universität wien

MASTERARBEIT

Titel der Masterarbeit

„Development of an online digestion method for the assessment the human embryo secretome“

verfasst von

Philipp Obermayr BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2014

Studienkennzahl lt. Studienblatt:	A 066 834
Studienrichtung lt. Studienblatt:	Molekulare Biologie UG2002
Betreut von:	ao. Univ.-Prof. tit. Univ.-Prof. Mag. Dr. Andreas Rizzi

"Phantasie ist wichtiger als Wissen, denn Wissen ist begrenzt"

Albert Einstein

Für Herta, Alice und Olivia

Danksagung

An dieser Stelle möchte ich mich bei allen Personen bedanken, die diese Arbeit erst möglich gemacht haben. Allen voran möchte ich meinem Betreuer, Goran Mitulovic, danken, der es mir ermöglicht hat, auf dem spannenden Gebiet der Proteomics zu arbeiten, und der mir immer mit Rat und Tat zur Seite stand. Herzlichen Dank an Profesor Andreas Rizzi, der mir im Laufe der Arbeit immer wieder mit Ratschlägen geholfen hat und der viel Zeit in die Verbesserung dieser Arbeit investiert hat. Vielen Dank an Chen Weiqiang, der als mein Mentor fungiert hat und von dem ich viel gelernt habe. Weiters möchte ich mich bei meinen Kollegen, Andreas Fichtenbaum, Tanja Panic und Axel Führer herzlich bedanken, die mir bei der Probenvorbereitung sowie der Datenaufarbeitung geholfen haben, und mit denen ich eine gute Zeit verbracht habe. Vielen Dank!

Zusammenfassung

Das Forschungsfeld „Proteomics“ befasst sich mit der qualitativen und (relativen) quantitativen Analyse der Gesamtheit der Proteine in Zellen, subzellulären Organellen, Geweben, in Körperflüssigkeiten und Sekreten oder sogar in gesamten Organismen. Derzeit werden dafür im Wesentlichen zwei unterschiedliche Strategien eingesetzt, einerseits die Analyse nach Auftrennung der intakten Proteine, der sogenannte „top-down“-Ansatz, und andererseits die Analyse der Proteine nach deren Spaltung in Peptide, der sogenannte "bottom-up" Ansatz. Letzterer umfasst zumeist eine enzymatische Spaltung der Proteine in Peptide an definierten Spaltstellen ("Proteinverdau"). Dieser Proteinverdau ist dabei in zeitlicher Hinsicht der Flaschenhals jedes "bottom-up" Experiments. Die vorliegende Arbeit befasst sich mit der Entwicklung einer Analysenmethode, die einen online Verdau beinhaltet, bei der Enzymreaktoren mit immobilisiertem Trypsin zum Einsatz kommen, um schnelle Proteom-Analysen zu ermöglichen. Zukünftig soll diese Methode unter anderem für die Analyse des Sekretoms von humanen Embryonen, die für In-Vitro Fertilisation gezüchtet wurden, eingesetzt werden. In diesem Falle ist das rasche Vorliegen von Daten zur Beurteilung der Embryonen besonders wichtig.

Der erste Teil dieser Arbeit beschäftigt sich mit Probenvorbehandlung und der Optimierung eines mobile Phase-Gradienten für die Auftrennung tryptischer Peptide mittels High Performace Liquid Chromatography (HPLC). Die Proben wurden unter verschiedenen denaturierenden Bedingungen behandelt, wobei sich Harnstoff für die Denaturierung der Proteine als sehr gut geeignet erwies, während sich das Surfactant Rapigest störend auf das Gesamtsystem auswirkte. Als mobile Phase-Gradient wurde ein ternärer Gradient mit Acetonitril, Methanol und Trifluorethanol als organische Lösungsmittelkomponenten entwickelt.

Der größte Teil dieser Arbeit befasst sich mit der Wahl und "Optimierung" der Verdau-Bedingungen im Enzymreaktor. Das pH Optimum lag dabei, wie erwartet, zwischen 7,5 und 8,5, und von den getesteten Puffersubstanzen erwies sich Ammoniumacetat als das am besten geeignete. Bezüglich des Einflusses von organischen Lösungsmittelkomponenten erwies sich, dass Acetonitril in einer Konzentration von 10% die Verdaueffizienz steigert, ebenso wie die Zugabe von 2M Harnstoff in die Pufferlösung. Als weitere Parameter wurde der Einfluss von Verdautemperatur und

Verdauzeit, bestimmt durch die Flussrate, untersucht. Dabei erwies sich eine Temperatur von 40°C als geeignet. Der Einfluss der Flussrate auf die Verdaueffizienz war geringer als erwartet, es stellte sich ein Volumenfluss von 7 µl/min, entsprechend einer Verweildauer im Trypsin-Reaktor von 14 min, als guter Kompromiss dar.

Als ein besonderes Problem stellte sich die Verschleppung der Probe durch Adsorption im Trypsin-Reaktor heraus. Weder das Waschen mit organischen Lösungsmitteln, noch das mit hoch konzentrierten Salzlösungen konnte diesen Effekt wesentlich reduzieren. Es wurde hingegen deutlich, dass mit der Steigerung der Verdau-Effizienz auch die Verschleppungseffekte reduziert werden konnten.

Ein wichtiger Punkt dieser Studie war die Konfigurierung eines Gesamt-Chromatographie-Systems, das hohen Probendurchsatz und online Verdau mit kurzen Aufenthaltszeiten im Reaktor erlaubt. Dafür wurde ein System entworfen, das aus jeweils zwei Trypsin-Reaktoren, zwei Vorsäulen und zwei Trennsäulen besteht. Mit diesem System wurde es möglich, die Wasch-und Equilibrierungszeiten zu erhöhen, die Analysendauer aber dennoch zu senken.

Schlussendlich wurde die neu entwickelte Methode anhand von IVF-Proben mit einer Standardmethode verglichen. Als vorläufiges Ergebnis konnte festgestellt werden, dass die vorgestellte Methode mit online Verdau für die Analytik von komplexen Proben geeignet ist.

Abstract

The research area of “proteomics” deals with the qualitative and (relative) quantitative analysis of the entire complement of proteins in cells, sub-cellular organelles, tissues, in body fluids or even in whole organisms. Currently, basically two different strategies are applied. On the one hand, in the so called “top down” approach, intact proteins are separated and then analyzed. On the other hand, in the so called "bottom-up" approach, proteins are enzymatically cleaved into peptides with defined cleavage sites ("protein digestion") before analysis. Protein digestion is the bottleneck of every bottom up proteomic experiment in terms of time consumption. In order to enable fast proteomic analyses, a method for online protein digestion is developed in this work, using immobilized trypsin reactors. In future, this method should be used for the analysis of the secretome of human embryos grown for *in vitro* fertilization (IVF). In this field of application, short analysis times are mandatory.

The first part of this work comprises the development of suitable sample preparation conditions as well as an eluent gradient for the separation of tryptic peptides with high performance liquid chromatography (HPLC). Samples were treated with different denaturing agents, of which urea was found to be well suited for the denaturing of proteins, whereas the Rapigest surfactant interfered with our system. As mobile phase gradient, a ternary gradient with acetonitrile, methanol and triflouroethanol as organic modifiers was developed.

The major part of this work deals with finding and “optimizing” appropriate digestion conditions for an immobilized enzyme reactor (IMER). As expected, the pH optimum for the IMER was between pH 7.5 and 8.5. Several different buffer substances were tested, of which ammonium acetate turned out being the most appropriate buffer salt. Regarding the addition of organic solvents to the buffer solution, it turned out that acetonitrile at a concentration of 10% could increase the IMER efficiency. A similar effect could be attained by addition of 2M urea to the buffer solution. As further parameters, the influence of digestion temperature and digestion time, controlled by the flow rate, have been investigated. A temperature of 40°C was shown to be appropriate. The flow rate did not have as great an impact as we had anticipated, being optimal at 7 µl/min, corresponding to a dwell time in the IMER of 14 min.

In the course of the study, memory effects caused by the IMER turned out being a major issue. Neither washing steps with organic solvents, nor those with high salt solutions could significantly reduce these effects. However, it was shown that the enhancement of digestion efficiency of the IMER was effective in reducing the carryover effect.

A further important point of the work was setting up a complete HPLC system suitable for high throughput online digestion analyses. Eventually, a system consisting of two IMER's, two trap columns and two separation columns has been developed, which allows parallel runs. In this way, the run time could be reduced while washing and equilibration time was enhanced.

Finally, the newly developed method was compared with a standard bottom up proteomic approach using IVF-samples. As preliminary result, it could be shown that the online digestion method is suited for the analysis of complex proteomic samples.

Table of Contents

Abbreviations

ACN acetonitrile
AFC affinity chromatography
BSA bovine serum albumin
CID collision induced dissociation
DTT dithiothreitol
ECD electron transfer dissociation
EmPAI exponentially modified protein abundance index
ESI electrospray ionization
ETD electron transfer dissociation
FA formic acid
FWHM full width at half maximum
HAc acetic acid
HCl hydrochloric acid
HILIC hydrophilic interaction chromatorgaphy
HPLC high performance liquid chromatography
IAA iodoacetamide
ICAT isotope coded affinity tag
ID inner diameter
IEX ion exchange chromatography
IMER immobilized enzyme reactor
IR infra red
iTRAQ isobaric tag for relative and absolute quantitation
IVF in vitro fertilization
LC liquid chromatography
MALDI matrix assisted laser desorption-ionization
MeOH methanol
Mr molecular mass
MS mass spectrometer
m/z mass to charge ratio
PMF peptide mass fingerprinting
PTM post translational modification
RPLC reversed phase liquid chromatography

PS-DVB polysytrene divenylbenzene

SEC size exclusion chromatography

SILAC stable isotope labeling in amino acid culture

TEAB triethylammonium bicarbonate

TFA trifluoroacetic acid

TFE trifluoroethanol

TOF time of flight

UV ultra violet

V valve

1 Introduction

The term ‘proteome’ was introduced in the mid-1990s in analogy to the term ‘genome’ [1], describing the entire complement of proteins of an organism, a tissue, a cell or a body fluid at a defined state. While the human genome consists of approximately 25.000 genes [2], the number of different proteins in the human body is estimated to be at around 1 million [3]. This is due to the fact that a single gene might code for several proteins. Genomic recombination, alternative splicing, differential transcription termination and initiation at different promoters can lead to the generation of several different mRNAs of a single gene [4]. Posttranslational modifications (PTMs) of the initial proteins further increase the complexity of the proteome. In contrast to the genome, the proteome is highly dynamic and changes during ontogenesis, in response to environmental stimuli and various pathological situations, making proteomics a highly challenging research field.

1.1 Strategies for Proteome Analysis

Two different approaches for proteomic experiments are currently existing, referring to the entities which are introduced into the mass spectrometer (MS) [5]. The so called ‘bottom up’ approach is currently the most widely used proteomic strategy. It implies the digestion of the proteins to peptides by biological or chemical agents, followed by liquid chromatographic (LC) separation and their mass spectrometric detection and identification. Protein digestion is most frequently done by the use of the protease trypsin. The generated peptides are usually separated using reversed phase liquid chromatography (RPLC), where volatile mobile phases are used, so that the high performance liquid chromatography (HPLC) apparatus can be coupled online to the MS via electrospray ionization (ESI) [6]. Although being the most mature proteomic method, the bottom up approach has some inherent limitations. Not all peptides of a protein are detected by the mass spectrometer and some with very low abundances only, making it challenging to distinguish between different isoforms of a certain protein. Furthermore, information on posttranslational modifications can easily be lost due to the limited sequence coverage and also by fragmentation during MS.

