nutrition is given to most patients.

Even though the symptoms of cancer cachexia are well described, the driving mechanisms behind it are still poorly understood. Cachectic development involves many different physical processes and nearly all organs of the body; this is why cachexia is often referred as "multifactorial".[15] The first stage of cachectic development is always a long and chronic inflammation and a persistent hypoxic environment. In this stage hypoxic cells, especially tumor cells, fundamentally change their metabolism to ensure a high survival rate.[16] As tumor cells endeavor to maintain a high proliferation and survival rate, they show an exceptional metabolism even before cachectic outcome. The characteristics of tumor metabolism are, that the tumor environment shows a low nutrient and oxygen supply due do the lack of vasculature.[16] In order to still maintain growth, tumor cells gain 90% of their energy by glucose dependent ATP production. This was firstly discovered in the 1950s by Otto Warburg and is known as the "Warburg effect".[17] As a consequence of the Warburg effect, mitochondrial activity decreases nearly to zero and tumor cells undergo mitophagy. Although it was assumed that these processes are restricted to occur in cancer cells only, recent studies show that tumor-associated fibroblast can undergo similar processes.[6] Due to the persistent oxidative stress in the tumor environment, fibroblasts turn into a survival mode. Thereby, they increase anti-oxidant defense and in order to protect themselves and neighboring cells from apoptosis. Further, they provide the tumor with energy-rich building

6

blocks for anabolic growth.[18] This system of tumor cells and associated fibroblast leads to a higher survival, mutation and proliferation rate of tumor cells under hypoxic conditions.[6, 18] Today this phenomenon has been proven by many independent approaches and is called "tumor-stroma co-evolution". As this tumor-stroma co-evolution is observed in nearly all cancer cases, it does not necessarily lead to cachexia. Nevertheless it is the origin of cachectic development.[1, 6, 16, 18]

This metabolic change paired with the tumor-stroma co-evolution lead in the long term to a metabolic switch, which affects the whole body. The first stage is the massive loss of body fat and skeletal muscle to maintain the high energy transfer into the tumor. With regards to the later, cancer patient`s muscles often show a disturbed adenosine triphosphate (ATP) production coupled with hypertrophy.[19] Subsequently pro-inflammatory proteins, such as interleukin-1, are released into the muscle. These mediators induce the expression of ligases (e.g. E3 ligase, MURF1) and promote protein degradation.[19] The muscle proteins become degenerated to serve as nitrogen and energy source for the tumor. This happens mainly over direct glutamine transfer to the tumor and supporting the liver with alanine.[1] Additionally, the synthesis of new muscle fibers becomes strongly down regulated and apoptosis of muscle cells is increased. The sum of these facts leads to a massive and mostly irreversible loss of muscle mass during cachectic outcome.

The loss of skeletal muscle in cachectic patients is always accompanied by massive wasting and browning of adipose tissue. The wasting of the white adipose tissue (WAT) is a first stage, which has three driving factors.[20] One is the increased lipolytic activity, which results in an activation of hormone-sensitive lipases. This leads to a release of glycerol and fatty acids into the blood system.[1] Next, the activity of the lipoprotein lipase is lowered, resulting in a hindered lipid uptake in WAT. Third, lipogenesis is down regulated in cachectic patients, leading to a decreased lipid deposition.[1, 20] Additionally to this loss and the hampered synthesis of WAT, newest studies imply that fat cells also undergo a browning process during cachexia.[4, 21] [4] Brown adipose tissue (BAT) is usually found in the neck of healthy humans and along the spine.[22] Its main function is the thermogenesis, thus the shivering free production of heat within the fat cells. The browning of fat cells is generally linked to the expression of UCP1, which contributes to the mitochondrial switch from ATP production to thermogenesis. Inflammation and tumor induced factors (e.g. IL-6, PTH-related proteins) have shown to increase the production of UCP1 and consequently lead to a massive increase

in BAT mass. The newly generated BAT itself then starts to produce heat, resulting in an increased body temperature and energy demand of the cachectic patient.[4] The loss and browning of adipose tissue always occurs together with muscle degeneration in cachectic patients. However, recent studies have proven, that decreased lipolysis results in a retention of muscular dystrophy. This implies that there is indeed a cross-talk between fat and muscle mass, which may serve as cachectic marker.[1, 20]

This cross-talk is surely conducted by adipokines and myokines, such as leptin or interleukins.[10] As many hormones and other signaling molecules are triggered in this muscle-fat signaling, also liver and brain must be involved to regulate this on an upper level. Moreover, the liver responses to the increased energy need of the body by increased production of short energy rich molecules, like glucose. This, together with the disturbed metabolism in the tumor environment, leads to a circulation of glycerol and ketone bodies (e.g. pyruvate) in cachectic patients. Furthermore, the flow of amino acids from the muscle to the liver leads to increased acute-phase protein synthesis. The acute-phase proteins enhance the inflammation process and thereby accelerate the energy wasting. The brain on the other hand is also massively involved in the altered metabolism in cachexia. [1 23] The high energy request from the body leads to a massive release of appetite stimulating hormones, like ghrelin and insulin. These permanently increased hormone levels may lead to the development or resistances. In consequence patient`s appetite and voluntarily food intake becomes reduced. Beside brain and liver also the heart is involved in cancer cachexia. During cancer and chronic inflammation the heart rate is increased resulting in a higher energy consumption and faster metabolism.[24] In addition barrier dysfunction, especially gut barrier dysfunction is observed in cancer patients, particularly in cachectic ones.[25] This leads to the release of bacterial toxins (e.g. LPS) into the blood system and thereby activates immune response. This steady activation of the immune system leads to inflammatory processes, which accelerate the cachectic development and furthermore decrease the energy uptake in the gut.[1, 25] As already mentioned, cachexia is indeed multifactorial, since not only liver, brain, heart, gut, muscle and fat tissue is involved, but furthermore the whole body metabolism. This multitude of cachectic effects is illustrated by Figure 2.
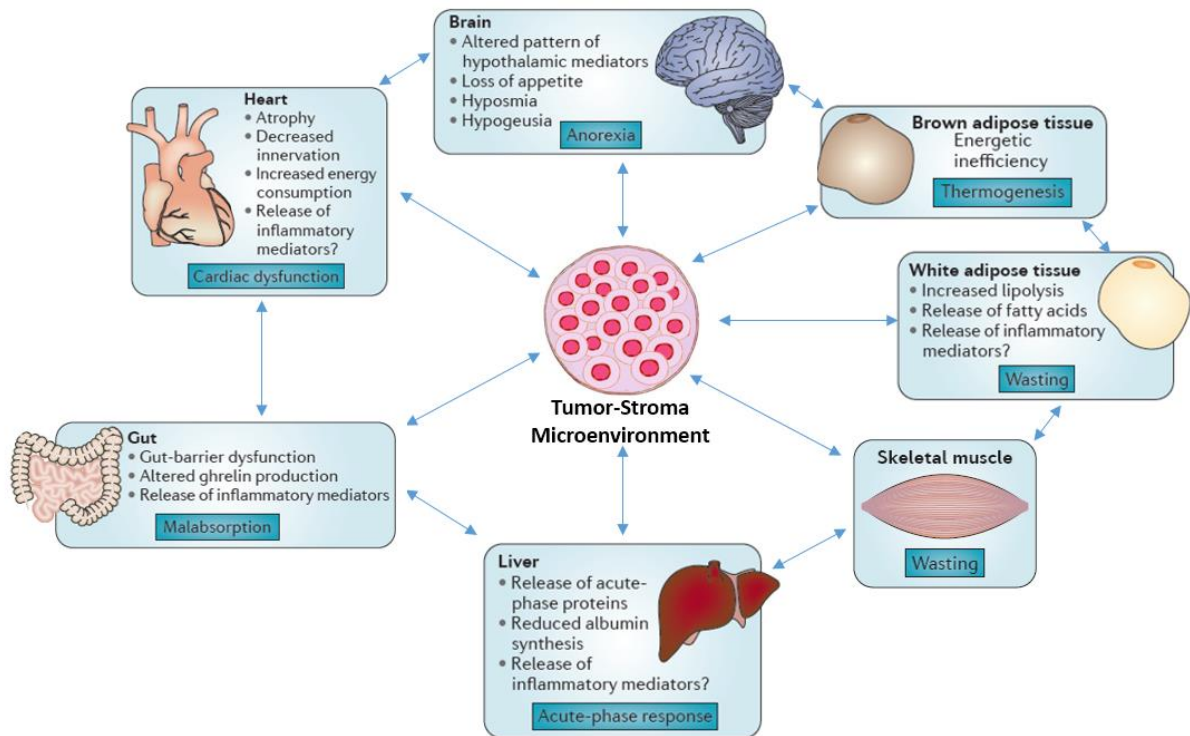
Figure 2: Cachexia: A multifactorial multi-organ syndrome[1]

The tumor and its associated fibroblast induce a wasting of adipose and muscle tissue, which is activated by the muscle WAT cross-talk. Thereby released amino acids from muscle degradation activate inflammatory response in the liver, which mediates the wasting process. This is further accelerated by inflammation processes in the gut, because of barrier dysfunction and the increased heart rate induced by the tumor. The constant demand for high energy leads to a certain resistance in the brain, which concludes in the loss of appetite. The cross-talk between the organs leads to more and more inflammatory response and energy wasting and late stage of cachexia starvation occurs even when enough energy is provided to the patient.

As mentioned in the introduction, the chance of cachectic development is strongly linked to certain tumor types (e.g. colon cancer.). Therefore, it was initially assumed that the ability to induce cachexia is dependent on the tumor cell type. However, more recent studies involving co-cultivation and single cell cloning proof that the cachectic development strongly correlates with the tumor location.[2, 6, 18] This means that not the tumor itself is able to induce cachexia, but rather the tumor host-factors and the tumor-associated fibroblast are the triggers for cachectic development. This understanding must lead to a tremendous change in the therapeutic strategies. Today cachexia treatment focuses on supply of high caloric nutrition and anti-inflammatory medication. Both strategies only cope with the symptoms of cachexia but not with the causality. A better understanding of how a tumor influences his environment and a strategy to block this tumor-stroma co-evolution might prevent the development of cachexia. This will not only decrease patient's burden, but furthermore increase overall

survival rate. This is due to the fact that many cachectic cancer patients cannot be treated against the tumor, because their bad physical condition is not suitable for most exhausting therapeutic tumor treatments (e.g. chemotherapy). [3] Furthermore, there is a valid hope that by blocking the tumor support from the associated fibroblast, a better immune defense against the tumor can be triggered. Additionally, the success rate in tumor therapy would strongly increase, as many modern tumor medications are known to fail because of the tumor-microenvironment. This would be applicable for many, also non-cachexia inducing, tumor types and paves the way towards the successful fight against the global cancer burden.

2.2 Human Blood Serum - Its Advantages and Challenges

The utility of blood in disease understanding and diagnostics has been known since ~370 B.C.. Hippocrates was the first who claimed that diseases can be caused by disorders in or between the body fluids. This idea of using blood remained over 1000 years and was reawaken in 1882 with the first synthesis of urea. With this synthesis, the distinction between living matter and chemicals began to disappear and the pathway to modern day blood diagnostics opened. Within the last 200 years, blood analysis has been consistently improved. This commenced with the discovery of single proteins, such as albumin in the 1830s, the fractionation of blood plasma by Cohn and Edsall in 1928, and the measurement of enzymatic activities and antibodies towards modern blood diagnostics in 1950.[26]

Blood is one of the most remarkable human proteomes. The body of a human adult contains in average 5 – 6 L of blood, which equals 8% of the total body mass. In modern day clinics blood is divided in three fractions: whole blood, plasma, and serum. Whole blood is the unmodified collected blood, from which plasma and serum are derived. Plasma refers to the liquid portion of whole blood in which cells and other insoluble substances are suspended. The plasma fraction constitutes 55% of the whole blood and is gathered by centrifugation in the presence of anticoagulants (e.g. heparin). Serum refers to the liquid, coagulation factor-free portion of the blood and is obtained by the removal of the coagulation factors from whole blood. Although plasma and serum is almost used interchangeably, plasma is the preferred blood proteome by the Human Proteome Organization (HUPO).[27] Reasons therefore are, that plasma is the more reproducible sample. The coagulation process adds a certain variability to the sample, which may affect the recovery of the target proteins.[12] However in clinical

diagnostics serum is the preferred sample type because of its higher stability and simpler matrix.

Blood in general it is not only the most complete and comprehensive, but also the most complex and challenging proteome. The liquid fraction of blood (be it plasma or serum) contains a multitude of different proteins spanning a dynamic range of 10 orders of magnitude in concentration. Due to its circulation through the whole body, plasma contains tissue derived proteins as well as the "true" blood proteome. The true blood proteome is defined as those proteins that carry out their function in the circulation and show an extended plasma life time.[26] On the upper end of protein concentration range albumin (41 mg/mL) and the immunoglobulins (11 mg/mL) can be found, together they contribute to over 80% of the whole blood proteins mass (~70 mg/mL). On the lower end, tissue-derived and messenger proteins can be found. They usually occur in blood below the low ng/mL level. A typical and important class of these low abundant proteins are interleukins (or cytokines). Even though they play a major role in human immune response, they only occur in the low pg/mL range in human blood. Figure 3 displays the protein concentrations observed in blood plasma and points out its high dynamic range. This plot in principle also applies to blood serum and illustrates its analytical challenges.
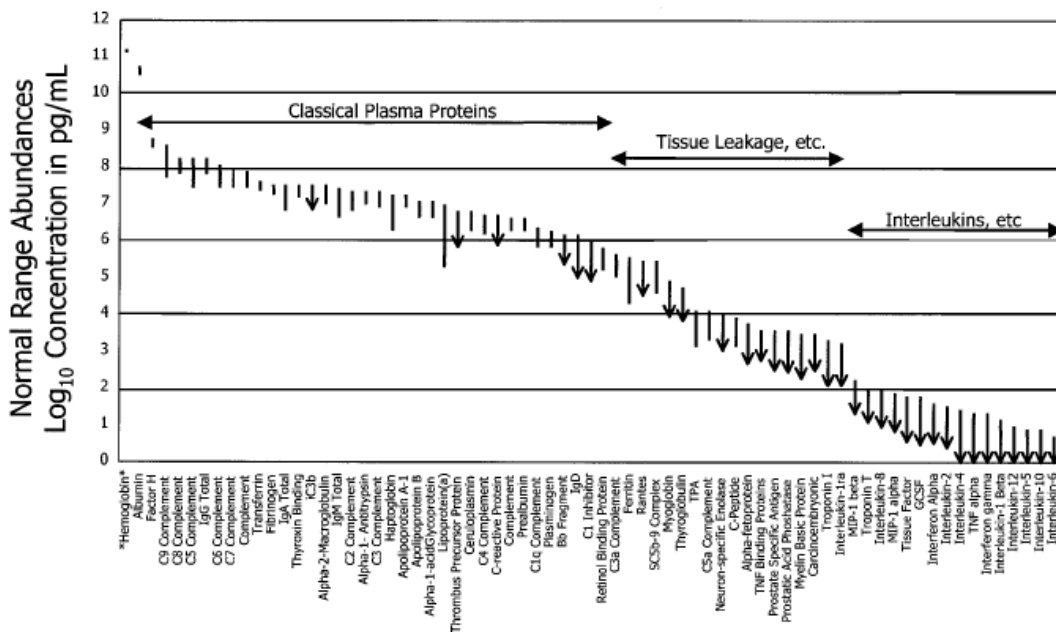

Figure 3: Reference values of protein concentrations in blood plasma combined by Anderson et al., 2002[26]
Proteins grouped according to their observed plasma concentration and biological function. Notably, protein concentrations in plasma span over more than 10 orders of magnitude.

To put this into perspective, identifying an interleukin like IL-1ß (1.2 pg/mL) among the higher abundance proteins is comparable to searching the entire world population for one specific individual. Despite these inherent challenges, the comparable and stable nature as well as the barely invasive gathering makes blood the most used diagnostic sample.

2.3 Serum Fractionation Techniques

As mentioned above blood serum is one of the most challenging matrices in modern day proteomics. With its dynamic concentration range spanning 10 orders of magnitude, even modern day LC-MS platforms reach their limitations. State of the art MS detectors can handle three to five orders of magnitude. [28] In combination with LC, the dynamic range of detection can be extended to maximum of six to eight orders of magnitude.[29] Unfortunately, most tumor derived proteins and inflammation markers, which are highly interesting for system biology, only occur in the lower concentration range (10 ng/mL and below). In order to be able to identify and quantify them as well, a pre-fractionation of serum is required. There are three different well-established strategies for serum pretreatment: serum fractionation, enrichment, and depletion.

Serum fractionation can be achieved by different separation techniques such as LC or GE. As usually a RPLC is interfered to the mass spectrometer, the first dimension of serum separation uses an alternative LC-method. Here mainly size exclusion chromatography (SEC) and ion exchange chromatography (IEC) are employed.[30] Alternatively also GE can be performed as fractionation step. Here one dimensional SDS **p**oly**a**crylamide **g**el **e**lectrophoresis (PAGE) and two dimensional GE are the most utilized ones. Enrichment can be performed on protein or peptide level and is achieved by the specific binding of the target analytes to a stationary phase. This can be accomplished by the use of antibodies or other binding molecules, such as lectins (for glycoproteins).[31] Depletion involves removal of the most abundant proteins, such as albumin and the immunoglobulins, from the blood serum. In doing so, the complexity of the matrix is reduced and the intensity of the lower abundant proteins becomes enhanced. Depletion is mostly performed by immune affinity capturing involving the removal of the six to fourteen most abundant proteins from the serum.

For this work two different serum fractionation techniques should be used and compared to each other; namely a SDS-PAGE separation and a depletion using immune affinity columns.

2.3.1 SDS - PAGE

SDS-PAGE is a widely used separation technique in biochemistry to separate proteins. In GE in general proteins are separated according to their electrophoretic mobility. The electrophoretic mobility is dependent of the net charge state, the size (molecular weight), and the form of the protein. Proteins show a variety of different charges, sizes and forms, so that SDS-PAGE was developed to reduce these factors. SDS is an anionic surfactant, containing a C12 carbon tail attached to a negatively charged sulfate head. When added to a protein mixture the negatively charged head neutralizes positive charges, whereas the neutral tail attaches it to hydrophobic parts of the protein. This leads to a distribution of negative charges along the protein and thus an unfolding of the proteins happens. SDS thereby cleaves hydrogen bonds and electrostatic interaction but no covalent bindings and disulfide bridges. In the end all proteins carry a net negative charge and a similar mass-to-charge-ratio. During the following PAGE, they can hence be separated according to their molecular weight.

For the separation itself polyacrylamide gels are used. Polyacrylamide is a copolymer accomplished by the polymerization of acrylamide with N,N'-methylenebisacrylamide in an approximately 40:1 ratio. The polymerization is started using ammonium peroxodisulfate (APS) as radical starter and tetramethylethylenediamine (TEMED) as catalyst. The pore size of the gel is determined by the percentage of the used acrylamide solution and can be altered according to the molecule weight range under investigations. Additionally, buffers can be incorporated into the polymer to ensure a certain pH value for separation. The gels are given in a cast with a separated anionic and cationic side and a buffer reservoir for the electrolyte. The protein mixture is loaded into gel pockets on the cationic side and a voltage is applied to start the separation. The negatively charged proteins start moving towards the anode and become separated along the way according to their size. Molecular weight markers, so protein mixtures with known weight, are typically put on the gel as well. These mixtures are called ladders and can be used to determine the molecular weight ranges of certain fractions on the gel. After the separation is completed, the proteins in the gel need to be stained in order to make them visible. Therefor different staining techniques like silver staining or coomassie staining can be used. [32] Afterwards the desired protein faction can be cut out of the gel and further processed for downstream MS analysis.

The use of a SDS-PAGE for serum fractionation has some serious drawbacks. GE shows considerable run-to-run variations, which add to the overall coefficient of variation (CV). Further, the cutting process is prone to errors and reduces reproducibility and the automation and throughput of the method is very limited. However, its high resolving power, its ease of use and the lack of the need for antibodies make the SDS-PAGE still an important method in modern day proteomic sample preparation.

2.3.2 Serum Depletion utilizing Pierce™ Top12 Depletion Columns

Depletion of serum or plasma sample is one well-established strategy to reduce the complexity of the matrix and enhance sensitivity. As already mentioned earlier, depletion can be achieved via different methods. For this work the immune affinity capturing via antibodies was chosen. Immune affinity capturing is a typical depletion method used in research projects as well as in routine analysis. So there are many commercial products available, which promise to remove the two to fourteen most abundant proteins out of plasma or serum. For this work, the "Pierce™ Top12 Abundant Protein Depletion Spin Columns" were used. These columns contain an antibody-coated resin suspended into a pH 7.4 buffer. The resin contains antibodies against the 12 most abundant proteins in blood, which are shown in Table 1. These antibodies should capture the target proteins with an efficiency of 95-99% and thereby remove about 90% of the total protein mass in serum.[33]

Table 1: Twelve most abundant proteins in blood[26, 33]

α1-Acid Glycoprotein	Fibrinogen
α1-Antitrypsin	Haptoglobin
α2-Macroglubulin	IgA
Albumin	IgG
Apolipoprotein A-I	IgM
Apolipoprotein A-II	Transferrin

Even though depletion is well accepted in the proteomic community, it should be mentioned that it also has some pitfalls. These drawbacks include increased costs and variability of the method, lower sample throughput, and increased sample loss resulting from the depletion of carrier proteins, such as albumin. Furthermore, the removal of 90% of the protein mass reduces the dynamic concentration range by one order of magnitude, leaving a still very challenging matrix. As mass spectrometric devices become more and more sensitive and

separation techniques increase their efficiency, there will be a reduced need for depletion in the future. However, its capability for automation and the current limitation in instrumentation, make depletion one of the favourable methods in today`s proteomic research.

2.4 Protein Quantification by modern Proteomics

2.4.1 Sample Preparation for Bottom - Up Proteomics

For untargeted and/or quantitative approaches bottom-up is the preferred strategy, due to its higher accuracy and the ability for multiplexing. For bottom-up analysis, the proteins need to be enzymatically cleaved into shorter peptides prior to the mass spectrometric analysis. In order to increase the efficiency of this enzymatic reaction, proteins are denaturized prior to the digestion. Denaturation is achieved by adding a mild surfactant or chaotropic agent (e.g. ammonium formate) to the protein mixture and reducing the disulfide bonds. The reduction of the disulfide bonds can be achieved by different agents, mainly dithiothreitol (DTT). Afterwards the free thiol groups become protected in order to prevent re-linkage. This can be done with iodoacetamide (IAA) for example. Figure 4 shows the reaction for the reduction of the disulfide bonds with DTT (A) and the protection of the thiol groups with IAA (B).
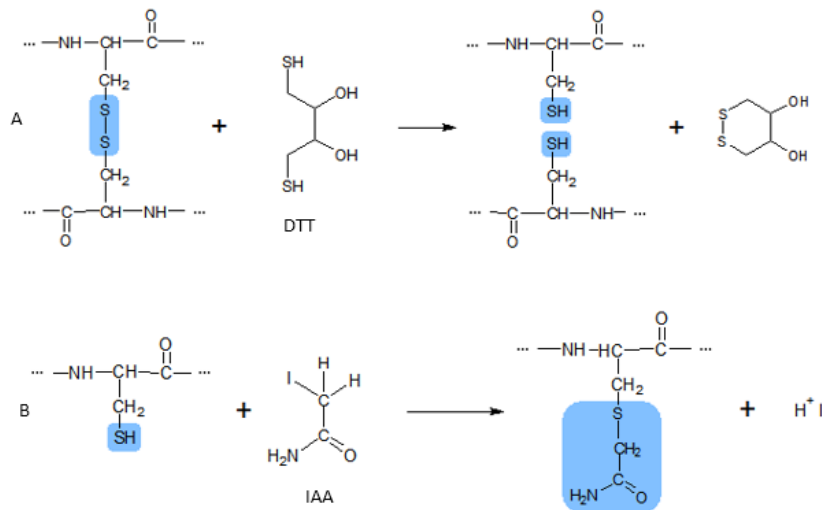
Figure 4: Reduction of the disulfide bonds and protection of the freed thiol groups
Prior to the tryptic digestion, disulfide bonds become reduced to thiol groups with DTT (A) in order to cleave cross-links within the protein and to unfold the protein. These thiol-groups are then protected with IAA (B) to prevent re-linkage.

The cleavage of the unfolded proteins is then performed by proteases, here mostly trypsin is used. Trypsin is a stable and aggressive serine protease that can be found in the pancreas of

vertebrates. It specifically cleaves the proteins at the C-terminal side of arginine (R) and lysine (K). In order to achieve a more complete digestion and to avoid missed cleavages, lysyl endopeptidase (LysC) is often used in combination with trypsin. LysC specifically cleaves at the C-terminus of K and is even more robust than trypsin. Tryptic peptides are very suitable for LC-MS analysis, because they always carry the basic amino acids K and R at the C-terminus, which promotes ionization and fragmentation. Further they show a good RPLC separation and their mass range allow a sensitive and accurate MS detection.[34] Even though tryptic digestion is a well-established and widely used proteomic workflow, it has some drawbacks. First of all not all proteins can be accessed by tryptic digestion, due to digestion resistance or sequence parts that contain no R or K. Additionally, the digest itself is an enzymatic reaction, which is very sensitive to reaction conditions like temperature, pH, or reaction time. This always adds a degree of uncertainty to the method and this is why highly standardized protocols are a must have in proteomics.

2.4.2 Nano RPLC and Nano ESI

2.4.2.1 Nano RPLC

High-performance liquid chromatography (HPLC) is a chromatographic separation technique based on the distribution of analytes between a liquid mobile and a solid stationary phase. Depending on the combinations of these two phases, different types of chromatography can be distinguished. Common types of LC are the separation according to size (SEC), ionic strength (IEC) or polarity (RPLC). In principle all of these techniques would be applicable for peptide separation, however the combination with a mass spectrometer leads to certain limitations. Due to the nature of the ESI process, it is not suited for solutions with a high salt content or ionic strength. Unfortunately this is needed for SCX and IEC making a direct coupling to an ESI source challenging and only possible by additional desalting steps. RPLC on the other hand does not need a high ionic strength for elution or an increasing salt content. Furthermore, the nature of RPLC, where analytes a distributed between a polar stationary phase (mostly C18) and an unpolar mobile phase, makes it suitable for a multitude of different organic molecules. This, the ability to be directly coupled with an ESI source, and the ease in method employment makes RPLC the most widely used separation technique today.[35]

Nano liquid chromatography (nanoLC) was firstly introduced in 1988 by Karlsson and Novotny.[36] Currently there is no general definition when to speak of nanoLC, but most "nano" applications use capillaries with an inner diameter of 10 to 100 µm. The use of miniaturized HPLC columns offers several advantages over conventional applications. First, higher separation efficiency is achieved in shorter separation time, combined with less mobile phase consumption. Second, less sample volume is required, which is preferable for biological samples, as they are very often limited in availability. The reduced flow rates also result in less chromatographic dilution, providing higher sensitivity and lower variability. When interfered to an ESI source the use of nanoLC enables a higher ionizations efficiency by reducing ion suppression and matrix-effect to a minimum.[36]

2.4.2.2 Nano ESI

The role of the ion source is to generate gas-phase ions out of solution-phase analytes for transfer to the mass selective units. ESI is considered to be the most effective and most preferable ion source for interfacing with RPLC, especially for the analysis of macromolecules, such as polypeptides. As ionization happens under atmospheric pressure and without fragmentation of the analytes, ESI is denoted as a "soft" ionization technique. In the ESI process the liquid sample eluting from the LC system flows through a capillary and is nebulized under atmospheric pressure by a spray needle. On the spray needle an electric current (3-6 kV) is applied leading to the generation of charged droplets. As the droplets move towards the vacuum system, in which the mass selective units are embedded, the solvent starts to evaporate. This can be supported by heating the ion source or using a drying gas depending on the flow rates. The evaporation of the solvent leads to an increasing charge density on the surface of the droplets resulting in coulomb explosions. The solvent droplets keep reducing their size by repeating this process till the droplet charge is transferred to the analyte molecules embedded in the solvent. This results in a charge distribution on the analyte molecules, depending on their structural properties. The (multiple) charged analytes are then released in the gas phase and can enter the mass selective unit. Depending on the applied charge on the spray needle, multiple protonated or deprotonated molecule ions are generated in the ESI process. Since the free ammonium groups of the basic amino acids and the N-terminal end make peptides prone for protonation, positive ESI is commonly used for peptide analysis. A schematic representation of the ESI process can be seen in Figure 5 (taken from Banerjee, 2011[37]).
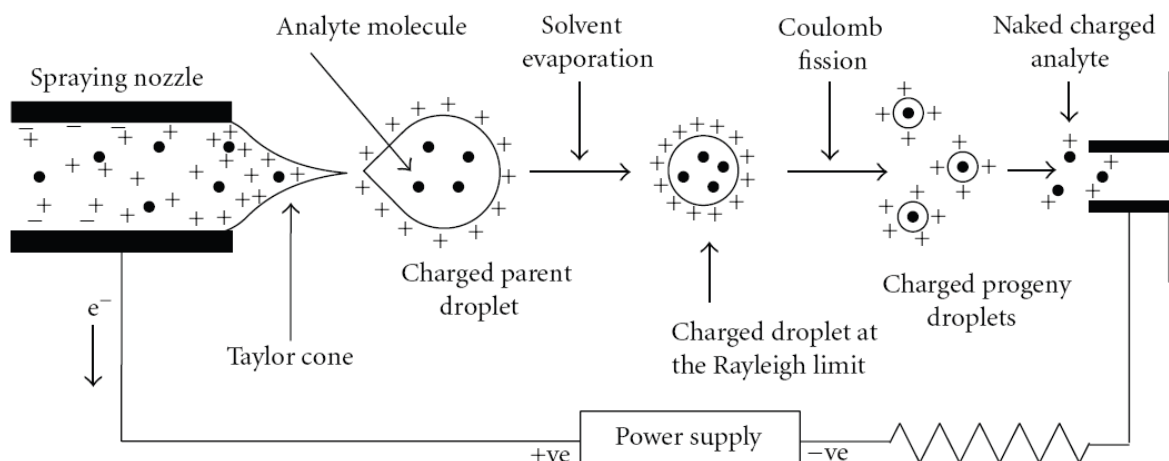
Figure 5: Schematic representation of the ESI process[37]

Surface charged droplets generated by the spraying capillary decrease in size because of solvent evaporation till they reach the Rayleigh limit. The high charge density on the surface leads to multiple coulomb explosions, till the charges are transferred to the analyte molecules. The charged analytes are then released into the gaseous phase either to subsequent coulomb explosions or ion evaporation.

Nano ESI (nESI) is a miniaturized version of conventional ESI that uses smaller capillaries (10-30 μm i.d.) and emitter tips. It is compatible with low flow rates (200-500 nL·min⁻¹) generally used in nanoLC. The reduced flow rates lead to the formation of smaller droplets that need less desolvatization for the generation of gaseous analyte ions. This results in a better charge transfer and thus in an increased ionization efficiency. The high ionization efficiency in nESI results in reduced ion suppression and matrix effects. As a consequence of the reduced ion suppression and the enhanced analyte ionization, the sensitivity is massively increased in nESI.[38]

2.4.3   ShotgunMS utilizing a Q Exective Orbitrap

The Q Exactive Orbitrap from Thermo Scientific is a state of the art hybrid quadrupole-orbitrap mass spectrometer. In the configuration used for this work, it consist of a nESI source, ion guide optics, a quadrupole mass selector, a C-trap, a higher-energy collisional dissociation (HCD) cell, and an orbitrap mass analyzer.[39] A schematic assembly of the Q Exective orbitrap is shown in Figure 6.
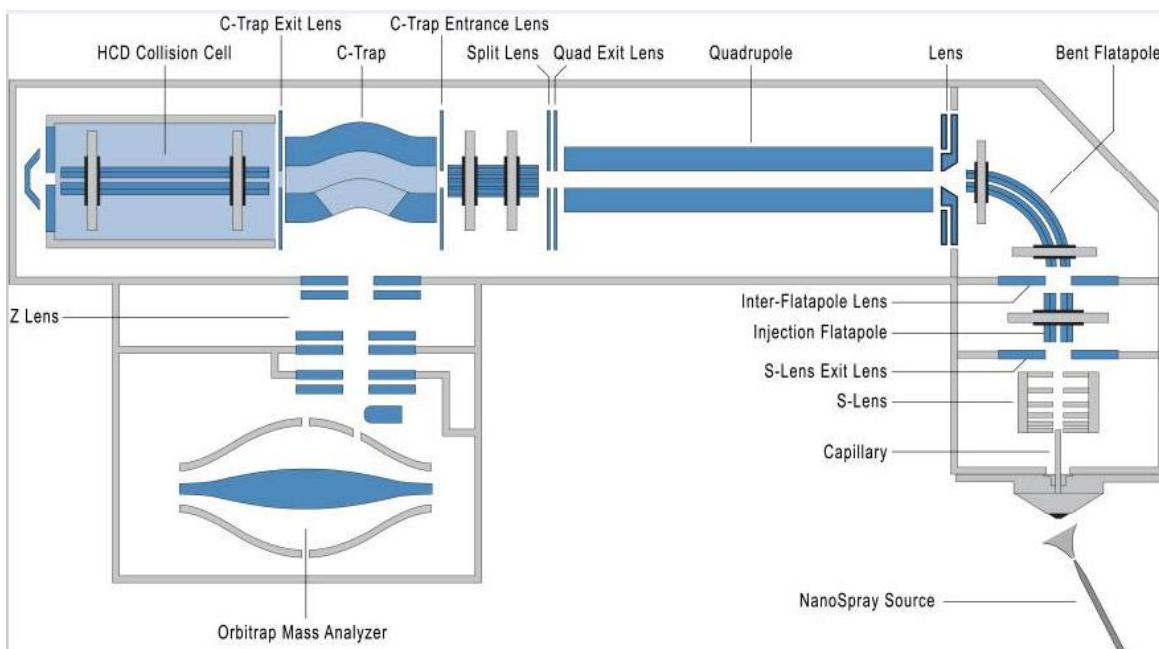
Figure 6: Schematic assamble of the QExactive orbitrap[39]

Ions generated in the ESI source are cooled in the S-lens and guided to the first analytical quadrupole. Here ion guidance or selection can be performed. Ions that passed through the quadrupole are further cooled and transferred into the C-trap. When the maximum infusion of the C-trap is reached, ions can be either sent to the orbitrap analyzer or to the HCD collision cell. In the HCD cell ions are further fragmented and send back to the C-trap from where they are finally send to the orbitrap. The orbitrap itself acts as high resolution mass analyzer and detector giving not only the ions m/z but also their abundance based on the signal gathered.

The ions generated in the ESI source are focused through the S-lens and transferred into the quadrupole via the bent flatapole. The quadrupole can act as ion guide or mass selective unit with a nominal isolation width. Followed to the quadrupole the ion beam travels through ion optics and a short octapole, which brings the ions then into the C-trap. The C-trap collects, cools, and stores the ions. When the C-trap has reached its maximum filling, the cooled ion stack can either be transferred to the HCD collision cell for fragmentation or into the orbitrap mass analyzer. Fragmentation in the gas filled HCD cell is achieved by applying an acceleration voltage to the ions. Thereby, the accelerated ions collide with the gas molecules and break into shorter fragments. The acceleration voltage can be set to certain values or ramped to cover more possible fragmentations. The HCD process is very similar to the collision-induced dissociation (CID) process and lead for peptides mainly to the formation of b- and y-type ions. The different ion types of peptides are displayed in Figure 7.
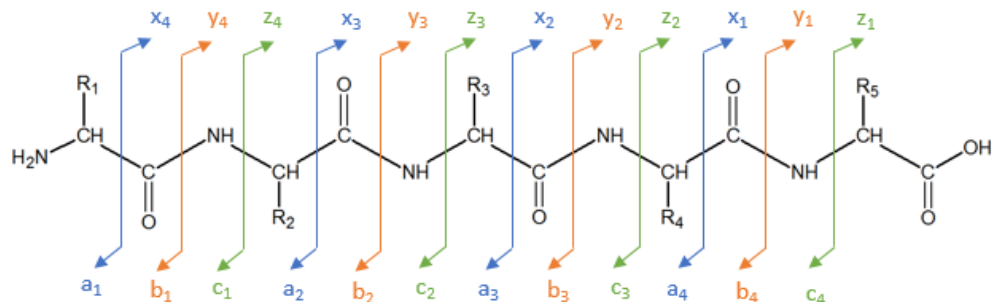
Figure 7: Peptide fragmentation sites and nomenclature of resulting ions
Depending on the cleavage site at the peptide backbone, fragment ions are denominated as a-, b- or c-ions when they contain the N-terminus of the peptide and as x-, y- or z-ions, when they contain the C-terminus of the peptide. The subscripted character specifies the number of amino acids contained in the fragment ion

These resulting fragment ions are cooled inside the HCD cell and transferred back to the C-trap. This filling of the C-trap can be performed while the previous orbitrap detection cycle is still ongoing, leading to significantly reduced cycle times. Further, more than one ion type per cycle can be stored in the C-trap before fragmentation, allowing the simultaneous fragmentation of many precursor ions. This increases the multiplexing ability and allows several new operation modes. To enhance sensitivity the fill-time of the C-trap can automatically be adjusted to the signal intensity of the precursor ions in a MS^1 scan. The ion detection and mass selection is then both performed in the orbitrap.

The orbitrap was developed by Makarov and is commercially available since 2005. It consists of a spindle-shaped electrode embedded into a small electrostatic field. The ion packages from the C-trap are injected with high energy into this field and orbit around the spindle electrode. The axial cyclic motion results in a current, which is measured up by the orbitrap detector. This signal is then Fourier transformed into the cycle frequency of each ion. By proper calibration the frequency can be converted to certain mass-to-charge ratio (m/z) values, yielding in a high resolution (≤1 ppm) mass spectra. Additionally the signal intensity of each frequency is recorded as well giving the ion current of each m/z. The orbitrap offers outstanding mass resolution of up to 140,000 for 200 m/z by maintaining a small benchtop suited size and the hybrid set up of the Q Exective Orbitrap allows many different operation modes.[39] For this work the data-depended-acquisition (DDA) was used and is further described below.

2.4.3.1 Shotgun Mass Spectrometry Data Acquisition

Shotgun proteomics, also known as discovery proteomics, refers to the analysis of protein digests by LC tandem mass spectrometry (MS/MS) operated in DDA. In a Q Exectatve orbitrap this is performed in a cyclic manner. At the beginning of every measurement cycle, a high resolution MS survey scan is performed. Therefore, the quadrupole guides all ions within a certain m/z range (mostly 400-1400) into the C-trap. When the system accounts the C-trap as filled, based on the ion current, previous data or reaching of the maximum infusion time, the cooled ion package is infused into the orbitrap and an MS[1] survey scan with high resolution (@70,000 for 200 m/z) is executed. According to pre-selected parameters, the 6-12 most abundant ions in this survey scan are determined. These most abundant ions are sequentially selected by the quadrupole (by using 1 m/z wide quadrupole isolation windows) and guided into the C-trap. When the C-trap is filled (filling is determined as described above), the selected precursor ions are transferred and fragmented in the HCD cell. Afterwards, the fragment ions are transferred to the C-trap, injected into the orbitrap and a product ion scan with medium resolution (@17,500 for 200 m/z) is performed. After a product ion scan of all pre-selected precursor ions is done, the next cycle starts again with a MS[1] survey scan. To avoid manifold fragmentation of the same precursor ions, the already fragmented species are mostly set on dynamic exclusion windows for a certain time frame. Additionally, static exclusion and inclusion lists can be used to trigger the fragmentation of as many different precursors as possible.[40]

2.4.3.2 Protein Identification and LFQ using MaxQuant

There are currently many commercial and free software packages available for protein identification and quantification from the gathered shotgun data, like OpenMS, Proteome Discoverer or MaxQuant. Even though these software packages rely on the same approach for the protein identification, they vary in the used algorithm. In here we utilized MaxQuant for protein identification and the implemented LFQ algorithm for relative quantification, due to its acceptance in the proteomic community and its reliable quantitative performance.

The first step of an each shotgun experiment is the protein identification; this is performed by searching against common protein databases (e.g. UniProt). Therefore, the software detects all features present in the gathered MS[1] shotgun data. The term feature is used to describe each peak with corresponding m/z, retention time (RT) and intensity. These features are

determined by fitting Gaussian peak shapes over three central data points. So a three dimensional (3D) peak is assembled. The software then detects the isotopic pattern and multiple charge states of the corresponding feature sets and deconvoluts them. This reduces the number of features by a factor of about 10. All corresponding MS^2 spectra for the detected features are gathered, peak centroids are generated (if not already gathered in the MS run) and the intensity-weighted average of all peak centroids is calculated. After all features are picked and assigned to their corresponding re-calibrated product ion spectra, the peptide identification can be performed.[41] Figure 8 shows the feature distribution identified by MaxQuant gathered from a typical shotgun MS run of depleted blood serum.
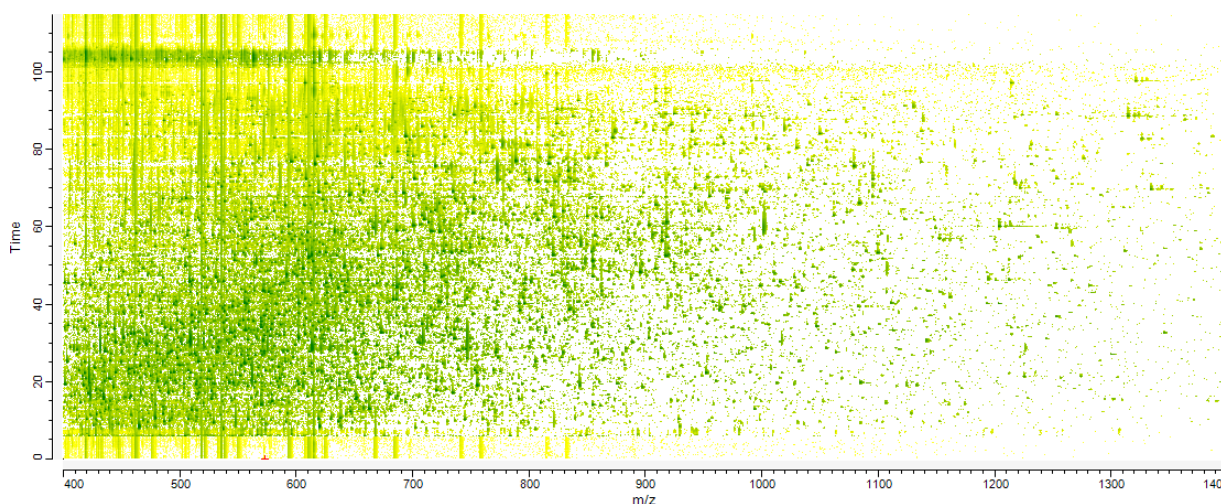
Figure 8: Feature distribution identified by MaxQuant for a blood serum sample analyzed by shotgun MS
On the x-axis the m/z of each identified feature is shown, whereas the RT time is displayed on the y-axis. Notably, this plot shows the detected features prior to the isotopic deconvolution. Thus serval features can correspond to one peptide.

For peptide identification the used protein database becomes *in-silico* digested and fragmented. Therefore, the MaxQuant is given the digestion conditions of the experiment, including used enzyme and the alkylation agent. Further restrictions, such as peptide length, possible charge states and number of tolerated missed cleavages can be set by the operator. During the *in-silico* digested, all possible peptides resulting from the enzymatic proteolysis considering the set modifications and restrictions are calculated. From these resulting precursor ions, possible fragment ions (preferable b- and y-ions) are calculated and theoretical product ion mass spectra are generated. The recorded mass spectra are matched against the theoretical spectra and scores are assigned based on their similarity and the observed mass error on MS^1 level. So each feature that can be assigned to a theoretical spectrum of a peptide accounts as identified with a certain score. These identified peptides

22

are then assigned to proteins, according to their sequence. For PTPs, the peptides can be directly assigned to the corresponding protein, while for shared peptides statistical methods are used. Shared peptides can either be assigned to a protein that is unambiguously identified by PTPs, it can be assigned based on multiple shared peptide patterns or just account for a protein group. For each protein or protein group identification, another probability score is calculated based on sequence coverage, score of the corresponding peptides and sequence uniqueness. To now determine how this probability score correlates with the number of false protein identification, each identified peptide is also searched against a revere version of the selected database, the so called "decoy-database". Also for decoy peptides a peptide score is calculated, as explained for the target peptides. Then for target and decoy peptide identifications a score distribution is calculated and this distributions are overlaid. Now for each peptide score the probability that this peptide is false identified can be calculated, this is the posterior error probability (PEP) for each individual peptide. This can then be used to set an acceptance level for possible false identified proteins, the so called false discovery rate (FDR). The FDR threshold can be configured by the user and is typically set to 1%. This means the search engine accepts proteins, which have a probability of being false positives of less than or equal to 1%.[41] Since the distribution of false identification overlaps with the distribution of correctly identified peptides, the concept of the FDR always leads to the loss of possible correct identifications. However, this concept is less rigid then classical statistical methods and by far the best method currently available in proteomics. For LFQ the software integrates the deconvoluted 3D peak areas for each identified peptide. The peak areas are then normalized based on the sum of all peak areas present overall samples. Therefore a matrix between all samples and all identified peptides is formed and an overall relative intensity is calculated for each sample under investigation. Peptide signals are then converted to protein level and a delayed normalization to this relative sample intensity is performed. This is done to account for run-to-run variations and different digestion efficiencies. The final result of these mathematical operations is an arbitrary unit for each protein or protein group abundance, the so called "LFQ intensity".