In the ‘top down’ approach, intact proteins are separated and introduced into the MS [7]. In contrast to the bottom up strategy, it is more likely to obtain complete se-

quence information, thus making this procedure more suitable for the analysis of different splice variants of proteins and the accessory of PTMs. Here, the time consuming step of protein digestion is avoided. However, there are major drawbacks to this approach. The main problem is the low solubility of hydrophobic proteins. In MS analysis, the charge state of the protein after ionization is not uniform, in fact there exists a charge state distribution. The dissociation behavior of intact proteins is not well characterized [5]. The spectral information generated by multiply charged proteins is very extensive, making top down strategies unsuitable for complex samples like cell lysates. Furthermore, bioinformatic tools provided for top down approach at the moment are quite limited compared to the ones available for bottom up strategies [7].

The 'middle down' approach can be seen a compromise between the two strategies mentioned above. Here, the proteins are digested into large peptides, for example using the protease ompT [8]. This approach is, for instance, well suited for the analysis of large proteins like histones [9].

Proteomics comprises both, qualitative and quantitative analysis of the proteome. For protein identification, a widely used method is peptide mass fingerprinting (PMF) [10, 11], where in a bottom up-manner peptides of digested proteins are separated and their accurate mass is determined by MS. These peptide masses are then compared to the theoretical masses of *in silico* digests, calculated by protein sequence databases by certain algorithms, of which MASCOT is the most popular [12]. This method can only identify proteins with known sequence, which are listed in the database of choice Protein identification can also be done using MS^2.. Therefore, the peptides are fragmented after MS, and the fragments undergo another round of MS. The fragment spectra are then compared to a database containing a spectrum library.

The field of quantitative proteomics comprises numerous methods, which can be divided into absolute and relative quantification methods, as well as in label-free and labeling based methods, where the labeling can be carried out either *in vivo* or *in vitro*. Stable isotope labeling by amino acids in cell culture (SILAC) [13] provides cells with a labeled essential amino acid, e.g. deuterated leucine, which will be included in their proteome. Proteins of cell cultures containing deuterated variants of amino acids can be compared with proteins from cell cultures fed by non-deuterated amino acids. Isotope coded affinity tags (ICAT) [14] are molecules reacting with the thiol group of

cysteines on one end and a biotin tag for affinity enrichment on the other end. These two groups are separated by a linker, which can be labeled by deuterium, thus enabling distinguishing of differently coded proteins. In contrast, the isobaric tags for relative and absolute quantification (iTRAQ) [15] method is based on the labeling of peptides with tags of identical mass. They will release reporter groups of different masses during fragmentation in the MS which can be detected and discriminated during MS^2. A simple variant of label free absolute quantitation is based on the exponentially modified protein abundance index (emPAI) [16]. It is calculated automatically during a database enquiry in MASCOT and delivers an approximation to the quantity for proteins identified in a PMF experiment.

1.2 Protein Digestion

Due to certain shortcomings of top-down proteomics, mostly originating from the large size of intact proteins and the limited solubility of hydrophobic membrane proteins, the bottom-up approach is at the moment the most widely used proteomic strategy. Advantages of bottom-up proteomics lie in the lower molecular mass and the narrower charge state distribution of peptides compared to proteins. Furthermore, peptides are more easily separated by LC than proteins, thus enhancing sensitivity of the analysis [17]. In bottom-up proteomics, protein digestion is an essential step, strongly influencing the outcome of the experiments [18].

1.2.1. Enzymatic Protein Digestion Using Trypsin

Despite the fact that many different proteolytic enzymes are currently available for controlled protein digestion, in the majority of experiments, trypsin is used as the enzyme of choice. This is because of the high cleavage efficiency, the cleavage site specificity and the nature of tryptic peptides, which are well suited for MS based analyses [19]. Trypsin specifically cleaves the protein backbone C-terminal of Lys and Arg, which are usually well distributed throughout a protein [18]. However, there are some exceptions from this rule, as trypsin does not cleave the peptide bond, when Arg or Lys are followed by Pro [20] or when the access to the cleavage site is sterically hindered, e.g. by PTMs. When two or more positive charged residues are next to each other, trypsin will cleave after the most C-terminal of this amino acid.. Tryptic

cleavage leads to peptides with an average length of about 14 amino acids, which have at least two positive charges [18] at physiological pH-conditions.

Protocols used for tryptic digestion usually include denaturation of the protein with chaotropic agents, such as urea, followed by the reduction of disulfide bonds with dithiothreithol (DTT) and subsequent alkylation of the cysteins by iodoacetamide (IAA). The actual process of digestion is carried out at neutral to slightly basic pH, since trypsin has an activity optimum at a pH range from seven to eight [21]. Standard protocols include incubation over about 16 hours at 37°C. The digestion is then stopped by the addition of formic acid (FA) or triflouroacetic acid (TFA). It has been shown that digestion efficiency of a given protocol varies for different proteins, and that optimal digestion conditions are protein dependent [22].

1.2.2. Alternative Enzymatic and Chemical Digestion

Under certain circumstances, like overabundance of Lys and Arg or due to pH incompatibilities, it can be favorable to use a different enzyme. For example, pepsin can be used as an alternative to trypsin when working at acidic pH [23]**.** As pepsin cleaves less specifically, a larger number of peptides is obtained and thus MS data become very complex. Endoproteinases like Arg-C, Asp-N, Glu-C and Lys-C show high specificity and cleavage efficiency and can be utilized instead of trypsin [24].

In order to access complex proteomes, combined digestion using two different enzymes can be a promising approach. This way, the number of identified peptides and proteins can be increased significantly [25, 26]. Common strategies are the use of typsin in combination with, Lys-N [27] or Glu-C [28].

Non enzymatic protein digestion can be carried out using acids like FA [29], acetic acid (HAc) [30], or hydrochloric acid (HCl) [31] or other chemicals, for example cyanogen bromide [32] and hydroxylamine [33]. Furthermore, electrochemical oxidation is another way of protein digestion, resulting in specific cleavage at Tyr and Trp [34, 35].

1.2.3. Enhancement of Enzyme Based Protein Digestion

Various ways to improve protein digestion using trypsin have been reported, implying the introduction of energy into the system. An increase of temperature generally leads to faster reaction kinetics. Besides, higher temperatures cause protein unfolding and thus increased cleavage site accessibility [36], which improves digestion. The practical application is limited though, because of the thermolability of most enzymes. This can be overcome by using for example reductive methylation of trypsin, which shifts the temperature of the activity maximum of the enzyme to 50-60°C [37], as well as the use of thermostable enzymes like thermolysin [38].

Application of pressure can have positive effects on protein digestion [39]. It has been shown that digestion under a pressure of six bar for 30 minutes yielded similar results compared to a classical over night approach [40]. Increasing the pressure up to 2400 bar enabled the digestion of bovine serum albumin (BSA) in as little as 60 seconds [39]. Unfolding of proteins may not only be achieved by heat, but also using digestion buffers containing organic solvents. A certain amount of organic solvents may therefore have a positive effect on the digestion process [41, 42]. Treatment of the sample with microwave irradiation during digestion can reduce the reaction time significantly [43]. The same has been sown for ultrasound [44] and infrared (IR) irradiation [45, 46].

1.2.4. Protein Digestion Using Immobilized Trypsin

The immobilization of enzymes in bioreactors is nowadays an established strategy utilized for enzymatic reactions. Using such an IMER for digestion of proteins implies some distinctive advantages over the classical over night digestion approach. It is time saving and reduces the risk of nonspecific cleavage [47]. The number of sample preparation steps is reduced, especially when coupled online to a LC system. Therefore the risk of sample contamination by external factors is minimized and the robustness of the analysis is thus increased. The digestion efficiency is enhanced due to the increased enzyme to substrate ratio without promoting autolysis of the enzyme [48]. It has been shown that immobilized trypsin is active at a higher pH range than free trypsin [49].

Various different materials have been shown to be appropriate for trypsin immobilization. Inorganic particles such as silica and glass beads [50], as well as inorganic monoliths based on silica [51] were used. Trypsin immobilized on silica based monoliths has been shown to have an activity increased 2000 fold over trypsin in solution [52]. Alternatively, trypsin can be immobilized directly onto the inner surface of silica-based capillaries [53], as well as on nanomaterials, such as nanoparticles [54] and nanofibers. Organic monoliths, based on polyacrylamide [55] or polymethacrylate [56] are also used as trypsin carrier [57, 58]. Besides, organic membranes [59] and organic particles such as cross linked polystrene divenylbenzene (PS-DVB, Poroszyme®) [60] should be mentioned. In practice, flow through IMERs based on monolithic carrier materials show the best performance [61]. The advantages of monolithic materials, referred to in the next chapter, apply for both separation and IMER performance.

Various different types of immobilization reactions exist up to date [62]. Basically, trypsin can be immobilized via covalent bondage or physical adsorption or can be incapsulated in sol-gel [63]. Covalent immobilization can be done e.g. by binding via epoxy groups, aldehyde groups, carbonyldiimidazole, activated ester groups, amongst others [64].

1.3 Basics of HPLC

Proteomic samples are usually very complex. They can consist of whole cell or tissue lysates, as well as body fluids, containing a multitude of different proteins. Furthermore, the dynamic range of protein concentration in these samples is usually very high; one cell can contain between 1 and 1 million copies of two different proteins [65], and the concentration range of human serum proteins has been estimated to exceed ten orders of magnitude [66]. Thus, the reduction of the complexity of biological samples is one of the most important and most challenging steps in proteomics. Before the development of capable HPLC systems, two dimensional (2D) electrophoresis was the method of choice for separating complex protein samples [67]. Nowadays, this technique is not as widely used anymore, because of some major drawbacks. It is very labor intensive and results are hard to replicate. Low abundant proteins and very hydrophobic proteins are underrepresented, as well as basic, very small and large proteins. [68-71].

1.3.1. HPLC Techniques for Separating Proteins and Peptides

For the separation of proteins and peptides, RPLC, ion exchange chromatography (IEX), affinity chromatography (AFC), hydrophobic interaction chromatography (HIL-IC) and size exclusion chromatography (SEC) are the techniques generally used [72]. Among them, RPLC is the most prominent separation method in proteomics [73], because of its superior efficiency compared to IEX and SEC [74], its robustness and the possibility of online coupling to MS.

1.3.1. Reversed Phase HPLC

While the RPLC of proteins and peptides is based on hydrophobic interactions between amino acid side chains and the immobilized n-alkyl ligands of the stationary phase, the exact molecular mechanisms are not yet fully understood [75].

The sample is applied to the column using an aqueous mobile phase, usually containing five to ten percent organic modifier. Elution is facilitated by increasing the amount of organic solvent components in the mobile phase. Therefore, the analytes are eluted in order of increasing hydrophobicity [76]. The most common organic modifier in RPLC is acetonitrile (ACN), while alcohols like methanol or isopropanol are less favorable because of their high viscosity when mixed with water, generating a high system backpressure. A small change in organic modifier content can alter the adsorption properties of a peptide drastically, and a typical proteomic sample contains a multitude of different peptides with a diverse range of retention factors. Therefore, gradient conditions are favored for elution [77].