This LFQ intensity is a relative measure of the abundance of each identified protein in the samples under investigation. The LFQ intensities can now be used to compare protein abundance between different samples. Here it is to note that because of the normalization and the relative nature of LFQ intensities, these can only be used to compare corresponding biological samples. So samples from the same biological proteome (e.g. cell lysate), gathered

and digested in the same manner and measured on the same, preferable the identical platform. Here the measurement of two or more corresponding groups, for example a treated against a control group has emerged a standard procedure. By statistical comparison of the LFQ intensities between the groups (e.g. t-test), significantly altered protein abundances can be determined.

2.4.4 Targeted MS utilizing Agilent`s 6490 QqQ-System

For quantitative targeted proteomics, triple quadrupole instruments (QqQ), operated in MRM mode are the most commonly used and most sensitive systems. A QqQ system consist of two mass selective quadrupoles with a collision cell sandwiched in between. In MRM mode, the first and the third quadrupoles are operated in selected ion monitoring (SIM) mode, meaning they let only pass pre-selected m/z values. Fragmentation in the collision cell is achieved via CID at set collision energies. The pair of a precursor ion and one resulting product ion at a fixed collision energy is called transition. For quantification in general three transitions per analyte are monitored to ensure high selectivity. These transitions can be monitored in a static cyclic manner throughout the entire run (static MRM). Thereby the cycle time increases and the dwell time on each transition decreases. This is why static MRM is limited in sensitivity and capability of multiplexing. A further development of MRM is called "dynamic" or "scheduled" MRM. In dynamic MRM, the RT of each analyte is taken into account and transitions for each analyte are only monitored within a certain timeframe. In this timeframe the dwell time for each transition depends on the set cycle time. Therefor a cycle time should be chosen, which gives enough points over the chromatographic peak for precise and reproducible integration. Scheduled MRM assays reduce the number of concurrent transitions and hence enable longer dwell times and higher sensitivity. Additionally, the number of analytes quantifiable in a single run tremendously increases and consequently a sensitive, accurate and selective high throughput assay is generated. Figure 9 illustrates the underlying principle of MRM.
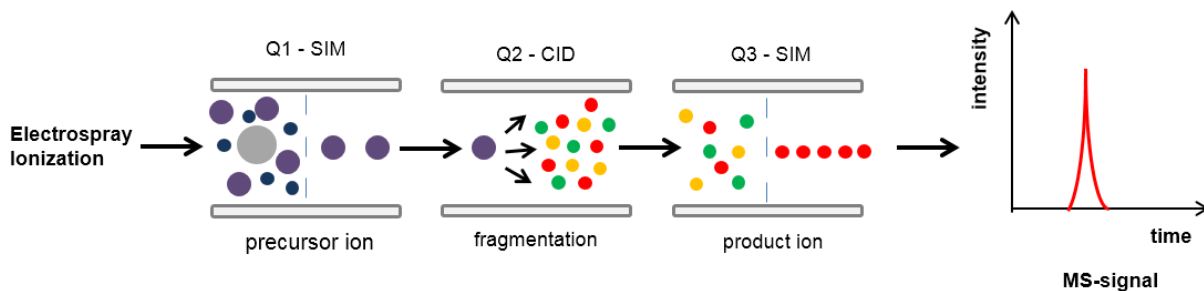
Figure 9: Principle of a triple quadrupole operated in MRM/MS mode
In Q1 only the selected precursor ion for a target peptide can pass through. The precursor is fragmented in Q2 via CID using a defined collision energy, thereby mainly b- and y-ions are formed. Q3 is set to monitor up to three product ions per precursor. These two mass selective steps, combined with an analyte specific fragmentation, offer highest selectivity and reduce background noise significantly.

Development of a dynamic MRM assay for targeted proteomics is a work intensive and time consuming task. Firstly, suited PTPs for each target proteins have to be selected. These PTPs have to be frequently observed by enzymatic digestion. Further, they should not contain possible post-translational modification sites, missed tryptic cleavages, or easily oxidized amino acids (e.g. methionine).[28] Additional they should be 8-20 amino acids in length and show sufficient hydrophobicity for RPLC seperation. The selection of these peptides can either be based on previous shotgun experiments or on bioinformatics. For the bioinformatics approach real experimental data can be accessed and compiled (e.g. PeptideAtlas) or predictive software tools can be used (e.g. PepFly). After suited PTPs are found, the most abundant fragment ions must be selected and the collision energy needs to be optimized for these fragment ions. This can either be done by carrying out previous optimization experiments or again with the use of databases and software packages (e.g. Skyline).[12] At least the RTs of each peptide need to be determined in order to schedule the assay. Here the most common way is to perform multiple unscheduled experiments to find the RT for each peptide individually. The use of databases is typically not applicable, since many different LC-setups and gradients can be used for RPLC separation. However, recent software packages for MRM method development offer RT prediction based on structural calculation (e.g. iRT calculation in Skyline).[12, 28]

As the workflow points out, the development of a targeted MRM assay is indeed an elaborate task. Nevertheless, MRM offers outstanding properties for targeted quantitative proteomics. The use of two mass selective steps with an analyte specific fragmentation enables highly selective and sensitive measurements. As a consequence, matrix effects and background

noise become reduced to a minimum. Thus, MRM/MS has the potential to provide very low detection limits, even when it comes to the measurements of highly complex samples, such as blood serum.

3 Materials and Methods

3.1 Materials

3.1.1 Chemicals and reagents

All chemicals and reagents employed in this thesis were of the highest grade available; Table 2 lists them, as well as their vendors and purities.

Table 2: Chemicals and Reagents

Chemical (Purity)	Abbreviation	Vendor
2-Iodoacetamide (crystalline)	IAA	Sigma Aldrich
2-Mercaptoethanol (electrophoresis, 99%)	BME	Sigma Aldrich
2-Propanol (HPLC grade)	iPrOH	VWR
Acetic acid (pro analysis)	AcOH	Merck
Acetonitrile (HPLC grade)	ACN	VWR
Acrylamide (99.9%)	AA	Bio-Rad
Ammonium bicarbonate (ReagentPlus, ≥99.0%)	ABC	Sigma Aldrich
Ammonium persulfate (pro analysis)	APS	Merck
Bradford 1x Dye Reagent	Bradford solution	Bio-Rad
Bromophenol blue sodium salt	Phenol blue	Merck
Cholamidopropyldimethylammonio-1-propanesulfonate (98%)	CHAPS	Gerbu
Dithiothreitol (high purity)	DTT	Gerbu
Formic acid (MS grade, ~98%)	FA	Fluka
Glycine (Ultra pure, ≥99.0%)	Gly	Fluka
Guanidinium hydrochloride (titration, ≥99%)	GHCl	Sigma Aldrich
Hydrochloric acid (pro analysis, ~37%)	HCl	Merck
Methanal solution (min. 37%)	HCHO	Merck
Methanol (LC-MS grade)	MeOH	VWR
N,N,N′,N′-Tetramethylethane-1,2-diamine (~95%)	TEMED	Sigma Aldrich
N,N'-methylene-bis-acrylamide (≥99.0%)	PDA	Sigma Aldrich
Potassium ferricyanide (ACS reagent, ≥99.0%)	$K_3[Fe(CN)_6]$	Sigma Aldrich
Precision Plus Protein™ Dual Color Standards	Marker	Bio-Rad
Silver nitrate (ReagentPlus, ≥99.0%)	$AgNO_3$	Sigma Aldrich

Chemical (Purity)	Abbreviation	Vendor
Sodium carbonate (≥99.5%)	Na_2CO_3	Sigma Aldrich
Sodium dodecyl sulfate (ReagentPlus, ≥98.5%)	SDS	Sigma Aldrich
Sodium thiosulfate pentahydrate (pro analysis, 99%)	$Na_2S_2O_3 \cdot 5\ H_2O$	Merck
Thiourea (ACS reagent, ≥99.0%)	CH_4N_2S	Sigma Aldrich
Trifluoroacectic acid (HPLC grade, ≥99%)	TFA	Sigma Aldrich
Tris(hydroxymethyl)-aminomethan (Ultrapure)	Tris	Gerbu
Trypsin/Lys-C Mix (MS grade)	TL	Promega
Ultrapure Water (Milli-Q + LC-Pak Polisher)	H_2O	Merck Millipore
Urea (for biochemistry)	CH_4N_2O	Merck

3.1.2 Human Serum Samples

Serum from cachectic and non-cachectic cancer patients was gratefully received from Dr Med Albrecht Reichle, Universitätsklinikum Regensburg. All serum samples were gathered from final-stage melanoma patients on a weekly basis, from the time they were accounted as beyond treatment till death. Classification in cachectic and non-cachectic cohorts was performed by the attending physician and by taking typical cachectic symptoms in account, such as loss of body weight, overall outlook, and progression of death.

Healthy serum was collected from three race and age matched donors, using VACUETTE® serum tubes (Greiner Bio-One, Germany).

3.2 Methods

3.2.1 Solution Preparation

3.2.1.1 Electrophoresis

Table 3 shows the solutions used for gel polymerization and GE.

Table 3: Solutions for gel polymerization and GE

Solution	Preparation
30% Acrylamide	29.2 g AA + 0.8 g PDA / 1000 mL H_2O
2 M Tris-HCl pH 8.8	242.28 g Tris, adjusted to pH 8.8 with HCl / 1000 mL H_2O
1 M Tris-HCl pH 6.8	60.57 g Tris, adjusted to pH 6.8 with HCl / 500 mL H_2O
20% SDS	20 g SDS / 100 g H_2O
10% APS	10 g APS / 100 g H_2O
10x Tris-Glycine	30 g Tris + 144 g Gly / 1000 mL H_2O
90% 2-Propanol	90 mL 2-propanol + 10 mL H_2O
Electrode buffer	100 mL 10x Tris-glycine buffer + 20% SDS 1000 mL H_2O

Table 4 shows the solutions employed for sample preparation in the GE.

Table 4: Solutions for sample preparation in the GE

Solution	Preparation
5x Laemmli buffer	5 mL 1 M Tris-HCl pH 6.8 + 2 g SDS + 0.05 g phenol blue + 71.4 µL BME + 17.5 mL H_2O
Sample buffer (SB)	22.5 g urea + 5.7 g thiourea + 0.77 g DTT + 2 g CHAPS + 125 mL 20% SDS / 50 mL H_2O

Table 5 displays the solutions used for staining and de-staining the gels.

Table 5: Solutions for gel staining and de-staining

Solution	Preparation
Fixing solution	500 mL MeOH + 100 mL AcOH + 400 mL H_2O
Washing solution	500 mL MeOH + 500 mL H_2O
0.02% $Na_2S_2O_3$	1 mL 2% $Na_2S_2O_3$ (using $Na_2S_2O_3 \cdot 5\ H_2O$) + 99 mL H_2O
0.1% $AgNO_3$	10 mL stock (10 mg/mL $AgNO_3$) + 90 mL H_2O
Developer Solution	3 g Na_2CO_3 + 130 µL HCOH / 100 mL H_2O
Stop solution	1 mL AcOH + 99 mL H_2O
De-staining solution	0.49 g $K_3[Fe(CN)_6]$ + 1.24 g $Na_2S_2O_3 \cdot 5\ H_2O$ / 100 mL H_2O

3.2.1.2 In-Gel Digestion

Table 6 shows the solutions utilized for the in-gel digest.

Table 6: Solutions for in-gel digestion

Solution	Preparation
25 mM ABC	0.099 g ABC / 50 mL H_2O
50 mM ABC	197.5 mg ABC / 50 mL H_2O
1 M DTT stock (in-gel)	1,54 g DTT / 10 mL 50 mM ABC
20 mM DTT	100 µL 1 M DTT stock + 4.9 mL 25 mM ABC
500 mM IAA stock (in-gel)	0.92 g IAA / 10 mL 50 mM ABC
100 mM IAA	500 µL 500 mM IAA stock / 10 mL 50 mM ABC
TL Stock	20 µg TL + 200 µL 50 mM AcOH

3.2.1.3 In-solution Digestion

Table 7 displays the solutions employed for the in-solution digestion of the depleted serum samples.

Solution	Preparation
500 mM ABC	39.5 mg ABC / 1 mL H_2O
50 mM ABC	197.5 mg ABC / 50 mL H_2O
DTT stock (in-solution)	278 mg DTT + 42.4 g GHCl / 50 mL H_2O
32 mM DTT	900 µL DTT Stock + 100 µL 500 mM ABC
IAA stock (in-solution)	555.6 mg IAA + 42.4 g GHCl / 50 mL H_2O
54 mM IAA	900 µL IAA Stock + 100 µL 500 mM ABC
TL stock	20 µg TL + 200 µL 50 mM AcOH

3.2.1.4 Equimolar Peptide Mixture

For internal normalization and MS quality control, synthetic peptides were spiked into each sample prior to MS analysis. Therefore, an equimolar peptide mix was prepared containing each 10 fmol of four synthetic standard peptides. The synthetic peptides [Glu1-Fribrinopeptide B - EGVNDNEEGFFSAR; M28 --TTPAVLDSDGSYFLYSK; HK0 - VLETKSLYVR; HK1---VLETK(ε-AC)SLYVR] were obtained from Peptide Specialty Laboratories GmbH and the final mix was stored at -20°C upon usage.

3.2.2 Bradford Assay

An in-house Bradford assay was performed on all serum samples before and after depletion in order to estimate their protein content. The determination of the total protein content is necessary to ensure correct protein content for depletion, digestion and peptide load on chromatographic column. Therefore, each sample was diluted (if necessary) and mixed H_2O and 50 µL Bradford solution to a final volume of 250 µL. The solution was vortexed for 30s and the observed coloring was compared to an in-house color table displayed in Figure 10.

Figure 10: Color scale for the Bradford assay
The figure displays the colors gathered in a Bradford assay with a standard solution using 1 µL sample volume. The optical distinction is best possible for protein concentrations between 1 and 4 µg/µL, because afterwards the blue shades cannot be distinguished by eye anymore.

The color table gives the protein concentration in µg/µL based on a sample volume of 1 µL. For different sample volumes, back-calculations were performed calculate the correct protein amount. As Figure 10 indicates, the color shades can be best distinguished for protein concentrations between 1 and 4 µg/µL. So sample dilutions were always prepared to fall within this preferred concentration range. The Bradford assay was performed for all samples prior use to ensure comparable sample conditions. Moreover, all depleted samples were investigated by Bradford assays to check for successful depletion and to enable optimal conditions for further sample treatment.

3.2.3 Sample Preparation

3.2.3.1 SDS-PAGE Fractionation and In-Gel Digestion

Gel and sample preparation

For serum fractionation via SDS-PAGE a well-established in-house protocol was used. In brief, for each SDS-PAGE experiment two gels were polymerized, loaded and run in parallel. The gel itself should constitute a focusing zone, where samples undergo isotachophoresis (ITP), and a separation zone. For the separation gel a 12% acrylamide gel was used. Therefor, 4.8 mL 30% acrylamide solution was mixed with 2.25 mL 2 M Tris-HCl (pH 8.8) and 4.83 mL H_2O. This solution was briefly mixed and an aliquot of 2 mL was transferred into a new flask. These 2 mL were mixed with 20 µL 10% APS and 5 mL TEMED and immediately transferred between the glass frames of the gel casting apparatus (Mini Protean Cell, Bio-Rad, USA). This was performed to give a polymer plug, which seals the gel casting stand. For the focusing gel, a 4% acrylamide gel was employed. Therefor, 1.066 mL 30% acrylamide solution was mixed with 1.0 mL 1 M Tris-HCL (pH 6.8) and 5.86 mL H_2O. Both, the separation and the focusing gel solutions were degased for 10 min under vacuum after preparation.

To start the polymerization of the separation gel, 50 µL 20% SDS, 45 µL 10% APS, and 8 µL TEMED were added to the 12% acrylamide solution. The solution was briefly mixed and transferred via pipette in between the glass plates of the gel casting apparatus. Each casting frame was filled with the separation gel till the solution was 2 cm below the top edge. To avoid the formation of a meniscus, 90% 2-propanol was poured on top of the solution. Polymerization was allowed to take place for ~1 h and was checked for completeness by slightly tilting the gel frames. Afterwards, the 2-propanol was removed with filter paper from the top of the separation gel. Next, the 4% acrylamide solution was mixed with 40 µL 20% SDS, 40 µL 10% APS, and 8 µL TEMED. The solution was filled in-between the glass frames till they were overfilled. A 10-well comb was placed between the glass plates and polymerization was allowed to take place for at least 30 min.

For sample preparation, 20 µL serum (~140 µg proteins) was mixed with 4 µL H_2O and 6 µL 5x Laemmli buffer, yielding in total 30 µL sample solution. To support the protein unfolding, all prepared samples were placed into boiling water for 10 min before loading onto the gels. In order to avoid band broadening, a blank was incorporated between the samples. Therefore, a blank solution containing 408 µL SB and 120 µL 5x Laemmli buffer was prepared.

Electrophoretic separation and protein staining

For electrophoretic separation gels were transferred from the gel casting stand to the electrode assembly (Power Pac Universal and Mini Protean Tetra System; Bio-Rad, USA). The apparatus was filled with electrode buffer and the combs were carefully removed. Samples were completely (all 30 µL) loaded onto the gels with at least one blank in-between and a molecular ladder (Precision Plus Protein™ Dual Color Standards) was also loaded onto each gel. Electrophoretic separation took place at 20 mA per gel and a restricted maximum voltage of 200 V. The run was stopped after the front marker (phenol blue) had migrated approximately 2 cm in the separation gel. The gels were then freed from the glass plates and placed into fixing solution overnight.

All steps of gel staining took place under mild shaking on a plate shaker (Shaking Plate 3016, GFL, Germany). First, gels were placed into washing solution for 10 min and then washed two times by placing them 5 min in water. To sensitize the protein bands, the gels were put into 0.02% $Na_2S_2O_3$ for 1 min and afterwards rinsed two times with water. For silver agglomeration, gels were placed 10 min into 0.1% $AgNO_3$ and afterwards shortly rinsed with

water. Subsequently, gels were placed in the developer solution till the proteins bands became visible (1-2 min). Thereafter gels were immediately put into stop solution and shacked for 5 min. The stained gels were photographed and kept in stop solution for storage or further processing. Figure 10Figure 11 shows a gel after silver staining. The cutting lines of the two fractions per sample are marked with black frames.
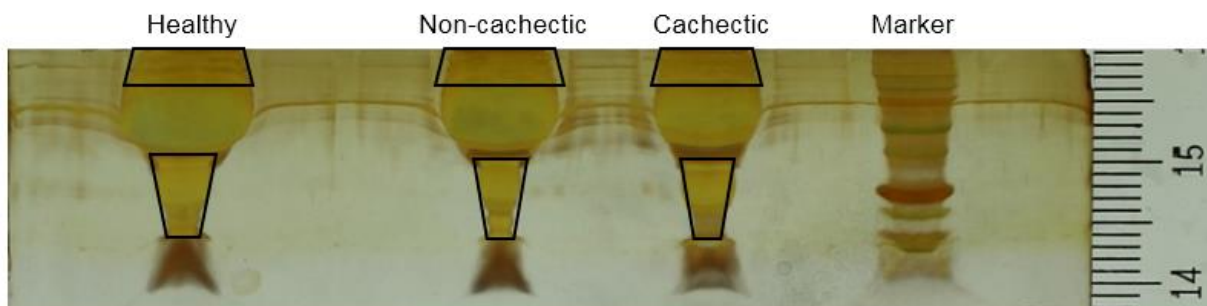

Figure 11: Stained gel of a sample set with indicated cutting lines
Picture of a silver stained SDS-PAGE of one sample set. The picture shows from left to right one healthy, one non-cachectic, and one cachectic sample. At the very right side the ladder is shown indicating the two fractions molecular weights. The cutting lines are shown by the black outlines.

Sample fractionation and in-gel digestion

For sample fractionation, the stained gel was carefully placed on a backlight and brought into a straightened form. Each serum sample under investigation was split into two fractions, one lighter then albumin (<70 kDa) and one heavier then albumin (>70 kDa). These fractions were clearly to distinguish from the slightly blue and very broad albumin band. Additionally, the protein ladder was used to ensure the correct fraction border and molecular weight (MW) cut-off. Each fraction was carefully cut out from the gel with a scalpel. The gel bands were then cut into roughly 1x1 mm pieces and each fraction was transferred into a 1.5 mL tube (Eppendorf, Germany).

For digestions, gel pieces needed to be de-stained. Therefore, 200 µL de-staining solution was added to each tube. Samples were vortexed (Vortex Genius, IKA, China) till the gel pieces turned colorless, shortly spun down (Centrifuge 5424, Eppendorf, Germany) and the solution was taken off via pipette. Next, 400 µL 25 mM ABC was added to the each tube. Samples were shaken for 10 min at 1400 rpm (Thermomix Comfort, Eppendorf, Germany), spun down and the supernatant was withdrawn. This was performed four times in total, alternating with 400 µL of 25 mM ABC and pure ACN. For reduction of the disulfide bonds, 200 µL of the 1 M DTT was given into each tube and samples were shaken for 30 min at

56°C. Afterwards, the tubes were spun down and the supernatant was removed. The gel pieces were washed twice with 25 mM ABC and ACN in the same manner as described above. For alkylation, 200 µL 500 mM IAA was added to each tube and the samples were incubated 30 min in the dark at 37°C. Thereafter, samples were spun down, the supernatant was withdrawn and samples were washed twice with 25 mM ABC and ACN. Next, gel pieces were brought to complete dryness, by placing the tubes 20 min at 40°C into the speedvac.

For enzymatic digestion, the dried samples were placed on ice and diluted TL solution (10 µL TL stock in 15 µL of cold 50 mM ABC) was added, giving an enzyme to protein ratio of approximately 1:20. After 15 min incubation on ice, the gel pieces had soaked up most of the TL solution and 25 mM ABC was added, till the gel pieces were completely covered with the solution (~20 µL). Digestion was allowed to take place overnight (~16 h) at 37°C. On the next day peptides were extracted from the gel pieces. Therefor, 40 µL 25 mM ABC was added, samples were vortexed, shaken for 15 min, spun down and the supernatant was transferred into a 0.6 mL siliconized tube. This procedure was repeated two times. The extraction procedure was then repeated again two times with 5% formic acid instead of 25 mM ABC to extract more hydrophobic peptides as well. All supernatants of a fraction were combined into a 0.6 mL siliconized Eppendorf tube and the samples were brought to complete dryness via vacuum centrifugation. The dried samples were stored at -20°C till further usage.

3.2.3.2 Serum Depletion and In-solution Digestion

For serum depletion, Pierce™ Top 12 Abundant Protein Depletion Spin Columns (Thermo Fisher Scientific, USA) were employed. The columns were stored at +4°C and brought to room temperature before use. For depletion, 7 µL of serum was put directly into the resin slurry of the depletion column. The column was capped again and softly shaken end-to-end by hand. Thereby it was controlled that the resin mix was freely moving and well mixing. Columns were placed into a rotator (RM Multi-1, Starlab, Germany) for end-to-end mixing over 60 min. In this time it was regularly (5-10 min) checked that the slurry moved freely to ensure maximum antibody/protein binding. Afterwards, the column closure caps were twisted off and the columns were placed into a 2 mL collection tube. Column caps were loosened and samples were centrifuged for 2 min at 1000 g. From each column approx. 500 µL of depleted serum flow-through was gathered and the protein concentration was determined by Bradford.

For digestion 250 µL of the depleted serum was placed onto a 3 kDa MW cut-off filter (Nanosep with Omega membrane, Pall, USA). The filter was conditioned before with 500 mL H$_2$O centrifuged through at 14,000 g for 15 min. In order to separate the protein content, the samples were centrifuged at 14,000 g till all liquid passed through the filter (~20 min). For reduction, 200 µL of 32 mM DTT were put onto the filter and mixed well with the pipette. Reduction was then allowed to take place for 30 min at 35°C on a thermal shaker (1000 rpm). Afterwards, the samples were centrifuged at 14,000 g till all liquid passed through (~30 min). The protein residue was washed by adding 200 µL 50 mM ABC and centrifuging again. For alkylation, 200 µL of 54 mM IAA were added onto the filter and the solution was well mixed with the proteins. Alkylation took place at 30°C for 45 min in the dark under constant shaking with 1000 rpm. The reaction mixture was then centrifuged, the residue was again washed with 200 µL 50 mM ABC and the filtrate was discarded. Now the filters were placed into new collection tubes and put on ice. For digestion, 5 µL of the TL stock solution was put onto the filter, giving an enzyme–to-protein ratio of 1:40. Next, 95 µL of cooled 50 mM ABC were added onto the filter and the solution was mixed with a pipette. Digestion was allowed to take place for 16 h at 37°C. Thereafter, samples were put on ice and 5 µL of TL stock solution (enzyme: protein 1:40) were added to each sample. Next, 45 µL of 50 mM ABC were added and the solution was well mixed with a pipette. The second digestion step was allowed to take place for 4 h at 37°C. Afterwards, samples were centrifuged at 14,000 g till all liquid has passed through (15-20 min). 50 µL of 50 mM ABC were added to each sample and samples were centrifuged again. The combined filtrates contained the extracted peptides, whereas the endopeptidase and intact proteins remained on the filter.

Peptides were cleaned-up using C18 spin columns (Pierce™ C18 Spin Columns, Thermo Fisher Scientific, USA). To wash the resin, columns were placed into a 2.0 mL tube and loaded with 400 µL of 50% ACN. Columns were centrifuged at 1500 g for 1 min and the flow through was discarded. The washing step was repeated and it was checked that all liquid had passed through before equilibration was started. For equilibration, the columns were loaded with 200 µL of 5% ACN (in 0.5% TFA) and centrifuged at 1500 g for 1 min. The equilibration step was repeated and the effluent was discarded. For sample binding, samples were acidified by adding 15 µL of 10% TFA, giving a final TFA concentration of approx. 1%. The acidified samples were loaded onto the column, centrifuged (1 min; 1500 g) and reloaded in the same manner. Samples were washed two times by adding 200 µL of 5% ACN (in 0.5% TFA) and centrifugation (1 min; 1500 g). After washing, the columns were transferred into

new 1.5 mL Eppendorf tubes. Peptides were eluted twice by loading 40 µL of 50% ACN (in 0.1% TFA) and centrifuge for 1 min at 1500 g. The combined filtrates were brought to complete dryness via vacuum centrifugation and the samples were stored at -20°C till further use.

In order to have reference samples to evaluate the sample fractionation techniques, also unfractionated serum was digested. Therefore, the serum was 100-fold diluted with H_2O and 35 µL of the dilutions were utilized for the in-solution digestion procedure as described above.

3.2.4 Shotgun LC-MS/MS utilizing a Q Exactive Orbitrap

For the shotgun measurements, an in-house protocol, initially developed for the analysis of cell supernatants[42], was utilized.

3.2.4.1 NanoLC Separation

For chromatographic separation, dried samples were reconstituted in 5 µL of the equimolar 10 fmol standard peptide mix and 40 µL of mobile phase A (98% H_2O, 2% ACN, 0.1% FA). The injection volume was set to 10 µL and the samples were separated on a Dionex Ultimate 3000 nano LC system coupled to the Q Exactive mass spectrometer via nESI source (Thermo Fischer Scientific, USA). For peptide concentration and further desalting, the samples were loaded onto a precolumn (2 cm x 75 µm C18 Pepmap100 precolumn, Thermo Fischer Scientific, USA). Therefore, a flow rate of 10 µL/min and 100% of eluent A was used. Peptide separation was then performed on a 50 cm x 75 µm C18 column (Pepmap100 analytical column, Thermo Fischer Scientific, USA) at a flow rate of 300 nL/min. An elution gradient was applied using 8-40% mobile phase B (80% ACN, 2% H_2O, 2% ACN, 0.1% FA) over 95 min. After each run, the system was flushed with 90% mobile phase B and re-conditioned to 100% mobile phase A.

3.2.4.2 MS/MS Data Acquisition

For data acquisition, data-directed-acquisition (DDA) on MS^2 level was used. MS^1 scans were recorded over the m/z range from 400 to 1400 m/z at a resolution of 70,000 (@200 m/z). For fragmentation, the 8 most abundant precursor ions were selected and MS^2 spectra were recorded at a resolution of 17,500 (@200 m/z) and saved as centroids. Fragmentation itself was achieved via HCD at 30% normalized collision energy. After fragmentation, the

corresponding precursor ions were excluded for 30 s from fragmentation triggering, by a dynamic exclusion list.

3.2.5 Targeted Analysis via nanoChip LC-MRM/MS

3.2.5.1 NanoChip LC Seperation

For LC-MRM/MS analysis, the dried samples were reconstituted in 30 µL (50 µL for unfractionated serum) of the 10 fmol peptide mix. For nanoChip LC, a 1260 Infinity Series HPLC system (Agilent, USA) coupled to the MS system via Agilent`s ChipCube was used. For the separation itself, a large capacity protein chip (G4240-62010) with a 160 nL enrichment column and a 150 mm x 75 µm separation column (5 µm ZORBRAX 300SB-C18, 30 Å pore size) was used. For peptide enrichment, 1 µL of the sample was injected and loaded onto the precolumn with 100% mobile phase A (97.8% H_2O, 2% ACN, 0.2% FA) at a flow rate of 5 µL/min in enrichment mode (backflush of the precolumn). For peptide separation, a 25 min gradient was applied starting with 8% eluent B (97.8% ACN, 2% H_2O, 0.2% FA). After 2 min, eleunt B is increased to 30% over 19 min followed by a flushing and conditioning step (overall 40 min runtime).

3.2.5.2 MRM Data Acquisition

Prior to all MRM measurements, the system was tuned for MS/MS in positive ionization mode and UNIT as resolution according to Agilent´s guidelines. The parameters gathered by this tuning were applied to the following MRM measurements. Only the peptide specific transition setting and the capillary voltage were changed according to the analytical question and the system performance. Capillary voltage was adjusted to spray performance based on visual inspection and ranged between 1750 and 1850 V. Peptide specific transition settings for scheduled and unscheduled measurements were gathered via Skyline (V3.1), which will be further outlined in the Bioinformatics section (see 3.3.2). The settings were exported into the MassHunter method editor, thereby it was checked that the cycle time is always 1300 ms, to ensure maximum reproducibility.

3.3 Bioinformatics

3.3.1 MaxQuant Label-free Quantification

For protein identification and quantification, the software package MaxQuant (V1.5.X) was used, including the Andromeda search engine and the Perseus statistical analysis tool. Therefor a well-established in house strategy was utilized.[42] In short, shotgun data gathered by the Q Exactive MS was inputted into MaxQuant and groups (e.g. cachectic) and replicates were assigned. Peptide spectra were searched against the SwissProt database (11.2014) for human taxonomy. Mass tolerances were set to recommended settings for Q Exactive instruments at the given resolutions. Carbamidation of cysteine was set as fixed modification and oxidation of methionine as variable modification. A minimum number of two peptides per protein (including one unique) was set together with a FDR of 0.01 on protein level. The match-between-runs feature was used to ensure as many identifications as possible.

Protein identification results of MaxQuant were further processed using Perseus. The proteins were filtered for known contaminates, reversed sequences and a minimum of three successful identifications overall biological groups (healthy, cachectic, non-cachectic). Missing LFQ-intensities were filled using Perseus` normal distribution filling and all LFQ-intensities were logarithmized to base 2. Statistical testing between the groups for regulated protein expression was performed using Perseus t-test feature with a two sided t-test for equal distributions.

3.3.2 MRM Assay Development

Based on the results of the sample preparation method evaluation, target panel selection and MRM assay development was only conducted based on depleted serum samples (reasons therefor will be further outlined in the R&D section).

3.3.2.1 Target Peptide Selection

For all peptides identified and assigned to a protein via MaxQuant, a spectral reference library was built in Skyline. From this library the peptide spectra assigned to the 96 candidate proteins gathered by LFQ (see 4.2 MaxQuant Results and Target Panel Selection) were extracted. Here only PTPs were accepted, which show no variable modification, no missed cleavages, contain no methionine and had 8 to 25 amino acids in length. Next, peptide

extracted ion chromatograms (XICs) on MS[1] and MS[2] level were manually inspected. Inspection on MS[1] level involved inspection of the right peak selection, based on co-eluting of all precursors. Further, identification triggers over all samples and technical replicates had to align within a certain time window. Additionally, a peptide should be identified only once over the whole chromatographic run (multiple identifications were only accepted for very broad peaks) and had to be identified in at least two of the three different sample types (cachectic, non-cachectic, healthy). Inspection on MS[2] level was based on number of assigned peaks, sequence coverage of the fragment ions and overall spectrum quality (e.g. noise). At last, a maximum of three most abundant peptides per protein were selected. Figure 12 shows an MS[1] XIC and the corresponding MS[2] mass spectrum of a peptide derived from monocyte differentiation antigen CD14. The precursor's isotopologues show consistent chromatographic peak shapes (A) and the product ion spectrum (B) provides good sequence coverage, so this peptide passes manual inspection. In the end a target panel of 188 peptides interfered from 93 proteins passed this inspection.
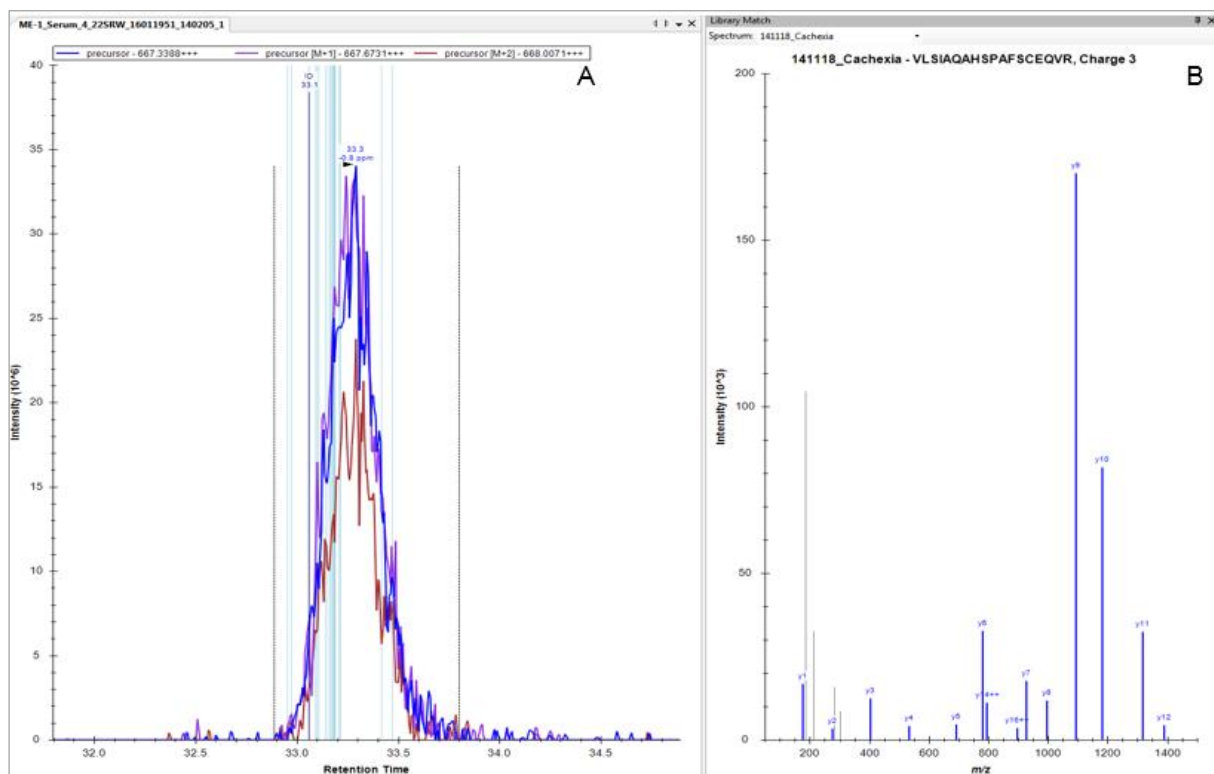
Figure 12: Peptide XIC and MS[2] product ion spectrum of a CHL1 peptide recorded by Shotgun MS
Figure A shows the MS[1] XIC of the [M]3+, [M+1]3+ and [M+2]3+ precursor ions for the peptide VLSIAQAHSPAFSCEQVR. The RT at which the peptide was identified in this measurement is indicated by the black line. Other identifications across all samples are shown by the light blue lines. The product ion spectrum (B) shows a high coverage of all theoretical ions.

For precursor selection of these 188 peptides, all precursor ions other than $[M+XH^+]^{x+}$ (monoisotopic precursor ions) were excluded. Next, all precursors that show interferences in the spectral library (e.g. inconsistent peak shape) were removed. From the remaining precursor ions, the one with the highest MS^1 intensity was selected for each peptide. Exceptions were made whenever two precursors did not differ strongly in their abundance. In such cases, the precursor ion with the lower charge state and thus with a higher m/z was selected to reduce possible interferences. In the next step, appropriate fragment ions were selected for the chosen precursor ions. Therefore, the five most abundant y-ions (based on the library) were selected for each peptide. The selection of y-ions is favorable whenever no labeled standards are available, since they are the preferred ion type formed by CID[43]. Transitions were exported from Skyline into a MassHunter compatible excel file using the Agilent export settings implemented in Skyline. Collision energies were calculated individually for each transition by Skyline using Agilent`s recommended QqQ 6490 settings (2 eV increments, slope 0.36, interception -4). Thereby, 940 transitions of 188 peptides (93 proteins) were sent for unscheduled MRM measurements.

3.3.2.2 Unscheduled MRM Measurements

For unscheduled MRM measurements, peptide precursors were sorted into three groups according to their maximum signal height in shotgun MS ($<10^6$; 10^6-10^7; and $>10^7$). Next, each peptide was assigned to the biological group (cachectic, non-cachectic, healthy), in which the highest signal was observed. According to this assignment in signal height, group specific transition lists were generated based on the transition list generated in Skyline. Group and sample type specific unscheduled MRM methods were generated using the MassHunter offline method editor. Therefore, transition settings were imported into the method file from excel, leaving the other parameters (e.g. gradient) untouched. The cycle time was set to 1300 ms and the number of concurrent transitions was manually set in order to ensure group specific dwell times. These dwell times were 100 ms for the group with the lowest signal intensities ($<10^6$), 50 ms for the group with moderate intensities (10^6-10^7), and 20 ms for the group with the highest signal intensities ($>10^7$). In order to achieve this, multiple method files were generated per signal group and sample type. In the end 13 sample type specific unscheduled MRM were generated for the 188 target peptides (940 transitions).

One biological sample of each type (cachectic, non-cachectic, healthy) that was also used for shotgun MS, was injected multiple times for unscheduled MRM measurements (see 3.2.5 Targeted Analysis via nanoChip LC-MRM/MS) using the sample specific methods. All MRM data gathered thereby was merged into on Skyline file containing all 940 target transitions.

3.3.2.3 Scheduled MRM Assay Development

For development of the scheduled MRM assay, peptide peak picking and transition selection had to be performed. Peaks were selected based on the following criteria: A co-eluting and consistent peak shape for at least three out of the five investigated product ions had to be present. Further, the library dot-product (dotp), which is a measure of the similarity between the acquired product ion spectrum and the library reference spectrum, had to be at least 0.8. Last criterion was the match between measured RT and calculated RT based on the Skyline SSRcalc 3.0. Here a correlation greater R= 0.9 had to be achieved. Figure 13 shows the gathered MRM product ion spectra for two peptides EATDVIIIHSK (A) and ALSIGFETCR (B). As illustrated in A, the y4 ion trace (red) shows a clear signs of interference, as its peak shape does not match the other four transitions. In B all five transitions are co-eluting and free from interferences. Next, three transitions were selected per peptide for the dynamic MRM assay, here always the three most intense product ions were selected. Exceptions were made whenever a selected transition showed clear signs of interference, like inconsistent peak shapes. Thereby, 372 transitions for 126 peptides interfered from 88 proteins and their corresponding RT could be gathered.
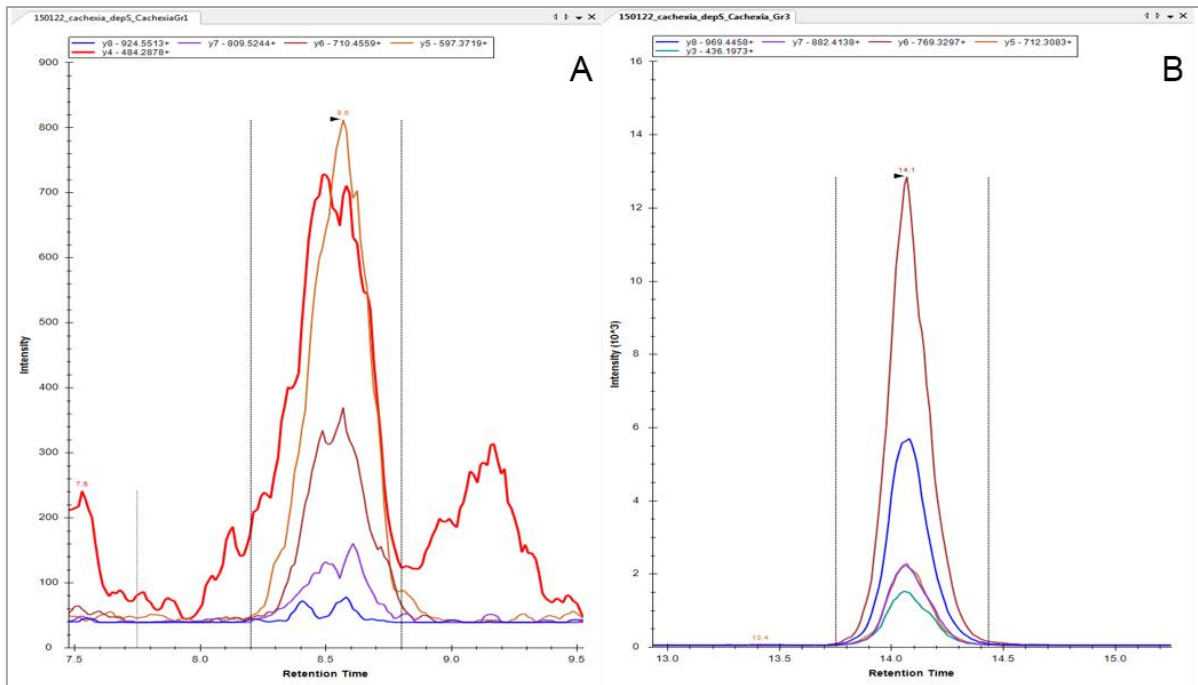
Figure 13: Transition and RT selection for the scheduled MRM assay development
Unscheduled MRM XICs for five transitions of EATDVIIIHSK (A) and ALSIGFETCR (B). The y4 transition of A shows clear signs of interference, because of its inconsistent peak shape compared to the other four transitions.

At last, the number of concurrent transitions over the whole run time was checked using the Skyline RT graph feature. These numbers should be no more than 60 at any given time, in order to ensure a minimum dwell time of 20 ms per transition. As this was not the case throughout the entire run time, the target panel was manually revised. This revision involved limiting the number of peptides per protein to two, removing peptides with insufficient signal intensities in all biological groups and removing proteins with low biological significance (based on literature and shotgun data). In the end, a dynamic MRM method for 92 peptides (276 transitions) interfered from 58 target proteins was achieved. To this method the in-house transitions and RTs for the four standard peptides were added in order to have internal calibrants. The method was implemented into MassHunter by exporting scheduled transitions settings from Skyline using a 3 min time window. This final dynamic MRM method was then used for the measurements of the clinical samples regarding to the analytical questions.

3.3.3 Scheduled MRM Measurements

Samples measured by the scheduled MRM assay were always injected in triplicates. The acquired data was imported into Skyline using the final MRM method Skyline file as starting point. Peaks were manually inspected regarding correct peak selection, interferences, and integration boundaries. In case of wrong peak selection, the correct peak was selected under the same criteria as described for unscheduled MRM. Transitions that showed signs of interferences were removed and incorrect integration boundaries manually adjusted. From this revised results, total peak areas (sum of all transition peak areas) were exported to excel. The utilization of total peak areas instead of a quantifier and two qualifier ions is preferred in proteomics. The reason therefor is that recent studies reviled, that the ratio in abundances of the different product ions is not constant over different concentration ranges[44]. All peak areas were normalized to the average peak areas of the four standard peptides, to compensate for variations in the system performance and ESI spray stability. Next, outliers of the triplicate measurements were removed based on Nalimov testing and average peak areas were calculated for each sample. Thereafter, the results were normalized to the used serum volume and dilution in order to achieve fully comparability. For protein results, the total peak areas of the corresponding peptides for each protein were summed up using excels pivot table feature. Further data evaluation was performed using common statistical tools implemented in excel, like student's t-test and conditional formatting.