In order to facilitate the retention of peptides on the stationary phase, it is crucial to add ion pairing agents to the mobile phase. By forming ion pairs, these agents mask the charged sites of the peptides. Therefore, the hydrophobicity of the peptide is increased, thus enabling binding to the hydrophobic stationary phase [78]. For peptide separation, TFA is the most widely used ion pairing agent because of its high ionic strength. Nevertheless, for online coupled LC-MS systems, TFA can be inappropriate. It forms very strong ion pairs with the analyte, preventing the ionization in the electrospray [79]. Alternatively, different fluorinated carbonic acids, like heptafluorobutyric acid, or weaker acids like FA can be used [80]. The loss of chromatographic

resolution using a weaker ion pairing agents is balanced with receiving a higher MS signal.

Stationary phases in RPLC are usually porous silica micro particles, sized around 2-5 µm. Pores for peptide analysis are usually sized between 100 and 200 Å [77]. The silica material is chemically modified with hydrophobic n-alkyl ligands, of which n-octadecyl (C18) n-butyl (C4) and n-octyl (C8) are the most common [81-83]. Alternatively, phenyl- [84] and cyano-groups [85] can be used as well. During the process of alkyl-immobilization, about half of the silica groups are modified, whereas the rest is further silanized. This procedure is referred to as end capping. The type of n-alkyl ligand has a significant impact on the retention of proteins [86]. The longer the ligand, the more hydrophobic the stationary phase, leading to a stronger retention of hydrophobic analytes. Silica packed columns suffer from certain shortcomings, which basically are the high pressure needed for separation, sensitivity of the silica against pH values above 8 [87] and nonspecific adsorption of analytes on silanol groups[88].

1.3.1.1. Alternative Stationary phase materials for PR-HPLC

To overcome the drawbacks of particle packed columns, monolithic columns provide an alternative to conventional packed columns, Monoliths are continuous bed structures, consisting of a single, mostly rod shaped particle, including highly interconnected pores. They can be made of silica and organic polymers, of which the latter are especially convenient for the separation of high molecular weight compounds like proteins and peptides [89, 90]. PS-DVB based monoliths were the first ones to be used in RP-HPLC. Because of their high chemical stability at pH range from 1 to 14 [88] and good analytical properties, they are widely used [91-93]. Moreover, methacrylate based monoliths are popular because of their short polymerization time and simple surface modification [88].

Due to the absence of mesopores [94], mass transfer in organic monoliths is driven mainly by convection rather than diffusion [88], yielding lower peak dispersion and back pressure at high flow rates. They show improved mass transfer rates and peak capacity [95]. Beyond that, monolithic columns are easy to fabricate and do not require retaining frits [88].

A further alternative to classical totally porous particles are core-shell particles. They are made of a solid core which is surrounded by a shell made of porous material [96]. Core-shell particles were introduced in the late 60's by Horvath and coworkers [97]. Columns packed with core-shell particles provide improved mass transfer kinetics and thus allow faster separation especially for macromolecules, compared to classical fully porous beads of the same size [96], [98].

Proteomic samples pose a special challenge to HPLC systems, because usually they are present in minute concentrations only. Therefore the system should be miniaturized in order to achieve low sample dilution and low flow rates compared to conventional HPLC [99]. For increased sensitivity, the inner diameter (ID) of the separation columns are reduced, allowing detection of proteins in the attomole and sometimes zeptomole range [100]. When working with miniaturized systems, it is crucial to avoid void volume and any dispersion effects from capillaries, detectors and switching valves as well as memory effects [101, 102].

1.4 Mass Spectrometry

Following the separation of peptides or proteins, the detection via MS is the technique of choice in virtually every proteomic experiment. A mass spectrometer basically consists of an ion source, a mass analyzer, which separates the ions according to their mass to charge ratio (m/z) in the gas phase and a detector, which counts the ions hitting the detector plate at a given time

1.4.1. Ionization

Ionization is the first step in a mass spectrometric analysis, bringing the analyte molecules in the gas phase and adding one or more positive or negative charges. Classical ionization methods like electron impact ionization often lead to breakage of covalent bonds. Therefore, MS analysis of proteins is only possible since the 1980’s, when soft ionization techniques like fast atom bombardment (FAB) and matrix assisted laser desorption-ionisation (MALDI) were introduced, enabling the ionization of non-volatile, heat-unstable proteins and peptides.

1.4.1.1. Matrix Assisted Laser Desorption-Ionisation

In MALDI, the sample is mixed with a matrix of organic molecules, which is usually dissolved in a mixture of water and organic solvent, and afterwards spotted onto a MALDI plate. Certain features are required for a substance to be used as a matrix for MALDI. It has to be water soluble, absorb ultra violet (UV) or IR irradiation at a given wavelength of the laser [103], and it has to be volatile so to evaporate under laser irradiation. Common matrices used for protein analyses are dihydroxybenzoic acid [104], 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) [105, 106], 4-hydroxy-3-methoxycinnamic acid (ferulic acid) [105, 106], and α-Cyano-4-hydroxycinnamic acid [107]. Alternatively, liquid ionic matrices (LIM), like nitrobenzylalcohol and nitrophenyloctlyether are in use [108], but have not established yet. After letting the solution dry, the crystallized matrix containing the sample molecules is targeted by pulsed laser irradiation. Usually, nitrogen lasers with a wavelength of 337 nm and neodymium-doped yttrium aluminum garnet (YAG) lasers with a wavelength of 266 or 355 nm respectively are used [109]. Less commonly used are IR lasers. The energy originating from the laser irradiation is absorbed by the matrix molecules, leading to fast evaporation and desorption of matrix and analyte molecules and the formation of a plume in which the sample molecules become ionized [110]. The exact mechanism of sample ionization in MALDI, however, is not yet fully understood. One model tries to explain the ionization process within a photo excitation model, where neutral analyte molecules are incorporated into the matrix. The laser irradiation leads to a photoionization of the matrix molecules, which in turn will ionize the analyte molecules [111, 112]. For the so called "lucky survivor" model, it is assumed that the analytes in the matrix exist as ions. Following a laser pulse, clusters consisting of matrix and analyte molecules are formed, and the analytes get neutralized. The analyte ions are the "survivors" of this process [113]. Generally, MALDI yields mostly single positive charged ions.

Advantages of MALDI are the high sensitivity, which goes into the low femtomole range and the high mass range up to 300.000 m/z. The fact that mainly single charged cations are produced facilitates the analysis of complex mixtures. The limitations are the low resolution of MALDI MS in the high m/z ranges over 15.000 as well as high background ions in the range below 700 Da.

1.4.1.2. Electro Spray Ionization

ESI is an alternative way of ionization, frequently used for peptides. In ESI, the analyte solution flows through a capillary, on top of which an electric potential of about 2-3 kV is applied. The resulting electric field between the spray capillary and the entrance capillary to the MS leads to a polarization of the solvent, and to an accumulation of positive ions near the surface of the meniscus. This causes the distortion of the meniscus into the so called Taylor cone [114, 115]. The tip of the Taylor cone emits a small jet of droplets of analyte solution [116], often assisted by a coaxial gas flow [117]. The droplets carry an excess of positive charged ions, mainly protons. The solvent of the droplets readily evaporates, generating even smaller droplets with a positive surface charge. At a certain droplet radius, the repulsion between the positive charges on the surface will overcome the surface tension. This condition is called Raleigh limit [119]. Droplets at the Raleigh limit disrupt into nanometer-sized droplets which are, through further evaporation, the source of completely desolvated gaseous analyte molecules [120-122].

Three different models exist for the transition of analytes into the gas phase, depending on size and shape of the molecule. Large, globular molecules like unfolded proteins undergo the transition via evaporation of the surrounding solvent molecules of the droplet at the Raleigh limit. The positive charge contained in the solvent is transferred to the analyte molecule. This reaction is referred to as the charged residue model [123, 124]. When the protein exists in an unfolded state, e.g. because it is eluted in an acidic or hydrophobic mobile phase, it is ionized via the chain ejection model [125, 126]. The hydrophobic character of unfolded proteins makes it unfavorable for them to reside within the droplet interior. Therefore it migrates to the droplet surface, where it is getting ejected stepwise. Analytes with low molecular mass (Mr), preexisting as an ion in the droplet, undertake the transition into the gas phase via the ion evaporation model, where the analyte gets ejected due to the electric field strength in the Raleigh droplet [121, 127].

For proteome analytics, ESI flow rates in the range of 100 to 300 nl/min are used [128]. This so called 'nano-ESI' has certain advantages like higher sensitivity and higher tolerance against buffer salts [129] as well as the compatibility to nano-LC.

A decisive benefit of ESI is the possibility of an on line coupling of the LC system to the MS. Furthermore, the presence of multiple charge states allows one to attain high mass accuracies. On the other hand, ESI shows a low tolerance against salts and detergents and requires high purity of the sample.

1.4.2. Mass Analyzers

After Ionization, the ions are separated in the mass analyzer according to their m/z ratio. One of the most widely spread mass analyzers used for MS is the quadrupole [130]. It consists of four parallel rod-shapes electrodes. The two opposing electrodes have the same potential. Between the neighboring electrodes a radio frequency (RF) voltage is applied, superimposed by a direct current (DC) voltage. In the resulting electrical field ions will move in certain trajectories trough the quadrupole. Only ions with certain m/z ratios have stable trajectories at a given electrical field while the others collide with the electrodes and get discharged. Quadrupoles have a limited mass range and rather low resolving power, however, triple quadrupole instruments are used for MS/MS [131], where the central quadrupole serves as collision cell for fragmentation of selected precursor ions. The first and the third quadrupoles are operated either in a mass filtering mode, allowing only analytes with a certain m/z to pass though, or in a scanning mode, where the electrical field varies and analytes with various m/z ratios are allowed to pass over a certain time span.

In quadrupole Ion traps [132, 133], an oscillating RF field keeps the ions in circular trajectories inside the trap rather than a movement through. Through manipulation of the electric field, ions get ejected. The trapped ions can be fragmented, e.g. by collision with inert gas atoms, allowing MS^n operations.

In Time-of-flight (TOF) [134, 135] analyzers, ions are accelerated in an electric field. The time, which is needed for the ions to move through a field free drift tube, is proportional to the square root of the m/z ratio. The fact that ions of the same m/z ratio can enter the drift tube with different kinetic energies impairs the resolution of TOF analyzers, particularly when ionized by MALDI. This shortcoming has been solved with the introduction of a reflectron [136], which accelerates the ions backward according to their kinetic energy, thus balancing the energy difference and achieving flight-time focusing.

The latest development in MS, frequently used for proteomics, is the Orbitrap mass analyzer [137, 138]. It consists of a spindle-shaped electrode surrounded by a barrel-like outer electrode. In the electrostatic field formed by the two electrodes, ions move in both axial and radial trajectories around the inner electrode, while the force of attraction applied by the inner electrode is balanced by centrifugal forces. The oscillation frequency along the z-axis of an ion is inverse proportional to the square root of the ion's m/z ratio. With a resolution power up to 100.000 full width at half maximum (FWHM) and a mass accuracy of down to 0.2 ppm [139, 140], the Orbitrap mass analyzer is a powerful tool in proteomic MS.