4 Results and Discussion

4.1 Evaluation of the Serum Fractionation Methods

For evaluation of the serum fractionation methods, one sample per biological group was chosen, so one cachectic, one non-cachectic and one healthy serum. These serum samples were split into two aliquots each, from which one was fractionated via SDS-PAGE and one was depleted using Pierce depletion columns. The subsequently digested samples were then measured via shotgun MS and the data were analyzed in MaxQuant (see 3.3.1). The two fractions per sample gathered by the SDS-PAGE approach were injected separately for LC-MS analysis. Assignment to the biological sample was then achieved post acquisition using the fraction settings in the MaxQuant group parameters.

4.1.1 Number and Quality of Identified Proteins

To choose the better suited sample preparation method, the number of identified proteins and their quality (e.g. known impurities) was assessed for both sample preparation methods. Therefore, the number of identified proteins across the three biological groups was compared for the depleted and SDS-PAGE fractionated serum samples. MaxQuant was able to identify and quantify 452 proteins across all depleted samples, but only 298 across all SDS-PAGE fractionated samples. 210 proteins were identified for both sample treatment strategies. These results are illustrated in Figure 14.
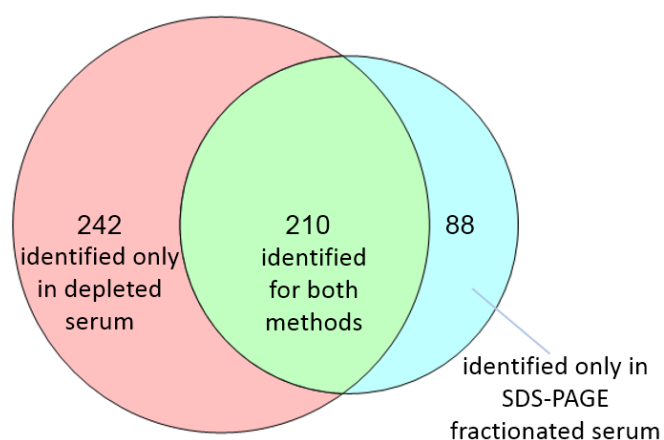

Figure 14: Protein identifications in depleted and SDS-PAGE fractionated serum
Venn diagram of the protein identifications achieved in depleted and SDS-PAGE fractionated serum. Significantly more proteins were detected in depleted serum compared to SDS-PAGE fractionated serum.

Next, the quality of the proteins that were identified by only one strategy was investigated. Proteins identified by both methods were excluded for this assessment, since they can also be evaluated later on in the target panel selection. One quality criterion was the number of immunoglobulins present in the data. Immunoglobulins show high individual and biological variations and can be expressed due to a multitude of reasons. Hence, they are not suited as marker proteins for a certain condition and have to be excluded from the panel. Furthermore, known impurities, like keratins or other skin derived proteins, were excluded together with proteins of low biological significance based on literature. From the 88 proteins identified only by the SDS-PAGE approach, 64 were excluded. From these 64 excluded proteins, 59 were immunoglobulins and 15 were either known impurities or low in biological significance. From the 242 proteins only identified in the depleted samples, only 35 were excluded, of which 27 were known impurities or low in biological significance and only 8 proteins were immunoglobulins.

These results point out, that the sample preparation via SDS-PAGE fractionation not only resulted in a lower number of identified proteins, but also led to identification of proteins with lower biological significance. Especially the high number of immunoglobulins present in the gel fractionated samples is a serious drawback. This was partly an expected result, since immunoglobulins cannot be excluded by gel fractionation due to their wide range of molecular weights. The higher number of impurities in the gel fractionated samples was surprising, but might stem from the method itself, as gel preparation, sample separation, gel staining, and band cutting are time-consuming processes, which make the method prone for impurities intake. During most of these processes, especially the cutting, the gel surface is exposed to the air and thereby impurities, mostly keratins from the skin, can contaminate the sample. Further, the high amounts of the immunoglobulins reduce the detection of low abundant proteins and decrease the dynamic range of detection of the measurement.

4.1.2 Throughput, Costs and Variability

Beside the results of the protein identification, also workload, sample throughput, costs, and variability of both sample preparation methods were compared. The fractionation by SDS-PAGE is a time consuming task and high in workload. Taking in account the gel preparation time (2 h), run time (1 h), staining procedure (2 h), and the cutting process (1-2 h), at least one day of lab work is needed. The depletion protocol, on the other hand, is

comparably short requiring 2 h processing time (1 h depletion + associated steps). The throughput of the depletion is also very high and only limited by the available instrumentation (e.g. size of the centrifuge). Additionally, more than one set of depletions can be done per day, accounting the short preparation time. However, the workload increases with the number of depletion performed in parallel. For SDS-PAGE, it is to mention that throughput can be easily increased by maintaining the main steps. The usage of larger gels allows loading more samples in parallel. This significantly increases the throughput by keeping the workload of gel preparation, electrophoresis and staining procedure constant. However, the workload of the sample preparation and cutting process would massively increase. The digestion procedures of both methods (in-gel or in-solution) are similar in workload and throughput and do not have major differences.

In terms of cost per sample, SDS-PAGE outperforms the depletion columns. Chemicals for the gel preparation are rather cheap and multiple samples can be fractionated on one gel. Further solutions incorporated in the staining process can be used for multiple gels and are also not very expensive. Depletion on the other hand is with 30 € per sample a rather expensive method. These high costs stems from the incorporation of highly specific antibodies. The usage of these highly specific antibodies combined with a standardized and vendor validated procedure makes depletion low in variability (also demonstrated in 4.3 Scheduled MRM Assay and Method Validation). The gel fractionation on the other hand is higher in variability. Firstly, not all SDS-PAGEs behave the same way under the electrophoretic separation leading to slightly different separation zones. Secondly, the nature of the fractionation process and the manual cutting of protein bands add a certain level of variability.

4.1.3 Summary of Serum Fractionation Method Evaluation

To conclude, the use of depletion columns leaded to a higher number of identified proteins and the proteins were better in terms of biological significance (e.g. no antibodies). Further, the throughput of the depletion is higher; depletion adds less variability to the sample preparation and is not as prone for contamination as the SDS PAGE. SDS PAGE in contrast is cheaper and offers the possibility for long-time storage. For the aim of this work, high-throughput accurate protein quantification should be achieved. Further possible clinical applications and automation should be considered. Here serum depletion using the Pierce

Top12 Depletion Spin Columns is clearly to prefer over the gel fractionation. Depletion was selected for sample preparation for all following steps, like target panel development and patient measurements. Nevertheless it should be mentioned that for research purposes and different analytical questions, SDS PAGE still offers a high resolution combined with a high level of robustness and ease of use.

4.2 MaxQuant Results and Target Panel Selection

For the identification and selection of possible altered proteins in cachexia, three samples per biological group were chosen. From each non-cachectic patient always the earliest available sample time-point was used in order to exclude possible cachectic developments in the late disease stage. For the cachectic patients always the latest available sampling date was used, as it was assumed that the cachectic outcome is strongest in the late disease stage. All nine samples were prepared in parallel by the depletion and in-solution digestion procedure described in the method section (see 3.2.3.2). Two technical replicates of each sample were measured by shotgun MS (see 3.2.4) and the data were analyzed via MaxQuant (see 3.3.1).

For statistical data evaluation (t-test), the MaxQuant results were imported into Perseus and further processed. Each replicate measurement was assigned to its respective biological group (cachectic, non-cachectic, or healthy). T-tests were performed for all three groups against each other and results were exported into excel (see 3.3.1). Protein alterations were accounted as significant, for p-values $p \leq 0.05$ and a minimum fold-change of two. 65 proteins were found to be significantly altered in non-cachectic patients compared to healthy donors. In cachectic patients, 136 proteins were significantly regulated in comparison to healthy protein levels, of which 40 had also been found to be altered in non-cachectic patients. When testing cachectic against non-cachectic patients, 80 proteins showed significant regulations, from which 57 were also found to be altered when comparing cachectic against healthy patients. Nine of these 57 shared proteins also show alterations in non-cachectic patients. These results are illustrated in the Venn diagram shown in Figure 15.
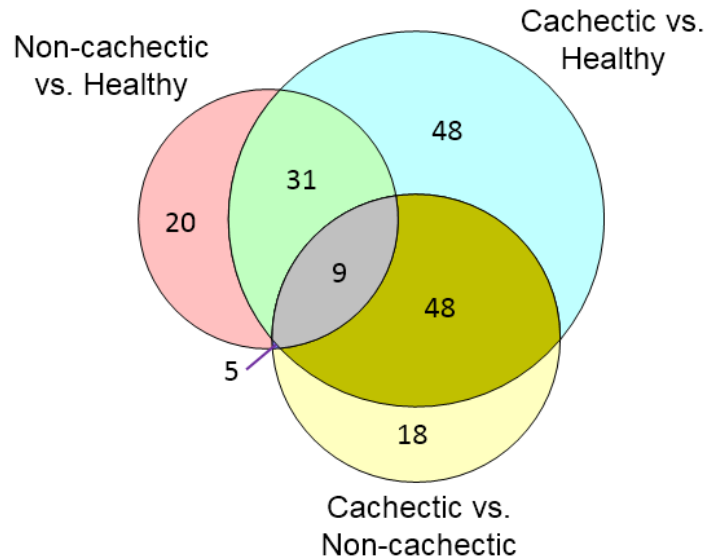
Figure 15: Protein expression regulations between the different biological groups
Venn diagram of significant protein expression regulations between the different biological groups found by MaxQuant LFQ and statistical evaluation in Perseus.

Based on these results, a panel of target proteins for MRM assay development was selected. Proteins that were regulated only in cachectic (compared to non-cachectic and healthy) patients were accounted most promising. Further, proteins that were found to be strongly up- or down-regulated in cachectic patients (>4-fold change) were accounted stronger. In such cases, even p-values slightly higher than p≤ 0.05 were accepted. All proteins found to be strongly or exclusively regulated in cachectic patients were then investigated for biological plausibility. Here, immunoglobulins were excluded, since they show high inter-individual expression variations. Furthermore, known contaminants (e.g. keratin-like proteins) were removed. Also proteins involved in the blood coagulation process were withdrawn, since differences might stem from the blood clotting used in the serum preparation. From the remaining panel, proteins with unknown or controversial function were double checked with literature and excluded if no meaningful relation to cachexia could be found. Based on these restrictions, the MaxQuant results leaded to a panel of 96 candidate proteins, displayed in Table 8.

Table 8: Candidate proteins for target panel

Candidate Proteins		
4F2 cell-surface antigen	Fibronectin	Peroxiredoxin-1
Afamin	Follistatin-related protein 1	Phospholipid transfer protein
Alpha-1-acid glycoprotein 1	Galectin-1	Plasminogen activator inhibitor 1
Alpha-1-acid glycoprotein 2	Galectin-3-binding protein	Platelet glycoprotein Ib alpha chain
Aminopeptidase N	Gamma-glutamyl hydrolase	Prohibitin
Angiopoietin-related protein 3	Glutathione S-transferase omega-1	Proteasome subunit alpha type-5
Apolipoprotein A-IV	Glutathione synthetase	Proteasome subunit alpha type-6
Aspartate aminotransferase, cytoplasmic	Heparanase	Proteasome subunit alpha type-7
ATP synthase subunit beta, mitochondrial	Hepatocyte growth factor receptor	Protein DJ-1
Beta-Ala-His dipeptidase	Hydroxymethylglutaryl-CoA synthase, mitochondrial	Protein S100-A8
Cadherin-2	Insulin-like growth factor-binding protein 1	Protein S100-A9
Cadherin-related family member 2	Insulin-like growth factor-binding protein 2	Protein Spindly
Calmodulin	Insulin-like growth factor-binding protein 5	Receptor-type tyrosine-protein phosphatase eta
Calumenin	Intercellular adhesion molecule 1	Receptor-type tyrosine-protein phosphatase zeta
Carbamoyl-phosphate synthase [ammonia], mitochondrial	Interleukin-6 receptor subunit beta	Retinol-binding protein 4
Cartilage acidic protein 1	Laminin subunit beta-1	Scavenger receptor cysteine-rich type 1 protein M130
Catenin beta-1	Latent-transforming growth factor beta-binding protein 1	Selenoprotein P
Cathepsin D	Leukocyte elastase inhibitor	Serpin B3
CD44 antigen	L-lactate dehydrogenase A chain	Serum amyloid A-1 protein
CD59 glycoprotein	L-lactate dehydrogenase B chain	Serum amyloid A-2 protein
Chondroitin sulfate proteoglycan 4	L-selectin	Serum amyloid A-4 protein
Coagulation factor XIII A chain	Macrophage colony-stimulating factor 1 receptor	Serum amyloid P-component
Collectin-11	Macrophage mannose receptor 1	Serum paraoxonase/lactonase 3
Complement C4-A	Matrix metalloproteinase-9	SH3 domain-binding glutamic acid-rich-like protein 3
Complement factor H-related protein 1	Melanocyte protein PMEL	Tetranectin
Complement factor H-related protein 4	Membrane primary amine oxidase	Thymosin beta-4
C-reactive protein	Metalloproteinase inhibitor 1	Transforming growth factor-beta-induced protein ig-h3
D-dopachrome decarboxylase	Monocyte differentiation antigen CD14	Transmembrane glycoprotein NMB
Dystroglycan	NAD kinase	Transthyretin
Electron transfer flavoprotein subunit alpha, mitochondrial	Neural cell adhesion molecule 1	Trem-like transcript 1 protein
Elongation factor 1-alpha 1	Neural cell adhesion molecule L1-like protein	Tryptophan--tRNA ligase, cytoplasmic
Ferritin light chain	Peptidyl-prolyl cis-trans isomerase A	Voltage-dependent anion-selective channel protein 1

4.3 Scheduled MRM Assay and Method Validation

As described in the Bioinformatics section (see 3.3.2), a scheduled MRM assay for 58 target proteins was developed. These target proteins are displayed in Table 9.

Table 9: Target proteins of the final MRM assay

MRM Target Proteins		
4F2 cell-surface antigen heavy chain	Fibronectin	Neural cell adhesion molecule L1-like protein
Alpha-1-acid glycoprotein 1	Follistatin-related protein 1	Phospholipid transfer protein
Alpha-1-acid glycoprotein 2	Galectin-3-binding protein	Platelet glycoprotein Ib alpha chain
Aminopeptidase N	Gamma-glutamyl hydrolase	Proteasome subunit alpha type-6
Angiopoietin-related protein 3	Glutathione S-transferase omega-1	Protein S100-A8
Apolipoprotein A-IV	Glutathione synthetase	Protein S100-A9
Aspartate aminotransferase, cytoplasmic	Heparanase	Receptor-type tyrosine-protein phosphatase eta
Beta-Ala-His dipeptidase	Insulin-like growth factor-binding protein 2	Retinol-binding protein 4
Cadherin-2	Intercellular adhesion molecule 1	Scavenger receptor cysteine-rich type 1 protein M130
Calumenin	Lactate dehydrogenase A	Selenoprotein P
Cartilage acidic protein 1	L-lactate dehydrogenase B chain	Serum amyloid A-1 protein
Cathepsin D	L-selectin	Serum amyloid A-2 protein
CD44 antigen	Macrophage colony-stimulating factor 1 receptor	Serum amyloid P-component
Chondroitin sulfate proteoglycan 4	Macrophage mannose receptor 1	Tetranectin
Coagulation factor XIII A chain	Matrix metalloproteinase-9	Transforming growth factor-beta-induced protein ig-h3
Collectin-11	Melanocyte protein PMEL	Transthyretin
Complement C4-A	Membrane primary amine oxidase	Trem-like transcript 1 protein
Complement factor H-related protein 1	Metalloproteinase inhibitor 1	Tryptophan--tRNA ligase, cytoplasmic
C-reactive protein	Monocyte differentiation antigen CD14	
Ferritin light chain	Neural cell adhesion molecule 1	

Before the MRM assay could be applied for the analysis of a multitude of patient samples, the method was evaluated. Method evaluation involved MRM assay specific parameters, such as chromatographic stability and signal reproducibility, as well as variability of the whole workflow (sample treatment & measurement).

4.3.1 Evaluation of the MRM Assay

For evaluation of the MRM assay specific parameters, a depleted and digested serum sample of a cachectic patient was injected as triplicate. Peak shapes were manually revised and the peptide signals were exported and normalized as described in the Bioinformatics section (see 3.3.3). The following exceptions were thereby made; outliers from the triplicate measurements were not removed and peak areas were not normalized on standard peptide areas. This was done, in order to get a better insight into measurement variations caused by the LC-MS system. 89 of the 92 target peptides could be asses, whereas three were below the limit of detection (LOD). The CVs of the peptide peak areas were calculated and plotted against the average RT, as shown in Figure 16.
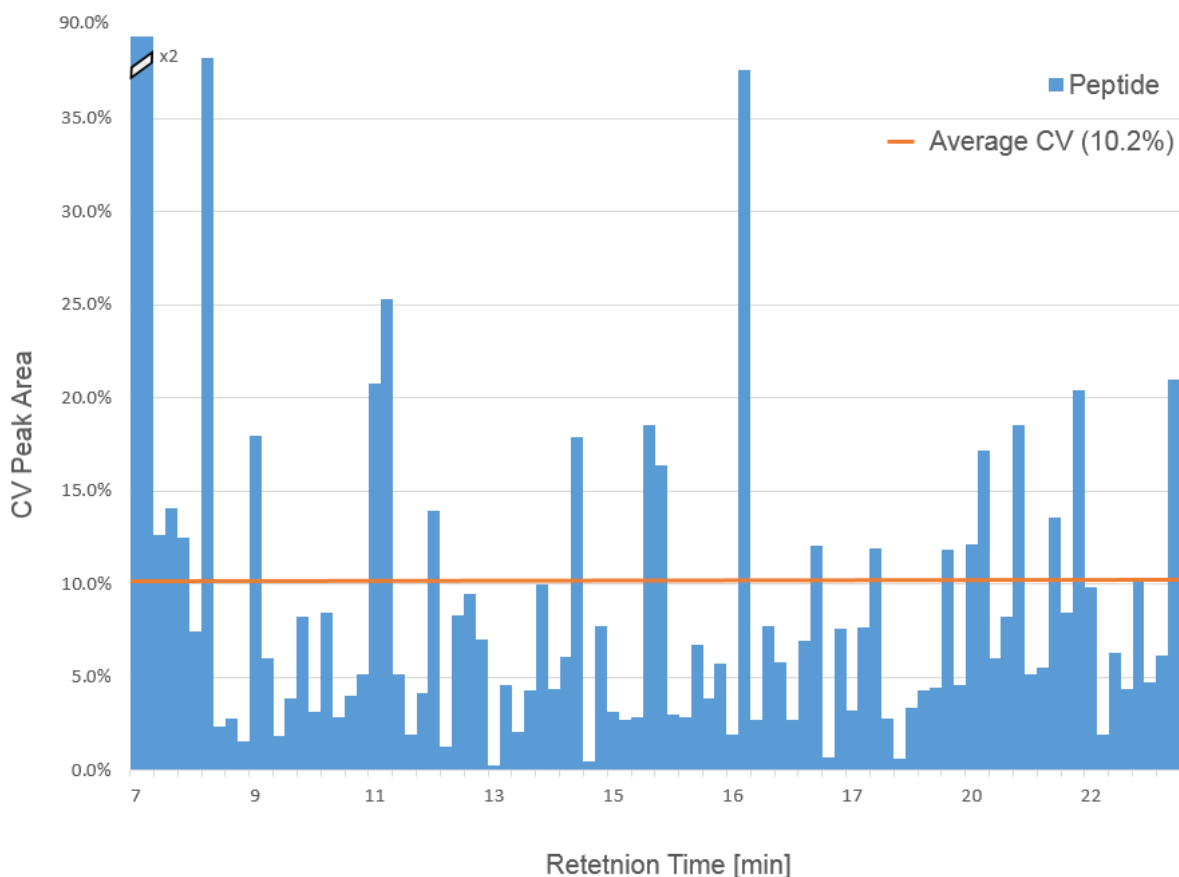

Figure 16: CVs of the peptide peak areas between 3 injections for 89 peptides

The peptide CVs were calculated based on triplicate injections and plotted against their average RT. Notably, early eluting peptides (<8 min) show unusual high CVs.

As the Figure 16 illustrates, an average CV of 10.2% was achieved for the triplicate measurement. CVs are evenly distributed over the whole chromatographic range, with exception for the early eluting peptides. Peptides eluting before min 7.5 of the chromatographic run show unusual high CVs (up to 80%). The reason therefore is the ESI spray instability at the beginning of the runs. The strong fluctuations in the spray stability before minute 7 of the run are caused by the high water content in the mobile phase. In order to improve spray stability, a shallower gradient could be used in the beginning of the run or a higher capillary voltage could be applied. The usage of a shallower starting gradient would give the spray more time to stabilize, but may also lead to peak broadening. The usage of a higher capillary voltage increases spray stability, though the ESI emitter tip of the chip could get damaged and chip life-time would decrease. As only two to three peptides elute before min 7.5, the spray stability was not further addressed. The vast majority of the peptides shows very low CVs and the chromatographic stability of the RTs was with an average CV of ± 0.12 min also very high. These results point out, that the developed MRM assay shows a high degree of stability and reproducibility.

4.3.2 Evaluation of the Serum Depletion

To evaluate the efficiency and reproducibility of the depletion procedure, each three aliquots of one healthy and one cachectic serum sample were depleted and digested in parallel. As reference, three aliquots of the same samples were also digested without prior depletion. For this sample preparation comparison (depleted vs. Undepleted serum), two different biological samples (cachectic and healthy) were used to demonstrate that depletion not only works in healthy serum, but also in the strongly altered matrix of cachectic serum. For MRM measurements, a modified scheduled assay was developed. This assay contained a subset of the target peptides from the MRM assay spread across the entire chromatographic range. In addition 22 peptides interfered from the following nine proteins; alpha-1-acid glycoprotein 1, alpha-1-antitrypsin, alpha-2-macroglobulin, apolipoprotein A-I, apolipoprotein A-II, fibrinogen alpha chain, haptoglobin, serotransferrin. These proteins should get removed to at least 95% by the depletion columns. For the three missing proteins; IgA, IgG, and IgM, no sufficient peptides could be found in the depleted serum. This demonstrates the high efficiency at which immunoglobulins are removed by the depletion columns.