1.4.3. Tandem MS

In tandem-MS, also referred to as MS^2 or MS^n, ions of a previously selected m/z ratio are fragmented, and the emerging fragments undergo another MS scan [141]. MS^2 leads to enhanced selectivity [141], and fragmentation patterns can reveal structural information of an analyte. In order to perform MS^2 experiments, mass spectrometers which combine two or more mass analyzers, can be used. Common instruments for MS^2 are the above mentioned triple quadrupole instrument, the LTQ Orbitrap, QqTOF [142, 143] as well as TOFTOF instruments [144].

Methods for fragmentation of peptides include collision induced dissociation (CID) [145], where analyte ions collide with inert gas molecules, mostly helium, nitrogen or argon. The collision energy thereby leads to bond breakage. In electron capture dissociation (ECD) [146, 147], multiply positive charged analytes are fragmented due to the reaction with a free electron in the gas phase. In electron transfer dissociation (ETD) [148, 149], electrons are transferred by collision with radical anions, leading to so called c- and /or z-ions. Like ECD, ETD requires multiple charged ions. Because labile groups like phosphates and sialic acids are not removed by ETD and ECD, these techniques are powerful tools for the study of certain PTMs, especially phosphorylation and glycosylation [150, 151].

1.5 In Vitro Fertilization

IVF is the major treatment for infertility in human reproductive medicine. Indications for IVF treatment are female infertility due to tubal disease, mainly caused by infection, inflammation and endometriosis, as well as male infertility caused by low sperm counts or low sperm quality.

The IVF procedure is usually proceeded by ovarian hyperstimulation, where follicle stimulating hormone (FSH) is administered to stimulate the development of multiple ovarian follicles [152]. Final oocyte maturation is induced by application of human chorionic gonadotropin (hCG), followed by retrieval of the oocytes from the body. An oocyte is then incubated with previously prepared sperm cells for fertilization. In case of male infertility, intracytoplasmic sperm injection (ICSI) can be performed, where a single sperm cell is directly introduced into the oocyte using a glass-micropipette [153].

The embryo can be cultured either in artificial medium or on a cell layer extracted from the mother's uterus [154]. The time point of implantation varies, from six to eight cell stadium to blastocyst stadium [155], according to local laws. The number of implanted embryos is varied, depending on age, infertility status and local laws. Implantation of two or more embryos increases the chance of pregnancy, but bears the risk of multiple pregnancy.

1.5.1. Embryo Selection

The evaluation of the quality of the embryo to be implanted is a crucial step in the IVF process. Up to date, the assessment of morphological criteria like cleavage rate and blastomere symmetry determined by microscopy [156] is still widely used. Besides, alternative techniques, like time-lapse microscopy [157] and gene expression profiling [158] are emerging. Furthermore, the search for biomarkers, which allow a qualification of an embryo using proteomic techniques are currently being explored by different groups [159-162].

1.6 Objectives

The aim of the study is to develop an automated online digestion system for the rapid analysis of the human embryo secretome. The goal is to use this method for biomarker discovery and, in future, to adapt it for routine testing of human embryo secretome. This would provide an alternative method for the assessment of the viability of human embryos grown for implantation in the course of IVF.

In the classical bottom up proteomic approach, the tryptic digestion of the sample proteins is the time-consuming bottleneck of every experiment, as it is usually done overnight, but at least needs several hours to be performed. In the field of IVF, the choice of an embryo for implantation is one of the crucial steps which has to be made shortly before implantation. Therefore, a fast analytical method is needed.

Furthermore, the introduction of online digestion into bottom up proteomics would reduce the sample handling steps of an analysis, minimizing the risk of sample contamination and increasing the robustness of the test system. Overall, the benefits of such an automated online digestion method could find many applications in the field of clinical proteomics, where fast and reliable protein identification of complex samples existing in minute amounts is necessary.

This study concretely comprises of:

- Determination of the optimal digestion conditions for the provided trypsin columns
- Adaptation of separation conditions to facilitate detection of peptides
- Development of a LC system which meets the requirements of an automated digestion and enables rapid sample analysis
- Comparison of the newly developed method with established proteomic methods

2 Material & Methods

2.1 Chemicals and Reagents

Ammonium acetate	Merck (Darmstadt, Germany)
Calcium chloride	Merck (Darmstadt, Germany)
Urea	Sigma Aldrich (Steinheim, Germany)
Formic Acid (Emsure)	Merck (Darmstadt, Germany)
Acetic Acid (Emsure)	Merck (Darmstadt, Germany)
Trifluoroacetic Acid	Sigma Aldrich (Steinheim, Germany)
Ammonia solution, 25 %	Merck (Darmstadt, Germany)
Acetonitrile	Sigma Aldrich (Steinheim, Germany)
Methanol	Merck (Darmstadt, Germany)
Isopropanol	Sigma Aldrich (Steinheim, Germany)
Trifluoroethanol	Alfa Aesar (Karlsruhe, Germany)
Dithiothreitol	Gerbu (Heidelberg, Germany)
Iodoacetamide	Sigma Aldrich (Steinheim, Germany)
Triethylammonium bicarbonate (TEAB)	Sigma-Fluka (Buchs, Switzerland)
RapiGest	Waters (Milford, MA, USA)
Trypsin (Sequencing grade, lyophilized)	Promega (Madison, WI, USA)
LTQ Velos ESI Positive Ion Calibration Solution	Thermo scientific (Marietta, OH, USA)
Complete Lysis-M protease inhibitor	Roche (Mannheim, Germany)
BSA	Sigma Aldrich (Steinheim, Germany)
Casein (from bovine milk)	Sigma Aldrich (Steinheim, Germany)
Ribonuclease A	Sigma Aldrich (Steinheim, Germany)
Ubiquitin (human, recombinant)	Sigma Aldrich (Steinheim, Germany)
Alcohol dehydrogenase (from yeast)	Sigma Aldrich (Steinheim, Germany)
Serotransferrin	Sigma Aldrich (Steinheim, Germany)
Lysozyme (from eggwhite)	Sigma Aldrich (Steinheim, Germany)
Carbonic anhydrase (from bovine erythrocytes)	Sigma Aldrich (Steinheim, Germany)

2.1 Apparatuses and Instruments

Directdetect infrared spectrometer	Merck-Millipore (Darmstadt, Germany)
Incubator	Techne (Staffordshire, UK)
LTQ Velos mass spectrometer	Thermo Scientific (Waltham, MA, USA)
Microcentrifuge	VWR (Radnor, PA, USA)
Sringe pump module	ProtHispra consortium (Vienna, Austria; Debrecen, Hungary; Moscow, Russia; Ajdovščina, Slovenia)
Thermomixer	Eppendorf (Hamburg, Germany)
Ultimate Nano HPLC	Dionex (Sunnyvale, CA, USA)
Centrifugal filter units 30 kDa	Millipore (Billerica, MA, USA)
Valve switching device	ProtHispra consortium (Vienna, Austria; Debrecen, Hungary; Moscow, Russia; Ajdovščina, Slovenia)

2.2 Sample preparation

2.2.1. Preparation of BSA samples

To prepare the bovine serum albumin (BSA) samples for online digestion, 10 µl of a stock solution containing 10 mg/ml BSA, were mixed with 30 µl 8 M urea and 96 µl of 50 mM triethylammonium bicarbonate (TEAB). 100 µl 10 mM DTT, dissolved in 50 mM TEAB, were added and the sample was heated at 60 °C for 30 minutes. After allowing the sample to cool at room temperature for 10 minutes, 6 µl 500 mM IAA, dissolved in 50 mM TEAB, were added and the sample was incubated in the dark for 30 minutes. Finally, 10 µl 10 mM DTT were added. The sample had a final concentration of 6 pmol/µl BSA and 0.95 M urea.. If not stated otherwise, for online digestion experiments, the sample was diluted with digestion buffer to a concentration of 1 pmol/µl, and 3 pmol were injected. For mixed protein samples, solutions of single proteins were made in a similar way and pooled afterwards.

For in solution digestion, the sample was prepared as described above. Additionally, 20 µl of trypsin reconstituted in the provided buffer, containing 2 µg of enzyme, was added. The sample was then incubated at 37 °C for 16 hours. To stop digestion, the sample was put on -20°C.

2.2.2. Preparation of IVF samples

From the samples received from the IVF clinic (about 70 µl), 40 µl were pipetted out instantly into 100 µl of protease inhibitor mix. The samples were pooled, so that media originated from seven embryos, all of them previously rated either viable or non-viable, represent one sample. The samples were filtrated through a 30 kDa cutoff filter, followed by centrifugation at 14.000 rpm (11.000 g) with the microcentrifuge.

The following three fractions were collected:

- The "filtrate" or flow through fraction
- The "concentrate" fraction was obtained by reversing the filter and centrifugation at 1.000 rpm for 2 minutes
- The "eluate" was obtained by pipetting 15 µl of 50 mM TEAB each side of the filter membrane, followed by centrifugation of the reversed filter

Protein concentration has been quantified in each fraction using the Direct Detect® IR-spectrometer, measuring the band at 1600-1690 cm^{-1} corresponding to the C=O stretching vibration of the peptide bond. 2 µl of sample were applied for measurements.

2.3 Reversed Phase Separation

2.3.1. Separation Conditions

For reversed phase separation of peptides, C18 Acclaim Pepmap columns with an inner diameter of 75 µm, a length of 25 cm, a particle size of 3 µm and a pore size of 100 Å were used. Samples were trapped in C18 trap columns of 300 µm inner diameter, 5 mm length, using particles with 5 µm size and 100 Å pore size.

A ternary mobile phase gradient was applied, consisting of the following eluent constituents:

- Eluent A: 5% ACN, 0.1% FA
- Eluent B: 15% ACN, 15 % MeOH, 0.1% FA
- Eluent C: 60% ACN, 30% MeOH, 10% TFA, 0.08% FA

The time course of the gradient was varied, depending on the sample. For complex biological samples, a longer gradient was applied: 3% B from 0 to 48 min, 3% B to 50% B from 48 to 160 min, 50% B to 60% B and from 0% C to 40% C from 160 to 220 min, and from 40% C to 100 % C from 220 to 280 min. 100% B were run for 10 min, and the column was equilibrated with 3% B for 15 min.

BSA samples and protein mix samples were separated with a steeper gradient: For example 0 to 34 min 5% B, 5% - 100% B from 34 to 50 min and 0% to 100% C from 50 to 60 min. Time of 100% C and afterwards 5% B was varied. The flow rate of the nanopump was set to 300 nl/min.

2.3.2. LC Setup

The initial LC system was operated by two pumps, a loading pump and a nanopump. It was set up in a way that the digested sample from the IMER gets trapped on a C-18 trap column. When switching valve (V) 2, the trap column gets reverse-flushed by the nanopump, delivering the trapped sample onto the separation column.

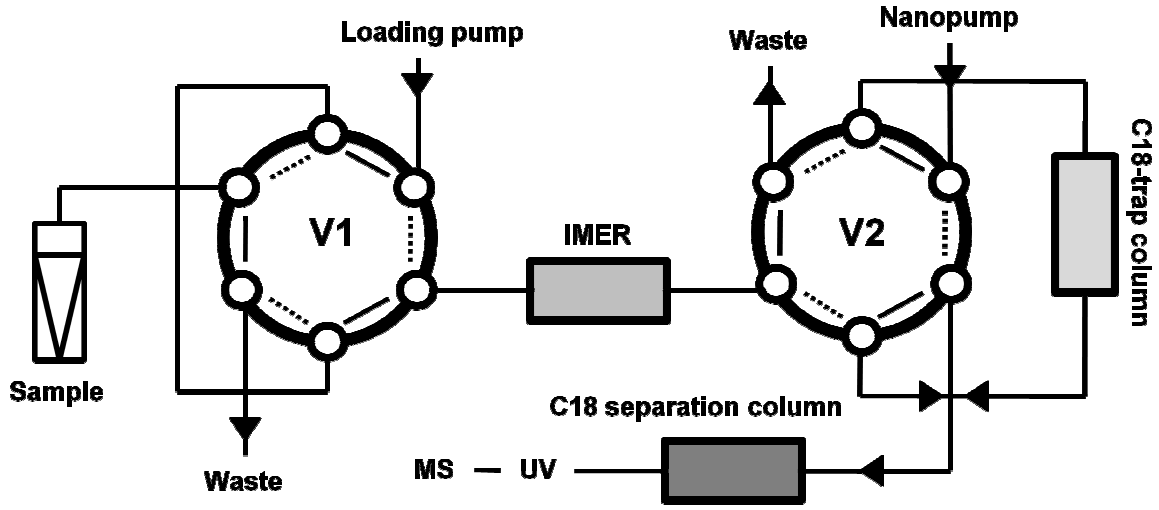

Figure 1: Scheme of the LC system for online digestion

2.4 On Line-Digestion

Protein online digestion was performed on methacrylate monoliths carrying immobilized trypsin. The exact monolith properties, as well as the immobilization chemistry cannot be stated here due to patent reasons. The IMER was included in the LC system as depicted in Figure 1:. Parameters of digestion are subject of the 'results and discussion' section of this work.

2.5 Mass spectrometry

MS was carried out using a LTQ Velos linear ion trap. The sample was ionized by nano-ESI at a flow rate of 300 nl/min. MS^2 spectra were obtained by CID for the ten most abundant peaks of the MS^1 scan. Following parameters were set for the MS:

- Single charged ions were excluded
- Active exclusion settings were: Repeat count 1, repeat duration 30 seconds, exclusion list size 300, exclusion duration 30 seconds.

- CID with fragmentation energy of 20 eV
- m/z range between 300 and 1800.

2.6 Database Search

Proteomic data processing and database search for the MS data was performed using the MASCOT software (Matrix Science Ltd., version 2.4). BSA and mixed protein samples samples were searched against the self-assembled Swissprot database, for biological samples the Swissprot human database was used in its most recent version. Search parameters were set as following:

- enzyme: trypsin,
- max. missed cleavages: 2
- peptide charge: +1, +2 and +3
- peptide tolerance ± 0.8 Da
- fixed modifications: carbamidomethyl on cysteine
- variable modifications: oxidation on methionine

If not stated otherwise, the following proteomic protein identification parameters were employed for data assessment of MS results:

- MOWSE score:
- Sequence coverage
- Number of "significant sequences" (number of protein specific unique peptides)

3 Results & Discussion

3.1 Optimization of Separation Conditions for Reversed Phase HPLC

For separation of both standard and biological samples, we applied a ternary gradient with mobile phases containing ACN, MeOH, trifluoroethanol (TFE) and FA as ion pairing agent. Despite the fact that in most similar works published, a binary gradient with ACN as organic solvent is used, we expected certain benefits from our eluents. The combination of the aprotic solvent ACN with a protic solvent, MeOH, should provide higher selectivity. Furthermore, the use of the strong eluent TFE should help to reduce carryover effects coming from trap column and separation column.

To ensure optimal separation of a digested protein sample, an experiment has been conducted in order to find the optimal gradient for peptide separation. Two and one pmol respectively of BSA overnight digest have been used and different mobile phase gradients have been compared Initially, the gradient has been set up starting with 5% mobile phase B, then rising from 5% to 100% B from 21 to 26 minutes (min), and finally getting from 0 to 100% C from 26 to 51 min.

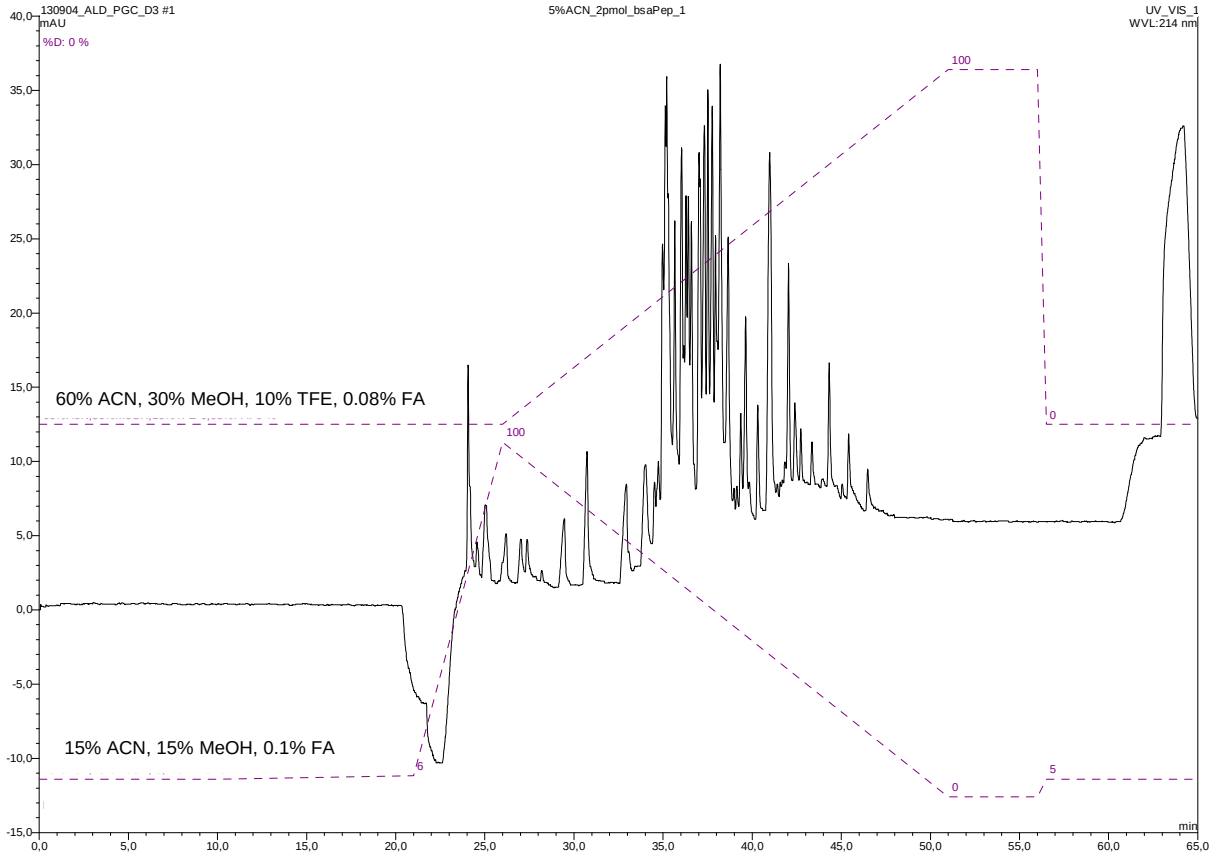

Figure 2: Chromatogram showing the separation of 2 pmol BSA digest with the following gradient: 0 to 21 min 5% B, 5% - 100% B from 21 to 26 min and 0% to 100% C from 26 to 51 min.

From the chromatogram seen in figure 2 it is obvious that the given gradient does not provide satisfactory separation. The majority of peaks are found at retention times between 35 and 39 min. Considering the delay of the gradient from the pump to the UV detector of about seven minutes, the mobile phase C (60% ACN, 30% MeOH, 10% TFA, 0.08% FA) is at very low percentages at these retention times.

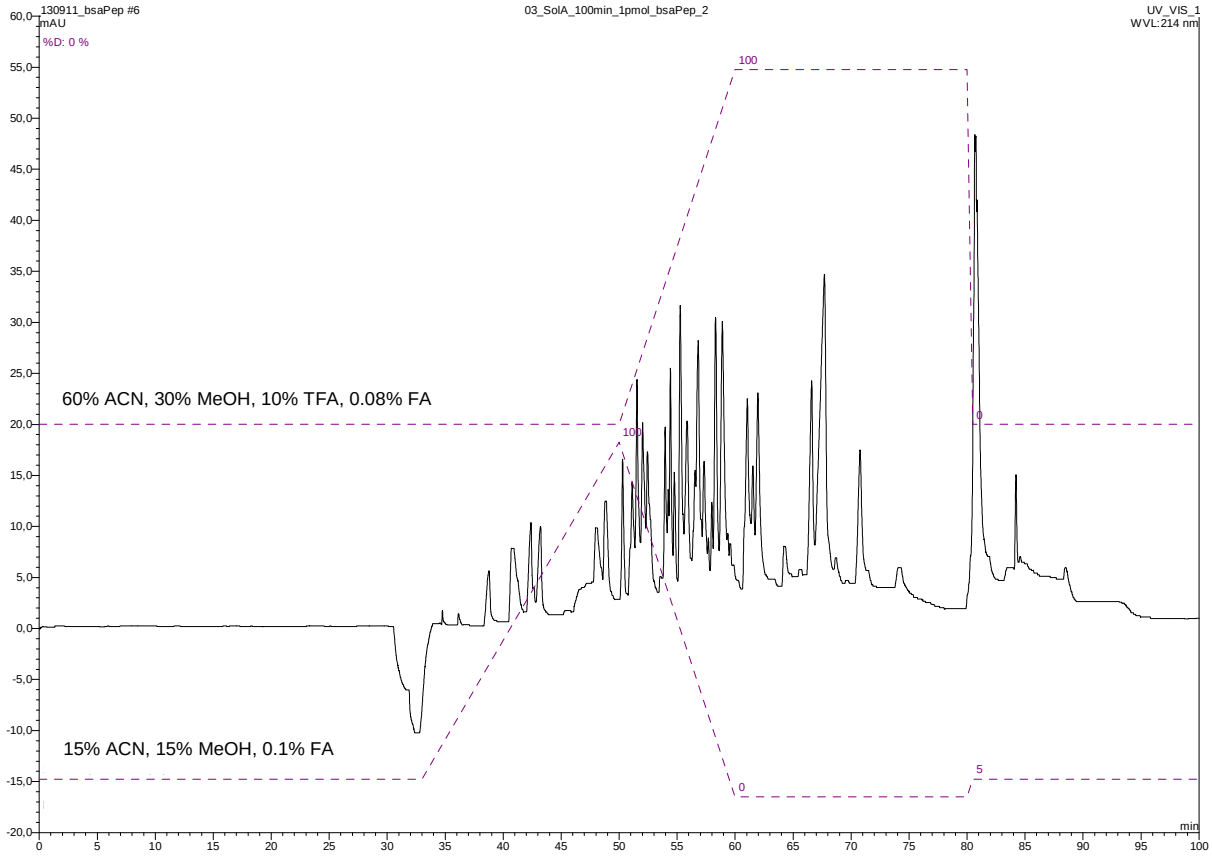

Figure 3: Chromatogram showing the separation of 2 pmol BSA digest with the following gradient: 0 to 34 min 5% B, 5% - 100% B from 34 to 50 min and 0% to 100% C from 50 to 60 min.

For the experiment shown in figure 3, the gradient of mobile phase B was extended from five and six minutes respectively to 16 minutes while the gradient of mobile phase C had been reduced to only ten minutes. Compared to the mobile phase gradient shown in figure 2, the separation is considerably better, suggesting that 100% B or 30% organic content already elutes the majority of tryptic BSA peptides. This experiment points out that the delay time is one of the most critical factors when setting up an elution gradient program.