To assess the efficiency at which the nine remaining depletion target proteins are removed from the serum, the signal intensities in depleted serum are compared to undepleted serum. Therefore, the measured peptide areas were normalized to the area of the standard peptides and to the used serum volume. Depletion efficiency was calculated based on normalized peak area comparison on protein level. The observed depletion efficiency for healthy and cachectic samples is shown in Figure 17.
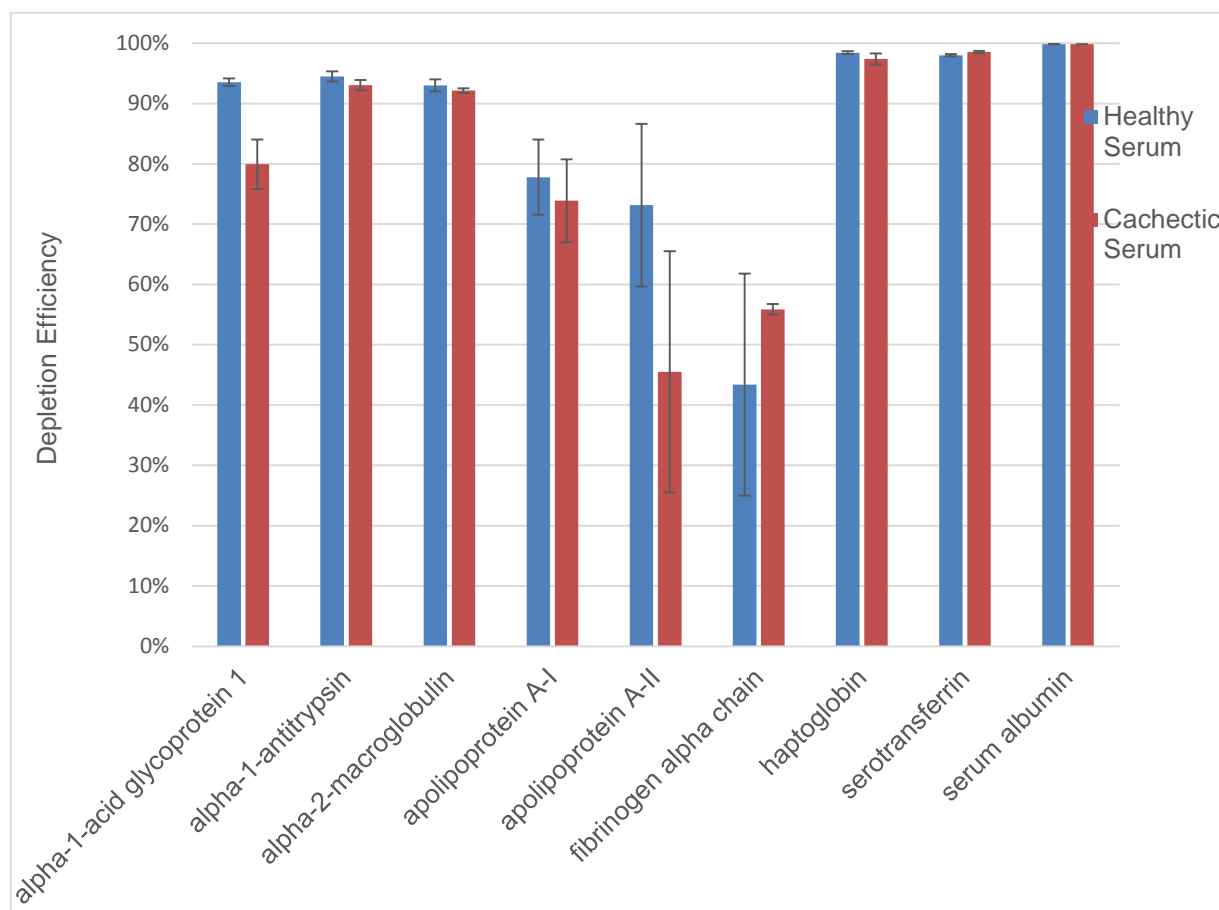

Figure 17: Depletion efficiency of the Pierce top12 spin columns
The efficiency, at which the depletion columns remove the target proteins, is shown for a healthy (blue) and a cachectic (red) sample.

As the figure illustrates, the target proteins were removed on average to 85.7% in the healthy and to 81.8% in the cachectic sample. Thereby an average CV of 4.6% for healthy serum and 3.8% for cachectic serum was observed between the three technical replicates. The significantly higher CVs observed for apolipoprotein A- II and fibrinogen are due to two reasons. First, for both proteins only peptides with very low signal intensities (slightly above LOD) could be found. As low signals contain a higher amount of noise, they always show

higher CVs. Second, for the quantification of fibrinogen only one peptide could be utilized. Protein results based on single peptide quantification often show higher variations, since no averages between multiple results can be formed. In summery it can be said, that the depletion of the target proteins shows a high efficiency (average ~84%) and low variation (average CV 4.2%). No significant differences were observed between healthy and cachectic serum.

Next the reproducibility of the depletion was assessed for both sample types. Therefor standard derivations between the technical triplicates were calculated for all measured peptides in the depleted and undepleted samples. This was done for each peptide present in the evaluation assay, the measured CVs are illustrated by a Tukey-boxplot in Figure 18.
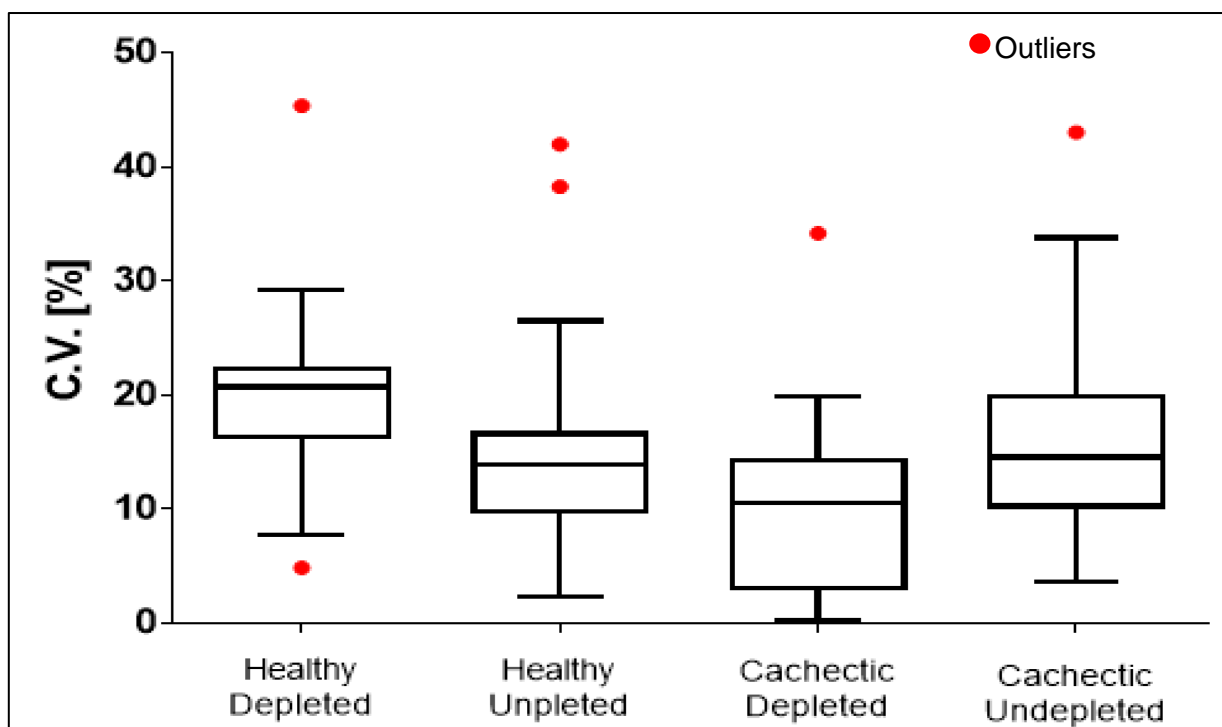

Figure 18: Tukey-boxplot of technical CVs with and without depletion
The CVs for all peptide peak areas between the three technical replicates of depleted and undepleted serum are illustrated by a Tukey-boxplot. No significant increase of the CVs could be observed, when sample were depleted prior to digestion. This is applicable for the healthy as well as for the cachectic serum sample.

As shown in the boxplot diagram, an average CV of 17% was observed throughout all measurements. This is 7% more compared to the 10.2% CV when only account MRM measurements variations. Though these result are not surprising, taking in account that the whole sample pre-treatment and digestion process adds certain variability to the results. For

the depletion process itself it can be said, that no significant increase in the CVs was observed when samples were depleted prior to digestion. These results point out, that the depletion process is highly reproducible and low in technical variation. Furthermore, signal enhancement and protein recovery achieved by depletion was evaluated. Signal enhancement is based on the fact that a higher serum volume can be loaded onto the chromatographic column after depletion. Thus, higher peptide signal intensities and greater peak areas are expected. On the other hand, depletion leads to a certain unselective loss of proteins, since many, especially small and signaling proteins are often bound to carrier proteins, like albumin. To calculate the signal enhancement and recovery, peptide signals were normalized to column protein load and to the used serum volume. Thereby, a theoretical signal enhancement of 10-fold was expected, since 10 times more depleted serum could be loaded on column compared to undepleted serum. The recovery is here defined as the ratio between the observed signal enhancement and the theoretical signal enhancement of 10-fold. For example, if a protein showed 7-fold signal enhancement in depleted serum, its recovery after depletion was 70%. The observed signal enhancement and recovery are displayed in Figure 19 for a selected subsets of high to mid abundant proteins.
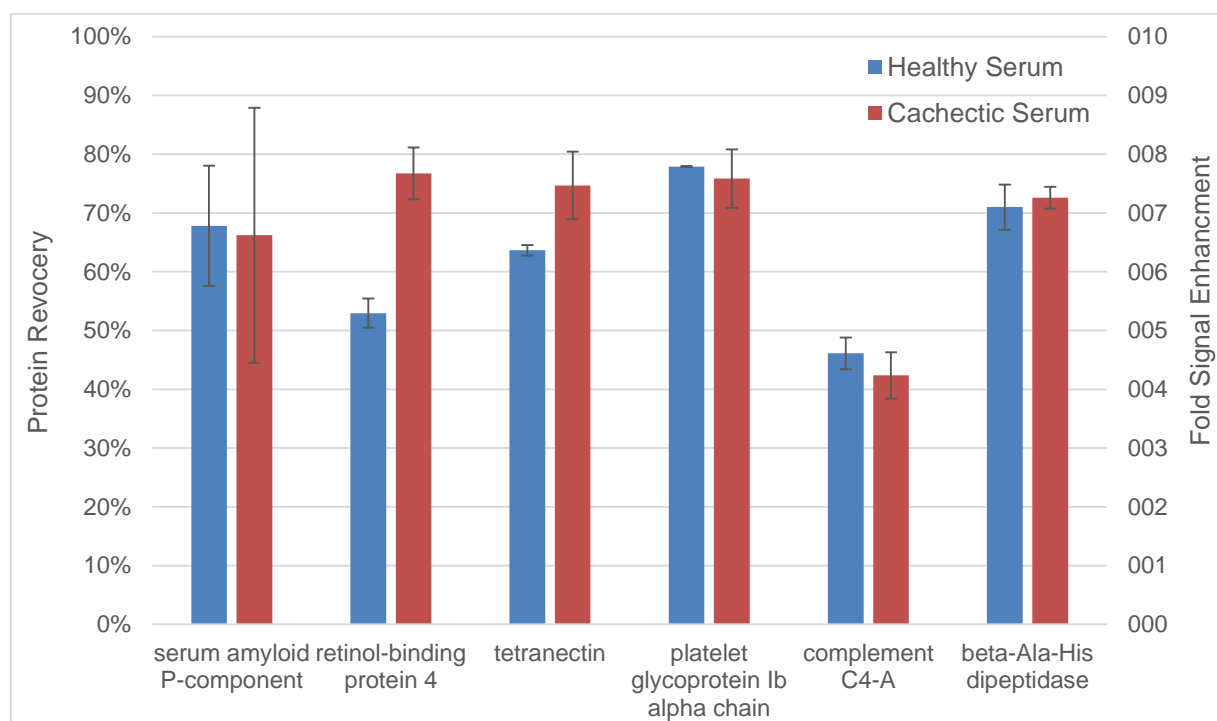

Figure 19: Protein recovery and signal enhancement in depleted sera
Signal enhancement by depletion and the thereof resulting protein recovery is shown for a selected subset of proteins. Error bars indicate the error for the recovery in percent.

On average a signal enhancement of 7.2-fold (72% protein recovery) was observed after depletion. As also shown in Figure 19, no significant variations in recovery were observed between healthy and cachectic serum.

4.3.3 Summary of the MRM Assay Development and Method Validation

A highly reproducible and precise MRM assay (average 10.2% measurement CV) was developed for the quantification of 58 target proteins in human serum. Thorough evaluation of the sample pre-treatment and digestion procedure demonstrated, that the utilized depletion is robust and stable throughout different biological sample types (healthy & melanoma serum) Further, depletion enabled an average 7.2-fold signal enhancement compared to undepleted serum. The evaluation of the complete workflow (including sample treatment and LC-MRM analysis) showed a precise protein quantification with an average CV < 25%. This was reached without the usage of SIS peptides, revealing the high accuracy in the developed MRM method and the highly standardized sample treatment.

4.4 Investigation of Serum Protein Alterations

4.4.1 Characteristic Serum Proteome Alterations in Cachectic Patients

For the investigation of serum proteome alterations caused by cancer cachexia, the developed MRM assay was applied to a subset of clinical patient samples. Three samples for each cohort (cachectic, non-cachectic and healthy) were chosen. Here the same strategy for sample selection was applied as for the shotgun measurements (latest time point for cachectic, earliest for non-cachectic patients). All nine samples were depleted and digested in parallel to reduce variability in the sample preparation process. The gathered, revised, normalized and outlier-corrected peptide peak areas, protein intensities were calculated. Protein intensities were logarithmized and differences in protein levels between the groups were calculated. To test, if a difference in the protein level between two groups was significant, a student's t-test was used. Therefor, a two sided t-test with homoscedastic variance was performed. A fold-change ≥ 2 with a p-value of $p \leq 0.05$ was accepted as significant.

In Figure 20, differences in protein expression between the biological groups are indicated by the color of cells. The darker the green tone becomes, the less is the up regulation; the

brighter the green tone becomes, the higher is the up regulation. As described above, the minimum accepted significant protein regulation was 2-fold, which is indicated by the darkest green tone. Black indicates no significant change in protein expression, either due to less than 2-fold expression difference or because of insufficient p-value. Red indicates down regulation of the protein expression in the investigated group. For red, only one tone is used indicating 2 to 5 fold down regulation, since no stronger down regulation have been found. The whole specific color code for the protein regulation is indicated on the upper right side of the figure.

In the very right column in Figure 20, proteins were further clustered according to their regulation into high specific cachexia regulation indicated by dark green and low specific cachexia regulation indicated in pale green. High specific cachexia alterations only were found to be regulated in cachectic patients compared to healthy as well as non-cachectic patients, but were not or opposite regulated comparing non-cachectic to healthy patients. This means, that these protein alterations are most probably induced by the cachexia itself and not by the melanoma. Low specific regulations show alterations in cachectic patients but also in non-cachectic and only differ in their outcome. This means that these protein alterations most probably stem from cancer itself, but are much stronger during cachectic outcome. However, these results have to be seen very critically, since most cachectic patients also show further cancer progression. So many of these stronger regulations are most probably induced by the cancer progression and might not be directly linked to cachexia. Further cluster are cancer or melanoma induced regulations indicated in yellow and unspecific findings in red. The yellow proteins show alterations in all cancer patients when compared the healthy donors, but no significant difference between cachectic and non-cachectic patients. So, these regulations stem from the tumor and inflammation process itself and not by any chance from the cachexia. Unspecific findings were proteins, for which no significant changes in expressions were found between the groups. Of the 58 candidate proteins, 13 high specific, 23 low specific regulated proteins and 9 tumor induced alterations were found. 13 Proteins showed no significant or meaningful changes in expression between the biological groups.

Figure 20: Heat map of protein expression between different groups
Proteins were clustered according to their regulations into high specific cachexia marker (dark green), low specific marker (pale green), tumor and inflammation marker (pale yellow) and unspecific findings (pale red).

From the 13 regulated high specific cachexia-induced protein alterations, 12 proteins were found to be up regulated. These proteins were further classified according to their function and biological significance. Here first to mention are intercellular adhesion molecule 1 (ICAM-1), neural cell adhesion molecule L1-like protein, and transforming growth factor-beta-induced protein ig-h3. All three play an essential role in cell adhesion and migration, mainly of leucocytes.[45] Their up-regulation in cachectic patients can be caused by the invasion of bacteria through the dysfunctional gut barrier.[1] These regulations can be a sign for tumor metastasis, since also here cell adhesion processes are triggered. Moreover, ICAM-1 is known to be expressed in inflammation processes, which also are present in tumor patients and especially accelerated in cachectic ones. Receptor-type tyrosine-protein phosphatase eta (R-PTP-eta) is also a regulator in cell adhesion. Further, it is involved in cell proliferation and growth. It is most strongly expressed in macrophages during inflammation and also shows some tumor suppressor activity.[46] Complement C4-A and scavenger receptor cysteine-rich type 1 protein M130 (sCD163) are part of the immune response. They are involved in the clearance of plugs and aggregates from the body. Complement C4-A is thereby involved in complement pathway and enhances the solubilisation of immune aggregates. The clearance of haptoglobin plugs is triggered by sCD163. The formation of these platelet plugs is a complex pathway in which also platelet glycoprotein Ib alpha chain (GPIbA) is involved. GPIbA shows also up-regulation in cachectic patients and is beside the plug formation also involved in cell adhesion processes. For the cause of platelet plug formation is the up regulation of trem-like transcript 1 protein (TLT-1) to mention. TLT-1 is a cell surface immune receptor located on the platelets[47]. Its up-regulation in blood may be caused by receptor shading during plug formation.

Next to cell adhesion, pro-inflammatory, and immune responsive proteins, the anti-inflammatory glutathione synthetase (GSH-S) was found up regulated. The expression of GSH-S is triggered by oxidative stress and the release of free amino acids. Both is observed in cachectic patients, especially the release of free amino acids due to the muscle degradation. GSH-H acts then as proteinase inhibitor and also catalyzes the formation of glutathione. Glutathione acts as radical scavenger and is involved in the cell metabolism. Also glutathione S-transferase omega-1, another enzyme involved in the glutathione metabolism was found up-regulated. However, the alteration of glutathione S-transferase omega-1 levels was present in all cancer patients and only stronger in cachectic ones. Phospholipid transfer protein is a lipid transporter and also involved in the formation of the high-density lipoprotein

(HDL) particles. It is thereby involved in the phospholipid transport from the WAT to the liver and other body parts.[48] Its up-regulation may be caused by the increased lipolytic activity observed in cachectic patients. Metalloproteinase inhibitor 1 and cathepsin D are the last two proteins found up-regulated only in cachectic patients. They both play a role in cell breakdown and apoptotic process regulation. Metalloproteinase inhibitor 1 thereby inhibits the function of metalloproteinase 1 to 3 and 7 to 13 by forming inactive complexes. Cathepsin D on the other hand is an acid protease active in intracellular protein breakdown. It is known to be involved in the pathogenesis of several diseases (e.g. breast cancer) and may also be an indicator for cachectic development. Transthyretin is the only protein which was found to be down-regulated in the high specific cachexia regulations panel. Transthyretin is a known thyroxine transporter.[49] It transports thyroxine from the blood through the blood-brain barrier and is thereby involved in the protein, fat, and carbohydrate metabolism.[49] Therefore, the decreased transthyretin expression in cancer patients may be a promotor for the metabolic changes during cachexia. Further, it might be the trigger for the insulin resistance observed in cachectic patients and crucial for the loss of appetite.

From the 23 low specific cachexia-induced protein alterations, only up-regulations were observed in tumor patients. From these first to mention is C reactive protein (CRP), a major acute-phase protein. CRP is part of the immune response and elevated in nearly all inflammatory processes. It is one of the most conserved plasma proteins and highly complex in its biological function. The primary role of CRP is to activate the complement pathway to degrade dying cells and bacteria. Further, it has been shown to interact with interleukins and thereby enhancing inflammatory processes. CRP is also involved in many other inflammation regulatory processes in humans.[50] There are many studies that suggest CRP levels as biomarkers for a multitude of diseases, ranging from cardiovascular diseases (CVDs) to various cancer types. Also for cancer progression and cancer cachexia, CRP was claimed to be a prognostic marker.[1, 51] However, the expression of CRP is caused by a multitude of inflammatory processes and can be linked to various diseases. Hence, CRP levels often failed to be used a biomarker for a certain condition. In this work, an up-regulation of CRP was observed in all cancer patients and therefor it cannot be exclusively linked to cachexia. Next to this major acute-phase protein, also alpha-1-acid glycoprotein 2, another acute-phase protein was found up-regulated in the low specific marker panel.

Many proteins which were present in the low specific cachexia induced alterations, are involved in the anti-bacterial defense. Macrophage colony-stimulating factor 1 receptor, macrophage mannose receptor 1, collectin-11 and monocyte differentiation antigen CD14 are here to mention. Macrophage colony-stimulating factor 1 receptor plays an important role in immunity and promotes the release of pro-inflammatory chemokines and superoxide species.[52] Macrophage mannose receptor 1 is a receptor for bacteria and mediates endocytosis of glycoproteins by macrophages. The same accounts for monocyte differentiation antigen CD14, which response to the presence of certain lipopolysaccharides (LPSs) present on bacterial shells.[53] Collectin-11 binds to various LPSs and thereby guides the macrophages to the bacterial invasion sides. These findings are a strong indicator for bacterial invasion observed in cancer patients. This may be caused by the barrier dysfunction of the gut and thereby enhances the ongoing inflammation processes. Beside these bacterial defense proteins, again proteins which are involved in cell migration and adhesion are present. Here to mention are membrane primary amine oxidase, galectin-3-binding protein, and CD44 antigen. Galactin-3- binding protein may also be involved in the immune tumor defense.[54] CD44 antigen is, besides its role in cell adhesion, also involved in tumor growth and progression. Further chondroitin sulfate proteoglycan 4 is to mention. Chondroitin sulfate proteoglycan 4 is involved in cell migration and proliferation. It has recently shown to be involved in melanoma invasion into type 1 collagen[55]. Therefore, its up-regulation might be a sign for tumor metastasis. Also the up-regulated aminopeptidase N, a metabolic peptidase, has shown to be involved in tumor invasion.[56]

For the metastasis of the tumor also the over expression of insulin-like growth factor-binding protein 2 is an indicator, since this protein is involved in insulin-like growth factors mediated cell proliferation. Also follistatin-related protein 1 can be mentioned here, since it may be involved in cell proliferation. A strong indicator for the tumor invasion is also heparanase, though its overexpression can also be caused by other inflammatory processes (e.g. bacterial invasion). Two exceptional findings of the low specific marker panel are gamma-glutamyl hydrolase and aspartate aminotransferase, cytoplasmic. Both are involved in the glutamate metabolism and increase the bioavailability of free glutamate. These findings could be an indicator for the muscle degradation observed in cachectic patients. However these proteins were also found slightly up-regulated in non-cachectic patients. In the purely cancer induced alterations, mainly inflammatory proteins are present. Here to mention are serum amyloid A-1 and A-2 protein (SAA1&2), as well as alpha-1-acid glycoprotein 1.