In Figure 4, the UV-chromatogram of four sample-replicates using the gradient shown in Figure 3, is depicted. This figure shows the good reproducibility of the separation, which is crucial for the further experiments.

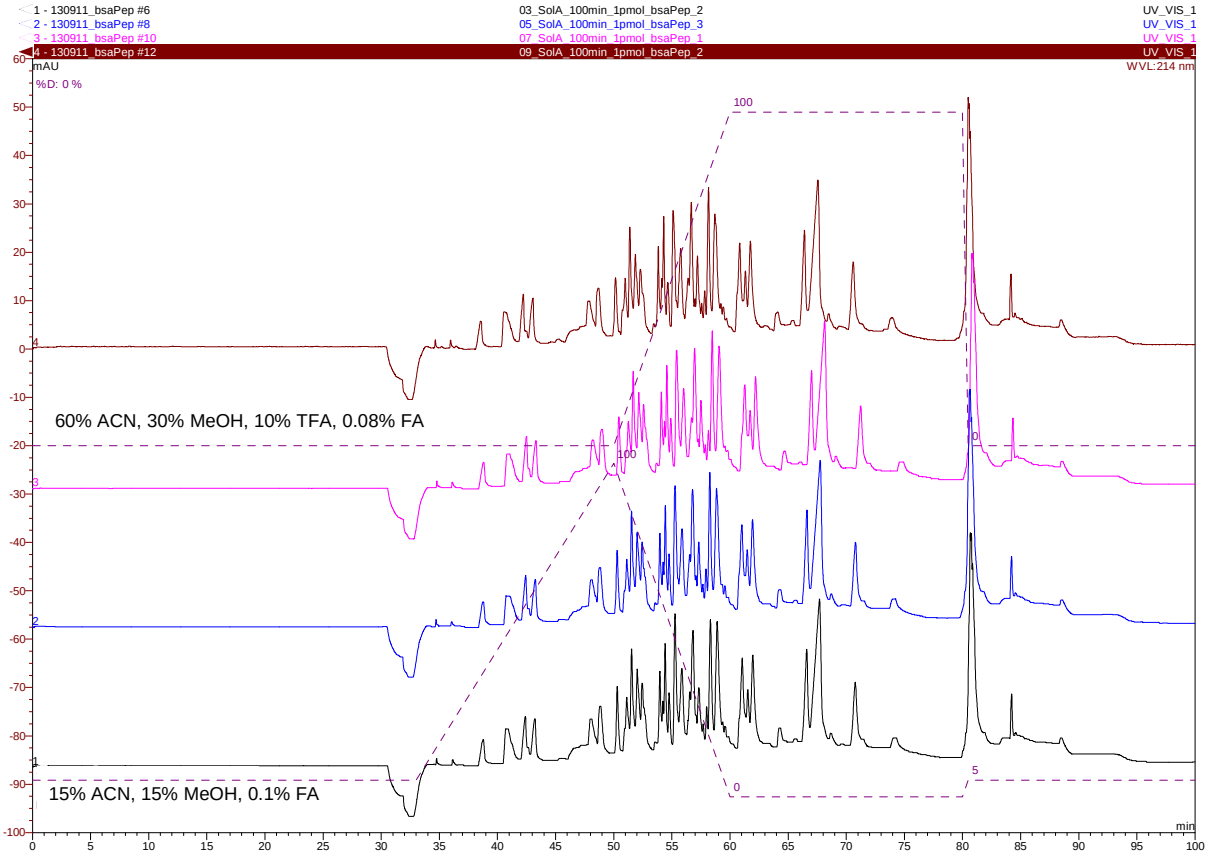

Figure 4: Chromatogram of four replicates using the newly developed gradient.

3.1 Optimization of Digestion Conditions

3.1.1. Sample preparation for online digestion

In order to improve the digestion efficiency, we have investigated whether the presence of certain detergents could support and improve the digestion process. For this purpose we chose Rapigest Surfactant, a commercially available detergent. Before using the substance for online digestion, it has been tested by adding it to a standard protocol offline digestion of BSA. 1 pmol of the resulting peptides was separated and analyzed by MS. The efficiency of digestion was addressed by the values of the protein identification parameters MOWSE score, sequence coverage and the number of significant sequences.

Sample Name	MOWSE score	Num. of sig. sequences	Sequ. Coverage (%)
1 pmol BSA peptides 1	831	27	45
1 pmol BSA peptides 2	1178	26	49
1 pmol BSA peptides 3	1138	28	56
1 pmol BSA peptides 4	1139	31	56
1 pmol BSA peptides average	1072	28	52
1 pmol BSA peptides RapiGest 1	1491	29	53
1 pmol BSA peptides RapiGest 2	1507	27	53
1 pmol BSA peptides RapiGest 3	1797	34	60
1 pmol BSA peptides RapiGest 4	1460	29	54
1 pmol BSA peptides RapiGest average	1564	30	55

Table 1: Impact of RapiGest in an offline digestion experiment. Four technical replicates were conducted

It can be seen that the detergent improves the digestion efficiency. One can conclude that the surfactant was effective in making the cleavage sites more assessable for the protease. The MOWSE score for the RapiGest treated samples was found to be 45% higher when using RapiGest surfactant. Both the number of significant peptides and the overall sequence coverage are slightly higher (30 compared to 28 and 55 compared to 52 respectively.

It has been tested as well whether the addition of RapiGest surfactant could aid in the process of online digestion. Comparison was made with urea treated and untreated samples, and samples treated with DTT and IAA only, in order to find out an appropriate sample preparation procedure for online digestion.

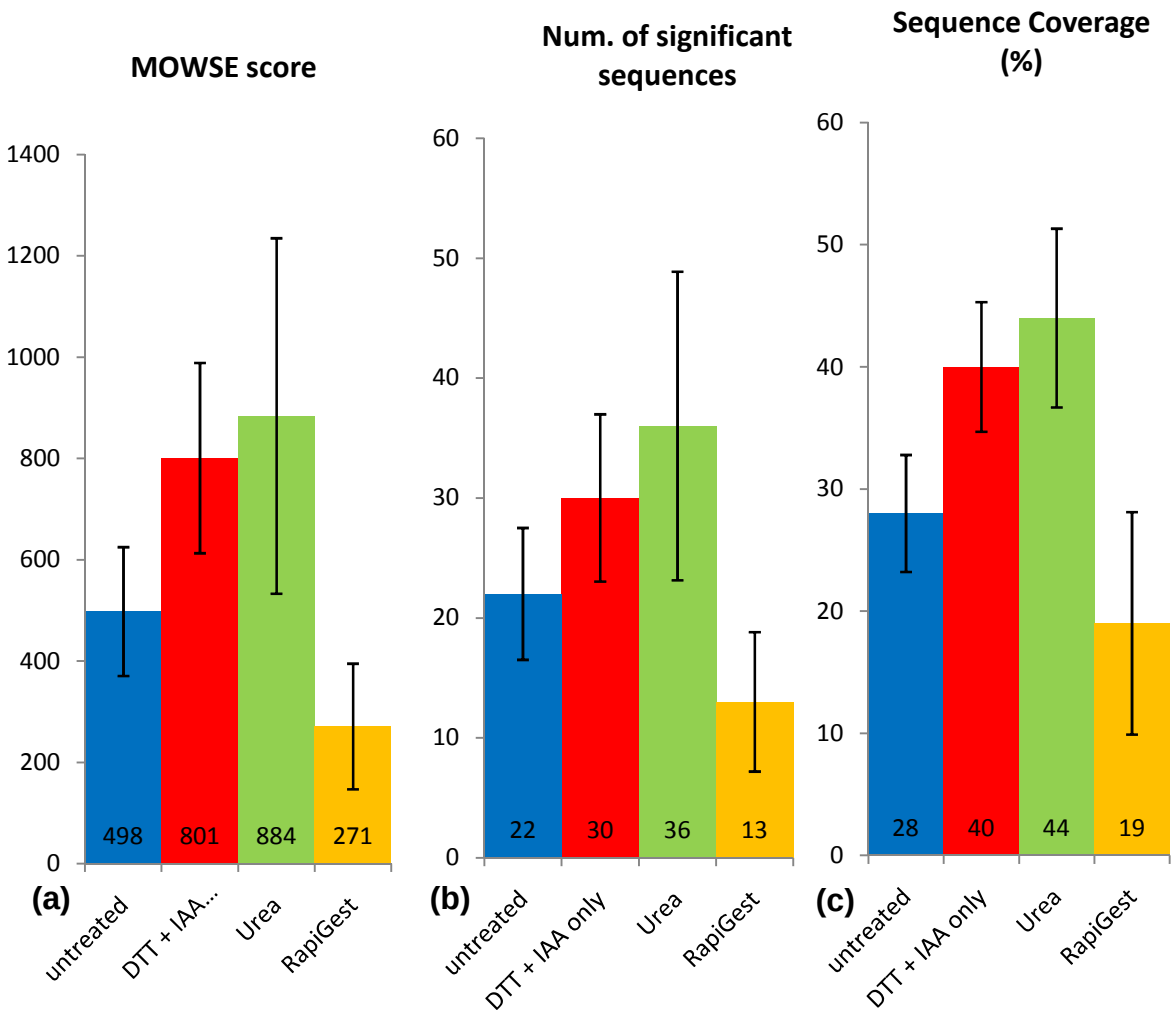

Figure 5: Comparison of sample treatment prior to online digestion; (a) shows the MOWSE score, (b) the number of significant sequences and (c) the sequence coverage. Shown values are the average of four or, in the case of RapiGest, five technical replicates. Error bars illustrate the standard deviation.

Unlike to the result from offline overnight digestion, the RapiGest treated samples give very low MOWSE scores, lower even than the untreated BSA samples. The same is true for the number of significant sequences as well as the overall sequence coverage. While working well in the classical overnight approach, the RapiGest surfactant somehow interferes with the process of online digestion. Urea seems to be the best choice for this purpose, even though the results for the samples treated with urea are insignificantly better compared to samples treated with IAA and DTT solely.

3.1.2. Optimizing the pH for digestion

The aim of the process is finding a compromise for the different needs of online digestion on the IMER and the subsequent trapping process on C-18 pre-columns. For trapping and separating on a silica-based C-18 column, the pH has to be below 8 [87] to prevent hydrolysis of the silica groups. It is in fact considerably lower due to the addition of organic acids, mainly TFA [78]. Trypsin activity, on the other hand, has an activity maximum between pH 7.5 and 8.5.

In an initial experiment, the impact of different pH values was evaluated using 300 pmol of a mixture of six proteins.