4.4.2 Protein Expression over Time

Next, the marker protein levels over time during tumor and cachexia progression were investigated. This was performed in order to check for intra-patients variations and also to identify possible progression markers. Diagnostic markers are required to be low in expression variations within a certain condition. Otherwise it cannot be used for the safe diagnosis of a certain physical conditions, because changes in expression could also be caused by other factors. On the other hand, if a protein shows a steady expression within a certain physical condition, but changes its expression as the condition progresses, it can be used as prognostic marker. To investigate variations in protein expression over time, five time points per cancer patient were measured. Therefor, three cachectic and three non-cachectic patients were selected, which show a long time of progression before death. From the weekly gathered serum samples, five time points per patient were chosen, in monthly intervals (4 to 5 weeks depending on the availability of clinical samples). For cachectic patients, the last time point was always the closest to death, since here the cachectic outcome should be the strongest. For non-cachectic patients the last time point was selected to be at least 2 month before death, to exclude possible late-stage cachectic developments.

All samples preparation steps (depletion, digestion, etc) were performed in parallel and all samples were analysed in one measurement sequence in order to minimize variations. Data was processed and normalized as described in the bioinformatics section (see 3.3.3) and compared on protein level. Proteins, which show unusual high CVs after data processing (>25%) or signals below LOD were removed from the panel. From the 58 proteins in the dynamic MRM assay, 45 were successfully quantified for all five time point and across all six donors. Next, only protein expression levels of proteins, which showed significant regulation between the biological groups, were assessed. This was performed since stability of these protein levels, especially of the high specific cachexia markers, is very critical for the biological significance of the findings. Further, chances that one of these proteins may change its expression during cancer progression are more likely than in a protein which shows no regulation at all when compared to the healthy group.

From the 45 proteins, which could be measured in all samples, 35 also showed significant regulations between the different biological groups (see 4.4.1 Characteristic Serum Proteome Alterations in Cachectic Patients). For these proteins, changes in protein expression between

the single time points were calculated. This was done by calculating the ratios of the protein intensities for each time point compared to the time point before in chronological order. For simplification, the protein signal for each protein of the second time point was divided by the signal gathered on the first time point and so on. Thereby, based on the five time points per patient, 4 fold changes were calculated (2:1; 3:2; 4:3; and 5:4) for each patient. These fold-changes were logarithmized (to the base 2) and plotted for each protein. These fold-change plots for the three cachectic and the three non-cachectic patients are shown in Figure 21. The graphs demonstrate, that nearly all proteins follow the same trend within a patient. For example, in the patient "cachectic 1" all proteins show nearly no alterations between the first two time points (displayed in point 1), but an approximate 2-fold up-regulation was observed between the third and fourth time point (point 3). The fact that all proteins follow the same trend between the time points, independent of their biological function, speaks for a systematic and not biological cause. These systematic changes in protein content are observed in all six investigated patients, independent from development of cachexia. Two proteins were found which do not follow the trend or show a much stronger regulation. These proteins are CRP (indicated by the dashed line) and SAA1 (indicated by the dotted line). Both are major acute-phase proteins, which are known to show huge variations in serum concentration during inflammatory processes. Their regulations are so strong, that their fold-changes overcome the systematic trend observed for the other proteins. However, in patients with lower changes of CRP and SAA1 (e.g. "Non-Cachectic 2"), these proteins follow the same systematic trend as the other proteins. The fold-changes found for all other marker proteins are mostly below or within a 2-fold change rate, indicated by the blue lines. In patient "Non-Cachectic 3" two other proteins show also very strong regulations, exceeding the general trend (see point 2). These are lactate dehydrogenase A and B chain and can be accounted as marker for dying cells. They also show huge regulations in other patients and have found to be up-regulated in all tumor patients, particularly in cachectic ones. This may be a sign of the massive cell break down induced by the metastatic tumor.
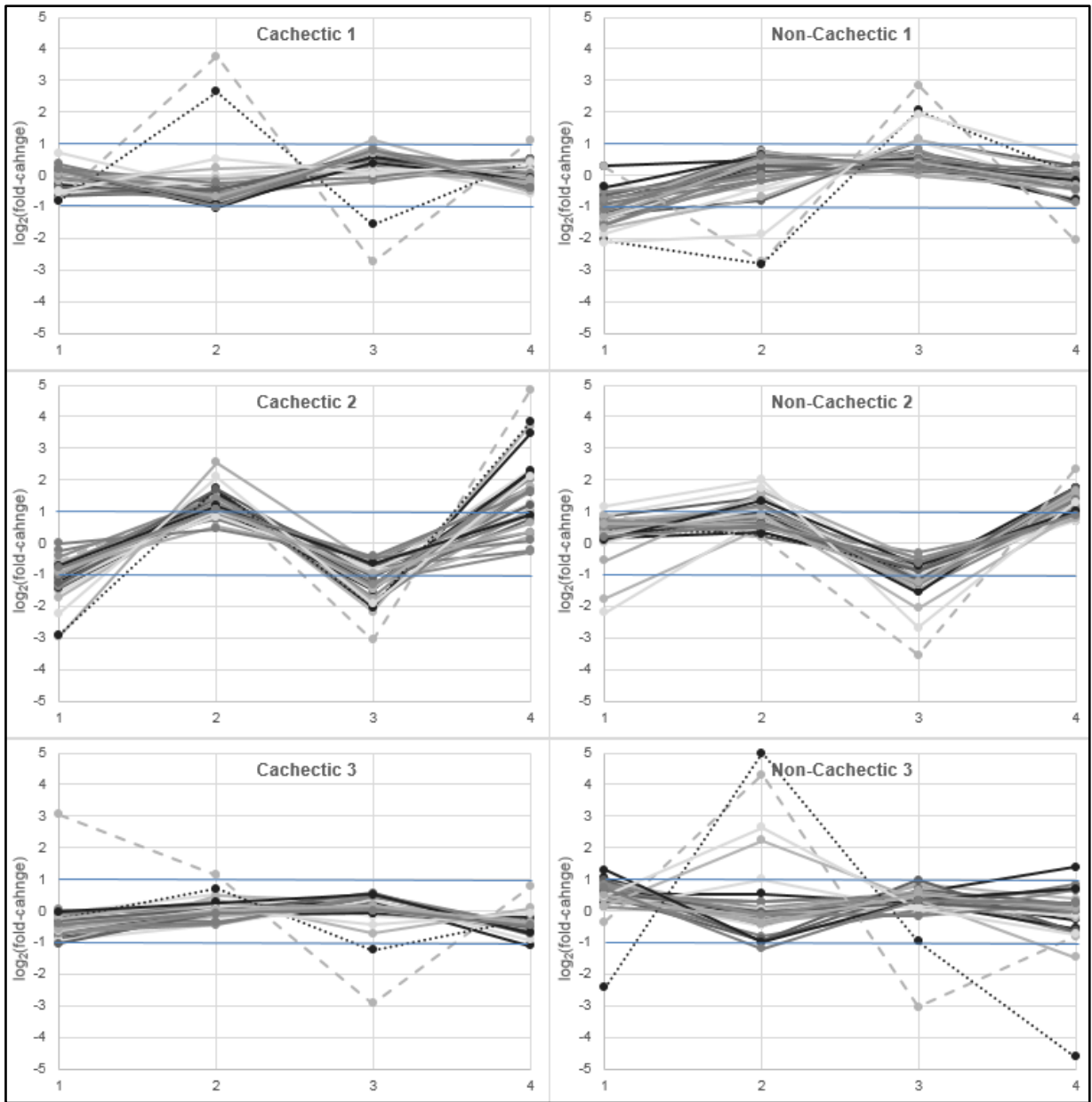
Figure 21: Protein expression over time in cachectic and non-cachectic patients

Fold-changes between each of the five time points compared to the time point before plotted in chronological order from 1 to 4. Each of the 35 regulated proteins is presented by a solid gray line. CRP (dashed line) and SAA1 (dotted line) are displayed extra, since the show exceptional fold-changes. The blue lines indicate a fold-change of 2, showing that the most protein changes were not significant in biological manner. However, it is to notice, that changes in protein expression follow the same trend within a single patient. This is a strong indicator for the presence of a systematic error.

For the observed systematic change of all proteins in the same manner within a patient, an error in the sample treatment and digestion can be excluded. As seen in the method evaluation of the digestion and depletion process (see 4.3.2 Evaluation of the Serum Depletion), only a CV below 25% is caused on average by the sample pre-treatment. Here,

the observed systematic changes or errors in protein expression are sometimes 2 to 4-fold, representing a variation greater than 100%. Since the depletion and digestion process can be excluded as cause for these systematic errors, another cause has to be found. Here, for sure the serum gathering process itself is to be mentioned. The serum gathering was performed in the clinic and not in-house. So, no details can be given here, however, there are no standardized protocols how to gather serum from whole blood. There are many commercial serum tubes on the market and many different protocols utilizing different coagulation times and centrifugation speeds, this put a lot of uncertainty in the use of serum. Further, as described in the theoretical part (see 2.2), the coagulation process itself is high in variability. During the formation of the blood clots, proteins are enclosed into the plugs in an unpredictable and inconstant fashion. This could also explain the general changes in protein content between the single time points. A last source of error could be serum storage and protein stability. Though, errors here are very unlikely, since all samples were at least in-house stored under same conditions, at -80°C and only thawed once (slow on ice). Further no trend of decreasing protein levels over time could be observed in the chronological measurement of patient samples. This excludes low stability as possible source for errors and leaving just a little uncertainty to the sample storage in the clinic and during transport. For further projects a closer collaboration with the clinic is desired, to ensure more standardized serum gathering or the providing of plasma samples.

These findings put a degree of uncertainty into the results of the serum proteome alterations in cachectic patients. Here, it is then not clear if a protein alteration is due to cachexia or just because of the error caused by the serum sampling. Nevertheless, it is to say, that most of these systematic changes are below the 2-fold change border set as significance level in the statistical testing. So, most findings are still caused by the biological condition itself and not by the sampling. However, these systematic variations in the protein content increase inter-patient variations within one biological group. This increased variation leads to increased p-values will decrease and thereby leads to less significant results. This means that some interesting protein changes might be excluded even if they were biological altered. Further it puts also an uncertainty in all findings which were just slightly above the 2-fold change cut-off. Here the sampling error cannot be completely excluded as source for the significant change in protein level. Since most marker proteins (with exception of CRP and SAA1) followed the same trend and showed nearly no changes above 2-fold, the results from the investigation for characteristic serum proteome alterations in cachectic patients are still mostly reliable.

However, as mentioned before, for future steps a more standardized and stable sampling or plasma should be chosen. Alternatively, standard proteins can be selected for normalization. These proteins need to show no alterations in expression during tumor and cachexia and their signals could then be used for normalization. Though this would work in theory, it will be hard to find proteins which are surely not altered during cancer or cachexia. Both conditions are multifactorial and affect the whole body, so nearly no biological process will be untouched. A more reliable sample or normalization would lead to more accurate results, with higher significance and possibly more significant findings.

The search for protein changes during cancer and cachexia progression cannot be performed with this data set. Reason therefore is the systematic trend in protein variations overlaying the physiological induced changes in protein expression. To access this, again more standardized samples or additional normalization would be needed.

4.4.3 Summary of the Investigation of Serum Protein Alterations

To conclude, serum alterations specific for cachexia were found by comparing serum protein levels in healthy, non-cachectic and cachectic patients. Thereby 13 high specific marker proteins were found to be significantly ($p<0.05$, 2-fold change) regulated only in cachectic patients. Here, mainly proteins involved in the cell adhesion and proliferation process were found to be altered. Another 23 proteins were found to be up-regulated in all melanoma patients, but with a stronger outcome in cachectic ones. In this low specific panel many pro-inflammatory and acute-phase proteins were found; further, indicators for bacterial invasion and immune response as well as indicators for tumor metastasis. At last, a panel of proteins which show up-regulation in all melanoma patients with no differences between cachectic and non-cachectic ones was assigned. In this panel mainly acute-phase and pro-inflammatory proteins were seen.

As second step, the protein levels of all regulated proteins within a patient over a time-span of roughly 23 weeks were monitored on a monthly basis. In doing so, it was discovered that all proteins follow certain a trend within a patient, independent from their biological function. This indicates a systematic error in the single time point samples and not a biological reason. As the in-house sample treatment and the MRM method were thoroughly evaluated, these systematic errors are most probably caused by the serum gathering process itself. Only two proteins were so strongly regulated, that their levels overcome the systematic trend, namely

CRP and SAA1, two major acute-phase proteins. The gathered results from the investigation for characteristic serum alterations in cachectic patients can still be accounted reliable, since these systematic variations are mostly under a 2-fold change.

4.5 Summary and Perspective

Aim of this thesis was the investigation of serum proteome alterations characteristic for cancer cachexia. This was achieved by combining high resolution shotgun MS conducted on a Q Exective orbitrap together with a targeted MRM method conducted on Agilent`s 6490 QqQ system. Two different serum pre-treatment techniques, namely SDS-PAGE fractionation and top12 depletion were tested and evaluated using the shotgun MS. Thereby, it was found that depletion lead to a significant higher number of identified proteins, which were also of higher biological relevance. Thorough evaluation further showed that the depletion process is highly reproducible, independent of the sample type (melanoma or healthy serum). Therefore, depletion was utilized for sample preparation for the untargeted screening as well as for the targeted MRM measurements. For the untargeted screening, three patients of the healthy, non-cachectic and cachectic group were selected and screened via high resolution shotgun MS. Based on the shotgun results, quality of the measured peptide signals, and biological relevance, a scheduled MRM assay was developed for 58 target proteins. Method evaluation showed that the MRM assay is very precise (average 10.2% CV). Good reproducibility was also proven for the whole workflow (including sample preparation and MS analysis) with an average CV below 25%.

The evaluated method was used to investigate alterations of the 58 target proteins characteristic for cancer cachexia. This was performed by measuring samples of three patients per group and comparing the acquired signal intensities on protein level between the groups. Thereby 13 high specific proteins only regulated in cancer cachexia could be found, which were mainly involved in cell adhesion and proliferation. Further, 23 low specific protein alterations, as well as very unspecific tumor-induced regulations were discovered. Here mainly acute-phase and pro-inflammatory proteins were found. 13 proteins, which were found regulated by the untargeted approach, show no or at least no meaningful regulations by the target MRM measurement. The measurement of the protein expression over time revealed an equal trend of the 58 investigated proteins within a patient, indicating a systematic error. As the method and sample treatment were thoroughly evaluated, this systematic error is most

probably a result of variations during the serum sampling process. For most proteins, this systematic trend was below the 2-fold change, which was set as significant level in the protein level investigation. Nevertheless, these systematic error increases variability and uncertainty and thereby may lowering the outcomes of the statistical evaluation.

For future projects, at first the problem of the protein variations resulting from the sampling, needs to be addressed. Here, normalization on known proteins with stable expression would be one possibility. Preferable, however, would be a closer collaboration with the clinic to ensure highly standardized sampling. A change from serum to plasma samples would most probably be the easiest solution to minimize biases during the sample treatment. After this issue was addressed, more samples should be analyzed with the developed MRM assay to increase the statistical significance of the findings. Further, additional proteins could be added to an MRM assay to test their suitability as marker proteins. The selection of these proteins could also be based on the shotgun data or on suggested markers from literature. Also cachectic samples from different tumor types should be measured, to evaluate, if the respective protein regulations are characteristic for cancer cachexia in general or just in melanoma patients. Furthermore, cachectic samples from patient with other chronic diseases, like AIDS, could be measured; here again to check if the findings are present in all cachectic outcomes or just in cancer-induced ones. As a final step, all significant findings of protein regulations with high specificity and high biological relevance should be combined in one MRM assay. For this comprehensive assay, SIS peptides should be incorporated for all target proteins, to perform absolute quantification and to pave the way towards clinical applications.

5 Abstract

Cancer cachexia is a serious wasting disorder, observed in 50-80% of all cancer patients. It is developed by final-stage cancer patients, leads to the massive loss of body fat and muscle and accounts for up to 20% of all cancer related deaths. The ability to induce cachexia is not only depended on the tumor type, but also on the host-factors. However, the driving mechanisms behind that are not fully understood and sufficient treatment methods are not yet available. Aim of this work was to investigate serum proteome alterations characteristic for cancer cachexia. Therefore, an untargeted high resolution MS-based screening (Q Exactive) was combined with a targeted MRM/MS analysis strategy (QqQ). Serum samples from non-cachectic and cachectic final-stage melanoma patients were kindly provided by Dr. Reichle (Universitätsklinikum Regensburg), whereas healthy serum was gathered in-house. Two different serum pre-treatment techniques, namely SDS-PAGE fractionation and depletion, were compared based on their shotgun MS results. Thereby, it was found that depletion leads to a higher number of quantifiable proteins with better quality. The serum depletion in combination with in-solution digestion was then used for all further experiments. For the target protein panel development, three samples per biological group (non-cachectic, cachectic, healthy) were analyzed via shotgun MS measurements. Protein identification was performed by searching the shotgun data against a human proteome database using MaxQuant. The implemented Label-free quantification algorithm further was enabled to perform relative protein quantification across the biological groups. Statistical evaluation was thereafter performed using Perseus. An MRM target panel was developed for all significantly regulated proteins (fold change ≥ 2; $p \leq 0.05$) with biological significance. Based on the high resolution shotgun data, interference-free peptides and precursor ions were selected for unscheduled MRM measurements. 93 proteins (188 peptides) were send for unscheduled MRM measurements and only interference-free transitions with sufficient signal intensities were processed further. Based on the unscheduled MRM data, a scheduled MRM assay was developed for 58 highly significant proteins (92 peptides). The MRM assay as well as the serum pre-treatment method did undergo thoroughly method evaluation, reviling a highly reproducible method. The evaluated MRM assay was used for rapid (20 min run time) and precise (<25% CV) measurements of patients samples. Three patient samples per biological group were analyzed and the determined protein levels were statistical evaluated for significant regulations (fold change ≥ 2, $p \leq 0.05$). Thereby, 13 serum proteins were identified

to be specifically regulated in cachectic samples compared to non-cachetic and healthy samples. Those included mainly cell adhesion-associated as well as pro-inflammatory proteins. In addition 23 regulated proteins with low specificity to cachexia were discovered. These proteins are mostly involved in the immune response and tumor metastasis. At least a panel of 9 very unspecific tumor-induced regulations was found, containing typical acute-phase proteins. As a last step, protein expression of all regulated proteins over time was assessed by the MRM measurements. This was performed for three cachectic and three non-cachectic patients, to ensure stability of the possible marker proteins as well as to screen for possible prognostic markers. Five serum samples (donated in monthly intervals) were analyzed per patient and variations in protein expression were calculated over time. The results showed systematic trends that are most likely caused by the serum gathering process and no evidence for biological reasons. As the outcomes of this systematic trend are not very strong, most findings of serum alterations can still be accounted as significant. However for further projects and for the search of progression markers this finding needs to be taken in account.

This work presents a robust workflow for fast and sensitive quantification of 58 proteins in human serum. The demonstrated strategy of combining untargeted screening with a precise target analysis can easily be implemented and thereafter used for the rapid and accurate measurements of a multitude of patient's samples. This would pave the way towards a better understanding of cancer cachexia and thereby point out possible clinical applications.

6 Zusammenfassung

Tumorkachexie ist eine schwere Stoffwechselstörung, welche in 50-80% aller Krebserkrankungen auftritt. Tumorkachexie wird nur im Endstadion von Tumorerkrankungen beobachtet und führt zu einem massiven Verlust an Körperfett und Muskelmasse. Kachexie ist verantwortlich für den Tod von rund 20% aller Krebspatienten und ihre biologischen Ursachen sind derzeit nur wenig verstanden. Neuste Studien zeigten jedoch, dass die Fähigkeit Kachexie zu induzieren nicht nur vom Tumortyp allein, sondern auch von seiner Umgebung abhängt. Ziel dieser Arbeit war die Untersuchung bzw. das Finden von Veränderungen im Blutserumproteom, welche charakteristisch für Tumorkachexie sind. Hierfür wurde ein ungezieltes hochauflösendes MS basierendes Screening (Q Exactive) mit einer gezielten MRM/MS Analysestrategie (QqQ) kombiniert. Blutserumproben von kachektischen und nicht kachektischen austherapierten Melanom-Patienten wurden dankenswerter Weise von Prof. Reichle (Universitätsklinikum Regensburg) bereitgestellt, während gesunde Referenzproben im Haus genommen wurden. Zwei unterschiedliche Probenaufbereitungstechniken, nämlich Fraktionierung mittels SDS-PAGE und Depletion, wurden basierend auf ihren Ergebnissen in der Shotgun-Analyse evaluiert. Dabei zeigte sich, dass Depletion zu einer höheren Anzahl an detektierbaren Proteinen mit größerer biologischer Signifikanz führt. Depletion der Serumproben in Kombination mit einem Proteinverdau in Lösung wurde daher für alle weiteren Experimente herangezogen. Zur Bestimmung potentiell regulierter Proteine wurden drei Proben aus jeder Kohorte (kachektisch, nicht kachektisch, gesund) mittels Shotgun-MS untersucht. Die Proteinidentifikation wurde mittels MaxQuant-Suche gegen die Humane Proteome Datenbank durchgeführt. Der in MaxQuant implementierte „Label-free quantification" Algorithmus wurde zur relativen Quantifizierung aller identifizierten Proteine in den biologischen Gruppen herangezogen. Eine statistische Auswertung der Ergebnisse erfolgte mittels Perseus. Zielproteine für die MRM/MS Analyse waren jene mit signifikanter Konzentrationsänderung (Unterschied ≥ 2-fach; $p ≤ 0.05$) und hoher biologischer Signifikanz. Basierend auf den hochauflösenden MS Daten, wurden Peptide und die korrespondierenden interferenzfreien Vorläufer Ionen für jedes der Zielproteine ausgewählt. 93 Proteine (188 Peptide) wurden für die statischen MRM-Messungen herangezogen und nur interferenzfreie Übergänge, die ein rauscharmes Signal zeigten, wurden weiter prozessiert. Aus den aufgenommenen Daten wurde ein dynamisches MRM Assay für 58 Proteine (92 Peptide) entwickelt. Die validierte

Methode wurde dann zur schnellen (20 min) und präzisen (<25% CV) Analyse von Patientenproben herangezogen. Drei Patienten pro Kohorte wurden dabei untersucht, und die gemessenen Veränderungen im Proteingehalt wurden statistisch auf Signifikanz (Unterschied ≥ 2-fach; $p \leq 0.05$) geprüft. Hierbei wurden 13 Kachexie-spezifisch regulierte Proteine gefunden. Diese 13 Proteine sind hauptsächlich in Prozessen der Zelladhäsion und der proentzündlichen Stimulierung beteiligt. Zusätzlich wurden weitere 23 Proteinänderungen mit einer geringen Spezifität für Kachexie entdeckt. Diese Proteine sind zumeist in der Immunantwort und der Tumor Metastasierung involviert. Auch 9 Tumor-spezifisch regulierte Proteine konnten gefunden werden. Diese sind typische Akutphase-Proteine. Um die gefundenen Unterscheide im Proteingehalt auf Stabilität zu prüfen bzw. mögliche Trends während des Krankheitsverlaufs zu erkennen, wurde die Expression aller regulierten Proteine über die Zeit bestimmt. Dies wurde für drei kachektische und drei nicht kachektische Patienten durchgeführt. Fünf Serumproben pro Patient (in monatlichen Intervallen) wurden mittels des entwickelten MRM Assays vermessen und Veränderungen im Proteingehalt über die Zeit errechnet. Die Ergebnisse zeigten einen systematischen Trend, der höchstwahrscheinlich durch die Probennahme verursacht wurde, ohne Evidenz für biologische Ursachen. Da die Effekte dieses systematischen Fehlers nicht sehr stark waren, können die meisten gefunden regulatorischen Unterschiede noch immer als signifikant angesehen werden. Trotzdem sollte diese Tatsache für weitere Schritte und zukünftige Projekte unbedingt berücksichtigt werden.