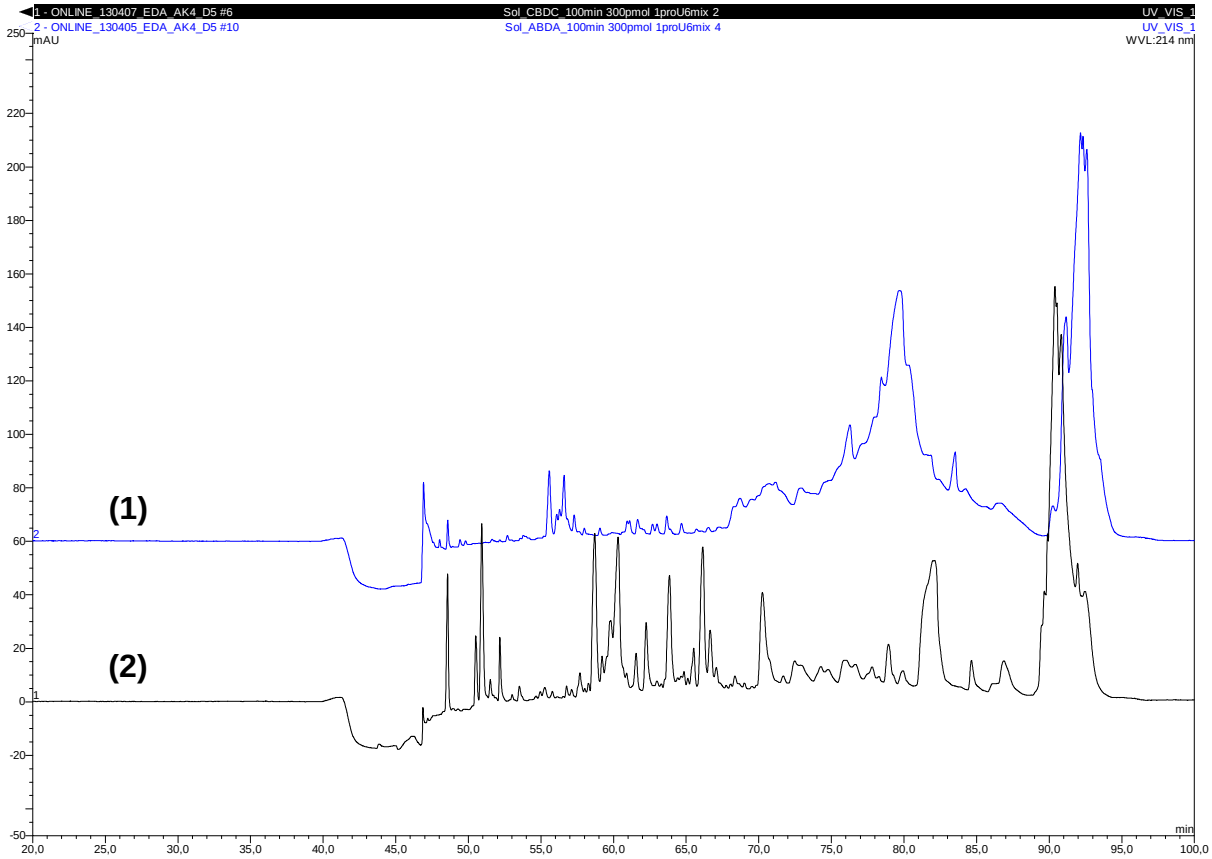

Figure 6: Chromatogram showing impact of pH value on online digestion: (1) pH 6, (2) pH 7.5

This experiment shows that while pH 7.5 is already high enough to provide sufficient trypsin activity, it is also adequate for trapping the sample in the pre-column. At pH 6 however, the enzyme activity has decreased considerably. The large group of peaks at 79 min is presumed to be poorly digested protein.

3.1.3. Choice of digestion buffers

3.1.3.1. Role of buffer ions and pH

TEAB, which is used as a standard buffer for overnight trypsin digestion, is compared against ammonium acetate. Both buffers are 50 mM in concentration and the pH was adjusted to 7.8. Three pmol of BSA had been injected per run.

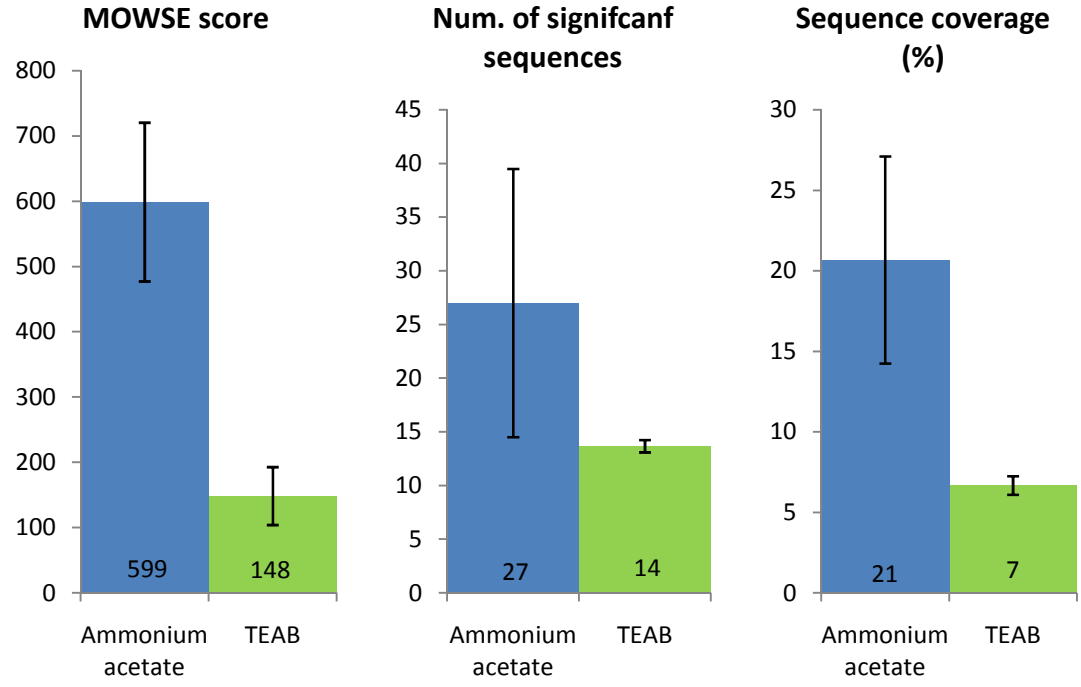

Figure 7: MS results comparing 50 mM ammonium acetate to 50 mM TEAB as digestion solution.

Considering the MS based proteomic identification parameters, there is a substantial difference between the results obtained with TEAB and ammonium acetate, apparent in all parameters. The average MOWSE score of the ammonium acetate samples is 599 compared to only 148 in the samples dissolved in TEAB. Also the sequence coverage (32%) and the number of significant sequences (15.7) were considerably higher than the respective values in the TEAB-samples (14% and 6.7 respectively). Despite being a widely used buffer for classical trypsin digestion, TEAB seems to interfere with one or the other steps within the process of online digestion. Despite of the fact that ammonium acetate is actually not buffering at pH 7.8, it is preferred over

TEAB because of its compatibility both with trypsin activity and ESI MS. Its buffer action is not crucial in the system used.

3.1.3.2. Role of ACN

Several groups report that the addition of organic solvents can lead to an increase in trypsin activity [42,163]. We also speculated that organic solvent components might reduce the unspecific attachment of proteins and peptides in the IMER, thus leading to a decrease of carry over effects. On the other hand, organic solvent components bring the risk of protein precipitation. We conducted the first experiments using 2% ACN as organic component in the digestion solution. Only a small amount of ACN was utilized initially because of the possibility of affecting the trapping capacity of the pre-column when using a higher organic content. Besides, the pH was adjusted to 8.5, which lies in the buffering range of ammonium acetate, to test whether the higher pH could improve digestion activity of the enzyme without reducing the trapping capacity of the pre-column.

These new digestion conditions were applied to a mixture of nine different proteins of different concentrations.

Protein	Molecular mass	Injected amount (pmol)	MOWSE Score	Num. of sign. sequences	Sequence coverage (%)
Sero-transferrin	75 kDa	5.00	837.0	25.5	44
BSA	66 kDa	4.87	535.0	18.0	36.5
lysozyme	14.3 kDa	5.19	197.5	5.5	44
ribonuclease-A	13.7 kDa	2.55	292.5	4.0	52
ADH	147 kDa	0.49	187.0	6.0	20.5
carbonic anhydrase	29 kDa	0.41	72.5	2.0	7.5
cytochrome C	12 kDa	0.50	54.5	1.0	15.5
β-casein	24.5 kDa	0.49	53.0	1.0	2.5
ubiquitin	8.5 kDa	0.51	0	0	0

Table 2: Proteomic identification parameters for nine proteins present in a mixture. The positive identification threshold value for the MOWSE score is 40.

The different proteins had a concentration range of one order of magnitude, and as expected, the higher concentrated ones received higher MOWSE scores and more significant peptides were detected. Ubiquitin, present in the sample solution in a small concentration only, could not be identified. It is the smallest protein in the sample and thus it has the fewest number of potential peptides. This and the fact that three proteins are close to the threshold MOWSE score of 40, shows that the sensitivity and dynamic range of the method is still to be increased.

In order to be able to increase the pH without damaging the trap column, a T-piece has been introduced after the IMER, so that the digestion solution could be mixed with 0.1% TFA and the pH of the digestion solution was decreased after the IMER in order to ensure an optimal pH for operating the trap column. Furthermore, this procedure allowed us to increase the ACN content to 10% without allowing too high a concentration of organic solvent to reach the trap column.

The next experiment was conducted in order to determine the impact of different ACN concentrations in the digestion solution. Slysz and Schriemer showed that organic solvent concentration of up to 45% can yield good results for online digestion [165]. Besides, ACN was compared to methanol regarding its ability to augment the digestion process. The digestion solutions were mixed with 0.1% TFA after the IMER, and the flow rate was varied so that the same solution reaching the trap column would have an ACN content of 2.5%. The digestion time was set accordingly. Digestion buffer containing 20% ACN was also run at 4 µl/min in order to determine the impact of the flow rate in this experiment. The overall flow rate when reaching the trap column was 20 µl/min in every experiment.

Sample	**MOWSE score**	**Num. of significant sequences**	**Sequence Coverage (%)**
10% ACN 4 µl/min sample	925	24	49
10% ACN 4 µl/min blank	282	8	16
15% ACN 3 µl/min sample	1291	24	44
15% ACN 3 µl/min blank	411	10	24
20% ACN 4 µl/min sample	1241	17	34
20% ACN 4 µl/min blank	516	9	22
20% ACN 2 µl/min sample	1466	20	42
20% ACN 2 µl/min blank	300	6	12

Table 3: Different amounts of organic solvent in the digestion solution have an impact on the digestion efficiency. Numbers shown are the average form four runs for every condition.

Apparently, the data show that higher concentrations of ACN do slightly improve the digestion efficiency of the IMER at given conditions. This is not unexpected, as proteins tend to unfold in presence of organic solvents, exposing their cleavage sites. Under both conditions i.e., 15 and 20% ACN, higher scores are achieved than with 10% ACN. When comparing the conditions with 20% ACN and different flow rates, one sees that at the lower flow rate the result is better considering both sample and carry-over score. The ratio of sample score compared to blank score is only 20% at 2 µl/min compared to 42% at 4 µl/min. The impact of different digestion flow rates will be addressed in depth later. This experiment also shows that 5% ACN does not decrease the ability of the trap column to bind the peptides arriving from the IMER. The use of 10% methanol instead of 10% ACN has brought no improvements, and no further experiments using higher methanol concentrations have been conducted because of the high viscosity of methanol/water mixtures and the resulting high system backpressure.

3.1.3.1. Role of Urea

It has been reported that the addition of chaotropic reagents, especially urea, can improve digestion efficiency in IMERs [164] and therefore urea was added to the digestion buffer in a concentration of 2 M. The concentration of ammonium acetate was increased to 100 mM, to amplify the buffering capacity of the digestion solution.