Diese Arbeit präsentiert eine robuste Methode zur schnellen und sensitiven Quantifizierung von 58 Proteinen im humanen Serum. Die eingesetzte kombinative Strategie aus ungezieltem Screening und gezielter Analyse lässt sich einfach implementieren und kann zur Messung einer Vielzahl von klinischen Proben mit hohem Durchsatz verwendet werden. Dies würde letztlich zu einem besseren Verständnis von Tumorkachexie führen und den Weg zu neuen klinischen Anwendungen öffnen.

References

1.   Argiles, J. M.; Busquets, S.; Stemmler, B.; Lopez-Soriano, F. J., Cancer cachexia: understanding the molecular basis. *Nat Rev Cancer* **2014,** *14* (11), 754-762.

2.   Chiba, F.; Soda, K.; Yamada, S.; Tokutake, Y.; Chohnan, S.; Konishi, F.; Rikiyama, T., The importance of tissue environment surrounding the tumor on the development of cancer cachexia. *International Journal of Oncology* **2014,** *44* (1), 177-186.

3.   Delano, M. J.; Moldawer, L. L., The Origins of Cachexia in Acute and Chronic Inflammatory Diseases*. *Nutrition in clinical practice* **2006,** *21* (1), 68-81.

4.   Kir, S; White, J. P.; Kleiner, S.; Kazak, L.; Cohen, P.; Baracos, V. E.; Spiegelman, B. M., Tumour-derived PTH-related protein triggers adipose tissue browning and cancer cachexia. *Nature* **2014,** *513* (7516), 100-+.

5.   Mondello, P.; Mian, M.; Aloisi, C.; Fama, F.; Mondello, S.; Pitini, V., Cancer Cachexia Syndrome: Pathogenesis, Diagnosis, and New Therapeutic Options. *Nutr. Cancer* **2015,** *67* (1), 12-26.

6.   Lisanti, M. P.; Martinez-Outschoorn, U. E.; Chiavarina, B.; Pavlides, S.; Whitaker-Menezes, D.; Tsirigos, A.; Witkiewicz, A.; Lin, Z.; Balliet, R.; Howell, A.; Sotgia, F., Understanding the "lethal" drivers of tumor-stroma co-evolution Emerging role(s) for hypoxia, oxidative stress and autophagy/mitophagy in the tumor micro-environment. *Cancer Biology & Therapy* **2010,** *10* (6), 537-542.

7.   Seok, J.; Warren, H. S.; Cuenca, A. G.; Mindrinos, M. N.; Baker, H. V.; Xu, W.; Richards, D. R.; McDonald-Smith, G. P.; Gao, H.; Hennessy, L.; Finnerty, C. C.; López, C. M.; Honari, S.; Moore, E. E.; Minei, J. P.; Cuschieri, J.; Bankey, P. E.; Johnson, J. L.; Sperry, J.; Nathens, A. B.; Billiar, T. R.; West, M. A.; Jeschke, M. G.; Klein, M. B.; Gamelli, R. L.; Gibran, N. S.; Brownstein, B. H.; Miller-Graziano, C.; Calvano, S. E.; Mason, P. H.; Cobb, J. P.; Rahme, L. G.; Lowry, S. F.; Maier, R. V.; Moldawer, L. L.; Herndon, D. N.; Davis, R. W.; Xiao, W.; Tompkins, R. G.; Inflammation, t.; Host Response to Injury, L. S. C. R. P., Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences* **2013,** *110* (9), 3507-3512.

8.   (a) Zhang, H.; Liu, A. Y.; Loriaux, P.; Wollscheid, B.; Zhou, Y.; Watts, J. D.; Aebersold, R., Mass spectrometric detection of tissue proteins in plasma. *Molecular & cellular proteomics : MCP* **2007,** *6* (1), 64-71; (b) Rifai, N.; Gillette, M. A.; Carr, S. A., Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotech* **2006,** *24* (8), 971-983.

9.   Liu, Y.; Huttenhain, R.; Collins, B.; Aebersold, R., Mass spectrometric protein maps for biomarker discovery and clinical research. *Expert Rev. Mol. Diagn.* **2013,** *13* (8), 811-825.

10.   (a) Mirza, S. P.; Olivier, M., Methods and approaches for the comprehensive characterization and quantification of cellular proteomes using mass spectrometry. *Physiol. Genomics* **2008,** *33* (1), 3-11; (b) Tang, J.; Gao, M. X.; Deng, C. H.; Zhang, X. M., Recent development of multi-dimensional chromatography strategies in proteome research. *J. Chromatogr. B* **2008,** *866* (1-2), 123-132.

11.   Michalski, A.; Cox, J.; Mann, M., More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority is Inaccessible to Data-Dependent LC-MS/MS. *J. Proteome Res.* **2011,** *10* (4), 1785-1793.

12.   Parker, C. E.; Borchers, C. H., Mass spectrometry based biomarker discovery, verification, and validation – Quality assurance and control of protein biomarker assays. *Molecular Oncology* **2014,** *8* (4), 840-858.

13. Percy, A. J.; Simon, R.; Chambers, A. G.; Borchers, C. H., Enhanced sensitivity and multiplexing with 2D LC/MRM-MS and labeled standards for deeper and more comprehensive protein quantitation. *Journal of Proteomics* **2014,** *106*, 113-124.

14. Fearon, K. C.; Carter, D. C., Cancer cachexia. *Annals of Surgery* **1988,** *208* (1), 1-5.

15. Fujiwara, Y.; Kobayashi, T.; Chayahara, N.; Imamura, Y.; Toyoda, M.; Kiyota, N.; Mukohara, T.; Nishiumi, S.; Azuma, T.; Yoshida, M.; Minami, H., Metabolomics Evaluation of Serum Markers for Cachexia and Their Intra- Day Variation in Patients with Advanced Pancreatic Cancer. *PLoS One* **2014,** *9* (11), 9.

16. Vasseur, S.; Tomasini, R.; Tournaire, R.; Iovanna, J. L., Hypoxia Induced Tumor Metabolic Switch Contributes to Pancreatic Cancer Aggressiveness. *Cancers* **2010,** *2* (4), 2138-2152.

17. Warburg, O., On the origin of cancer cells. *Science (New York, N.Y.)* **1956,** *123* (3191), 309-14.

18. Martinez-Outschoorn, U. E.; Balliet, R. M.; Rivadeneira, D. B.; Chiavarina, B.; Pavlides, S.; Wang, C.; Whitaker-Menezes, D.; Daumer, K. M.; Lin, Z.; Witkiewicz, A. K.; Flomenberg, N.; Howell, A.; Pestell, R. G.; Knudsen, E. S.; Sotgia, F.; Lisanti, M. P., Oxidative stress in cancer associated fibroblasts drives tumor-stroma co-evolution: A new paradigm for understanding tumor metabolism, the field effect and genomic instability in cancer cells. *Cell cycle (Georgetown, Tex.)* **2010,** *9* (16), 3256-76.

19. Constantinou, C.; Fontes de Oliveira, C. C.; Mintzopoulos, D.; Busquets, S.; He, J.; Kesarwani, M.; Mindrinos, M.; Rahme, L. G.; Argiles, J. M.; Tzika, A. A., Nuclear magnetic resonance in conjunction with functional genomics suggests mitochondrial dysfunction in a murine model of cancer cachexia. *International journal of molecular medicine* **2011,** *27* (1), 15-24.

20. Argiles, J. M.; Lopez-Soriano, J.; Almendro, V.; Busquets, S.; Lopez-Soriano, F. J., Cross-talk between skeletal muscle and adipose tissue: a link with obesity? *Medicinal research reviews* **2005,** *25* (1), 49-65.

21. Petruzzelli, M.; Schweiger, M.; Schreiber, R.; Campos-Olivas, R.; Tsoli, M.; Allen, J.; Swarbrick, M.; Rose-John, S.; Rincon, M.; Robertson, G.; Zechner, R.; Wagner, Erwin F., A Switch from White to Brown Fat Increases Energy Expenditure in Cancer-Associated Cachexia. *Cell Metabolism 20* (3), 433-447.

22. Virtanen, K. A.; Lidell, M. E.; Orava, J.; Heglind, M.; Westergren, R.; Niemi, T.; Taittonen, M.; Laine, J.; Savisto, N. J.; Enerback, S.; Nuutila, P., Functional brown adipose tissue in healthy adults. *The New England journal of medicine* **2009,** *360* (15), 1518-25.

23. Molfino, A.; Laviano, A.; Rossi Fanelli, F., Contribution of anorexia to tissue wasting in cachexia. *Current opinion in supportive and palliative care* **2010,** *4* (4), 249-53.

24. Eschenhagen, T.; Force, T.; Ewer, M. S.; de Keulenaer, G. W.; Suter, T. M.; Anker, S. D.; Avkiran, M.; de Azambuja, E.; Balligand, J. L.; Brutsaert, D. L.; Condorelli, G.; Hansen, A.; Heymans, S.; Hill, J. A.; Hirsch, E.; Hilfiker-Kleiner, D.; Janssens, S.; de Jong, S.; Neubauer, G.; Pieske, B.; Ponikowski, P.; Pirmohamed, M.; Rauchhaus, M.; Sawyer, D.; Sugden, P. H.; Wojta, J.; Zannad, F.; Shah, A. M., Cardiovascular side effects of cancer therapies: a position statement from the Heart Failure Association of the European Society of Cardiology. *European journal of heart failure* **2011,** *13* (1), 1-10.

25. Klein, G. L.; Petschow, B. W.; Shaw, A. L.; Weaver, E., Gut barrier dysfunction and microbial translocation in cancer cachexia: a new therapeutic target. *Current opinion in supportive and palliative care* **2013,** *7* (4), 361-7.

26. Anderson, N. L.; Anderson, N. G., The human plasma proteome: history, character, and diagnostic prospects. *Molecular & cellular proteomics : MCP* **2002,** *1* (11), 845-67.

27.   Omenn, G. S., The Human Proteome Organization Plasma Proteome Project pilot phase: Reference specimens, technology platform comparisons, and standardized data submissions and analyses. *PROTEOMICS* **2004,** *4* (5), 1235-1240.

28.   Lange, V.; Picotti, P.; Domon, B.; Aebersold, R., Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecular Systems Biology* **2008,** *4*, 222-222.

29.   Domanski, D.; Percy, A. J.; Yang, J.; Chambers, A. G.; Hill, J. S.; Freue, G. V.; Borchers, C. H., MRM-based multiplexed quantitation of 67 putative cardiovascular disease biomarkers in human plasma. *Proteomics* **2012,** *12* (8), 1222-43.

30.   Tang, J.; Gao, M.; Deng, C.; Zhang, X., Recent development of multi-dimensional chromatography strategies in proteome research. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2008,** *866* (1-2), 123-32.

31.   Huttenhain, R.; Malmstrom, J.; Picotti, P.; Aebersold, R., Perspectives of targeted mass spectrometry for protein biomarker verification. *Current opinion in chemical biology* **2009,** *13* (5-6), 518-25.

32.   Brunelle, J. L.; Green, R., One-dimensional SDS-Polyacrylamide Gel Electrophoresis (1D SDS-PAGE). In *Laboratory Methods in Enzymology: Protein, Pt C*, Lorsch, J., Ed. Elsevier Academic Press Inc: San Diego, 2014; Vol. 541, pp 151-159.

33.   Scientific, T., Pierce Top 12 Abundant Protein Depletion Spin Columns Instructions. **2012**.

34.   Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R., Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews* **2013,** *113* (4), 2343-2394.

35.   Claessens, H. A.; van Straten, M. A., Review on the chemical and thermal stability of stationary phases for reversed-phase liquid chromatography. *Journal of chromatography. A* **2004,** *1060* (1-2), 23-41.

36.   Hernandez-Borges, J.; Aturki, Z.; Rocco, A.; Fanali, S., Recent applications in nanoliquid chromatography. *Journal of separation science* **2007,** *30* (11), 1589-610.

37.   Banerjee, S.; Mazumdar, S., Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry* **2012,** *2012*, 40.

38.   Wilm, M.; Mann, M., Analytical Properties of the Nanoelectrospray Ion Source. *Analytical Chemistry* **1996,** *68* (1), 1-8.

39.   Michalski, A.; Damoc, E.; Hauschild, J.-P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S., Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. *Molecular & cellular proteomics : MCP* **2011,** *10* (9), M111.011015.

40.   Domon, B.; Aebersold, R., Options and considerations when selecting a quantitative proteomics strategy. *Nat Biotech* **2010,** *28* (7), 710-721.

41.   Cox, J.; Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotech* **2008,** *26* (12), 1367-1372.

42.   Bileck, A.; Kreutz, D.; Muqaku, B.; Slany, A.; Gerner, C., Comprehensive assessment of proteins regulated by dexamethasone reveals novel effects in primary human peripheral blood mononuclear cells. *J Proteome Res* **2014,** *13* (12), 5989-6000.

43.   Sherwood, C. A.; Eastham, A.; Lee, L. W.; Risler, J.; Vitek, O.; Martin, D. B., Correlation between y-type ions observed in ion trap and triple quadrupole mass spectrometers. *J Proteome Res* **2009,** *8* (9), 4243-51.

44.   Roemmelt, A. T.; Steuer, A. E.; Poetzsch, M.; Kraemer, T., Liquid Chromatography, in Combination with a Quadrupole Time-of-Flight Instrument (LC QTOF), with Sequential Window Acquisition of All Theoretical Fragment-Ion Spectra (SWATH) Acquisition: Systematic Studies on Its Use for Screenings in Clinical and Forensic Toxicology and Comparison with Information-Dependent Acquisition (IDA). *Analytical Chemistry* **2014,** *86* (23), 11742-11749.

45.   Uniprot Database., http://www.uniprot.org/. **03.05.2015**.

46.   Dave, R. K.; Dinger, M. E.; Andrew, M.; Askarian-Amiri, M.; Hume, D. A.; Kellie, S., Regulated Expression of PTPRJ/CD148 and an Antisense Long Noncoding RNA in Macrophages by Proinflammatory Stimuli. *PLoS One* **2013,** *8* (6), 13.

47.   Barrow, A. D.; Astoul, E.; Floto, A.; Brooke, G.; Relou, I. A.; Jennings, N. S.; Smith, K. G.; Ouwehand, W.; Farndale, R. W.; Alexander, D. R.; Trowsdale, J., Cutting edge: TREM-like transcript-1, a platelet immunoreceptor tyrosine-based inhibition motif encoding costimulatory immunoreceptor that enhances, rather than inhibits, calcium signaling via SHP-2. *Journal of immunology (Baltimore, Md. : 1950)* **2004,** *172* (10), 5838-42.

48.   Albers, J. J.; Wolfbauer, G.; Cheung, M. C.; Day, J. R.; Ching, A. F.; Lok, S.; Tu, A. Y., Functional expression of human and mouse plasma phospholipid transfer protein: effect of recombinant and plasma PLTP on HDL subspecies. *Biochimica et biophysica acta* **1995,** *1258* (1), 27-34.

49.   Herbert, J.; Wilcox, J. N.; Pham, K. T.; Fremeau, R. T., Jr.; Zeviani, M.; Dwork, A.; Soprano, D. R.; Makover, A.; Goodman, D. S.; Zimmerman, E. A.; et al., Transthyretin: a choroid plexus-specific transport protein in human brain. The 1986 S. Weir Mitchell award. *Neurology* **1986,** *36* (7), 900-11.

50.   Black, S.; Kushner, I.; Samols, D., C-reactive Protein. *The Journal of biological chemistry* **2004,** *279* (47), 48487-90.

51.   Roxburgh, C. S.; McMillan, D. C., Role of systemic inflammatory response in predicting survival in patients with primary operable cancer. *Future oncology (London, England)* **2010,** *6* (1), 149-63.

52.   Chitu, V.; Stanley, E. R., Colony-stimulating factor-1 in immunity and inflammation. *Current opinion in immunology* **2006,** *18* (1), 39-48.

53.   Kelley, S. L.; Lukk, T.; Nair, S. K.; Tapping, R. I., The crystal structure of human soluble CD14 reveals a bent solenoid with a hydrophobic amino-terminal pocket. *Journal of immunology (Baltimore, Md. : 1950)* **2013,** *190* (3), 1304-11.

54.   Ullrich, A.; Sures, I.; D'Egidio, M.; Jallal, B.; Powell, T. J.; Herbst, R.; Dreps, A.; Azam, M.; Rubinstein, M.; Natoli, C.; et al., The secreted tumor-associated antigen 90K is a potent immune stimulator. *The Journal of biological chemistry* **1994,** *269* (28), 18401-7.

55.   Iida, J.; Pei, D.; Kang, T.; Simpson, M. A.; Herlyn, M.; Furcht, L. T.; McCarthy, J. B., Melanoma chondroitin sulfate proteoglycan regulates matrix metalloproteinase-dependent human melanoma invasion into type I collagen. *The Journal of biological chemistry* **2001,** *276* (22), 18786-94.

56.   Pasqualini, R.; Koivunen, E.; Kain, R.; Lahdenranta, J.; Sakamoto, M.; Stryhn, A.; Ashmun, R. A.; Shapiro, L. H.; Arap, W.; Ruoslahti, E., Aminopeptidase N is a receptor for tumor-homing peptides and a target for inhibiting angiogenesis. *Cancer research* **2000,** *60* (3), 722-7.

Curriculum Vitae
Martin Eisinger

Professional Experience

- Oct 2014 – Apr 2015   **Master Thesis**
"Investigation of serum proteome alterations characteristic for cancer induced cachexia. Combining a HRMS-based screening (orbitrap) with a targeted LC-MRM/MS strategy."
Supervision by Prof. Christopher Gerner and Besnik Muqaku.
University of Vienna, Department for Bioanalysis and Separation Techniques - Vienna, Austria

- Oct 2012 – Feb 2013   **External Bachelor Thesis**
"Development of a 2D LC-MRM/MS method for the Multiplexed Quantitation of NCD-linked Biomarkers in Undepleted Human Plasma."
Supervision by Dr. Christoph Borchers and Dr. Andrew J. Percy.
UVic Genome BC Proteomics Centre – Victoria, B.C., Canada

- Jan 2012 – Jul 2012   **Part-time Position as Technical Assistant**
Laboratory assistant in the preparation of river water samples for pesticide analysis via LC-MS/MS.
Institute for Analytical Research (IFAR) - Idstein, Germany

- Mar 2011 – Jul 2011   **Practical Semester**
"Establishment of analysis parameters for the rapid characterization of crude oils and heavy oil residues."
Supervision by Ing. Martina Jüttner and Prof. Leo Gros
OMV R&M AG - Schwechat, Austria

- Sept 2010 – Feb 2011   **National Service**
Completion of basic military service in the military command Vienna as a pioneer and a truck driver.

Education

- Sept 2013 – ongoing   **Master of Chemistry**
Modular masters course "Chemistry" with personal focus on analytical chemistry.
University of Vienna - Vienna, Austria

- Mar 2011 - Feb.2013 **Bachelor of Applied Chemistry**
  Lateral entry in the 5th of 8th bachelor semesters, after successful graduation from the Technical High School.
  Final mark: 1.2
  University of Applied Science Fresenius - Idstein, Germany

- Sept 2005 – June 2010 **Technical High School for Chemistry**
  Technical High School (post-secondary, non-tertiary education) for chemistry.
  Major subjects: environmental analysis and process engineering
  Final mark: 1.5
  HBLVA Rosensteingasse - Vienna, Austria

- Sept 2001 – June 2005 **Gymnasium**
  Participation in the school pilot "Bionik" (Interdisciplinary learning in all STEM-subjects).
  GRG16 Maroltingergasse - Vienna, Austria

Poster Publications

- M. Eisinger, B. Muqaku, A. Bileck, A. Reichle, C. Gerner**: Investigation of proteome alterations characteristic for tumor associated cachexia: Combining high resolution MS-based screening (Orbitrap) with a targeted analysis strategy (MRM).**, ANAKON 2015, Mar 23 – 26, 2015, Graz, Austria

Additional Qualification and Competences

Language Skills: German (mother tongue)
English (business fluent, TESPiS English Certificate Level B2)

Software Skills: MS Office
Common software for HPLC, GC and MS data acquisition and evaluation (e.g. Mass Hunter, Skyline, MaxQuant and the like)

Certificates: Basic Certificate Project Management (GPM)
Certified Quality-Coach (Systemcert)
ECDL Core

Driving License Classes: A/B/C1/C2/F

Referees

- Prof. Dr Christopher Gerner
  Email: christopher.gerner@univie.ac.at   Phone: +43 1 4277 52302

- Dr Christoph Borchers
  Email: christoph@proteincentre.com   Phone: +1 250 483 3221

- Prof. Dr Thomas Knepper
  Email: knepper@hs-fresenius.de   Phone: +49 6126 35264

X

Appendix