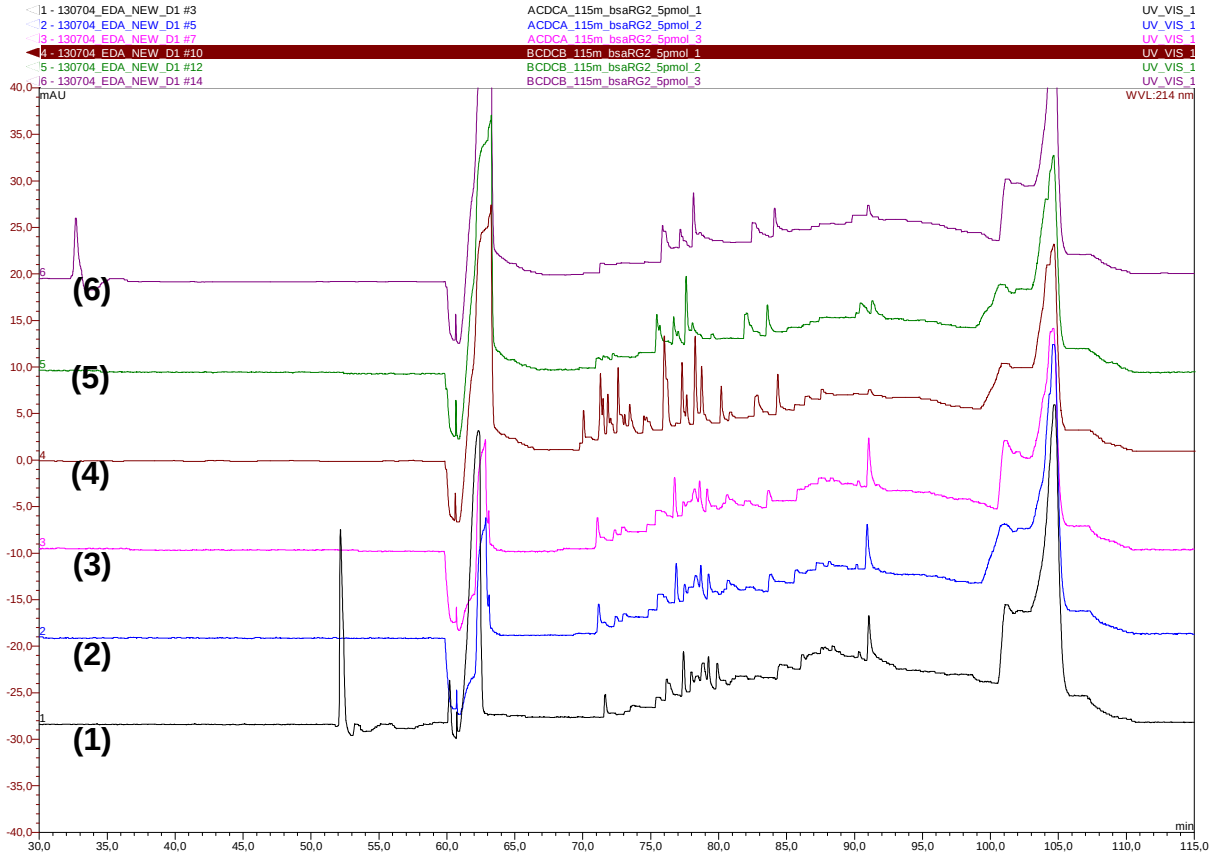

Figure 8: RP-UV Chromatograms of 5 pmol BSA with different digestion buffers: (1-3): 100 mM ammonium acetate, 2 M urea, 10% ACN, pH 8.5, three technical replicates; (4-6)50 mM ammonium acetate, 2% ACN, pH 8.5, three technical replicates.

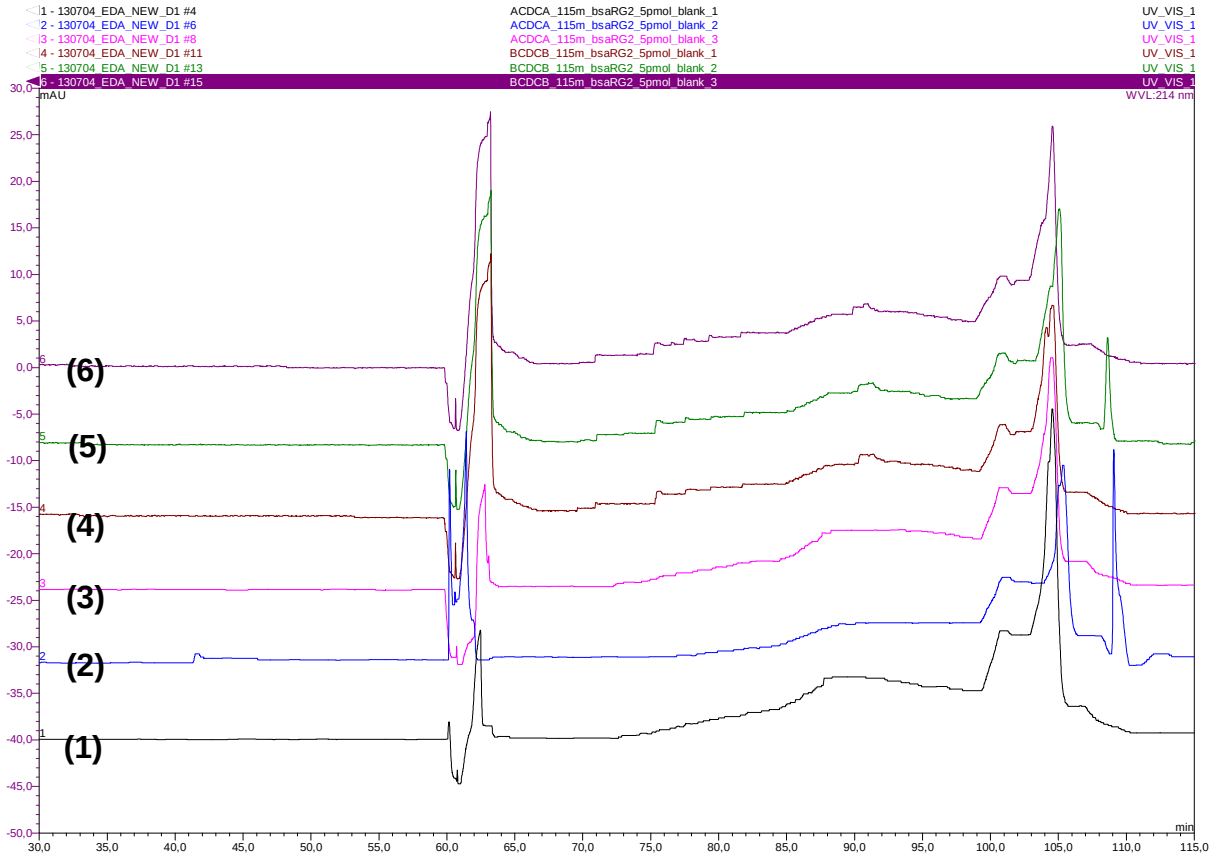

Figure 9: Comparison of the blanks of modified digestion buffer: (1-3): 100 mM ammonium acetate, 2 M urea, 10% ACN, pH 8.5; (4-6)50 mM ammonium acetate, 2% ACN, pH 8.5)

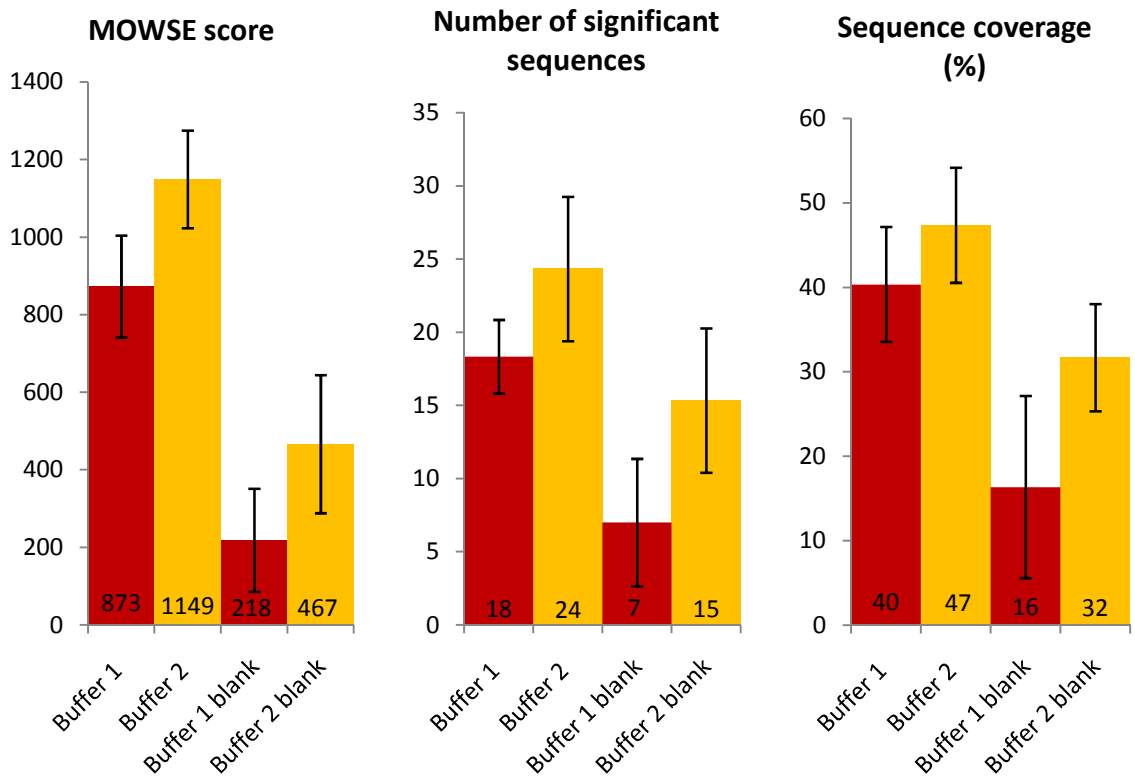

Figure 10: Protein identification parameters of 5 pmol BSA with modified digestion buffer (buffer 1; 100 mM ammonium acetate, 2 M urea, 10% ACN, pH 8.5; 1-3) compared to previously used digestion buffer (digestion buffer 2; 50 mM ammonium acetate, 2% ACN, pH 8.5). The values are the average of three technical replicates.

The UV chromatograms of the samples, seen in figure 8, show more peaks with the urea containing buffer, and very few minor peaks can be seen in the blanks without urea but not in the blanks with urea. Surprisingly, the MS data show that the digestion efficiency seems to be higher when the 50mM ammonium acetate buffer without urea had been used. The average MOWSE score of these samples is 1149 compared to 875 in the 100 mM ammonium acetate buffer containing urea. On the other hand, the MOWSE scores of the blanks with the urea buffer are considerably lower with 217 in average, representing 25% of the average sample score. The blank scores of the non-urea buffer average 466, representing 40% of the average sample score. Considering the data of these two different conditions, there seems to be a compromise to be made between IMER activity and carryover effect, taking into account that the carry over effect is increasing to a higher degree than the IMER efficiency.

In the next experiment, based on the digestion buffer containing 100 mM ammonium acetate, 2 M urea, 10% ACN, pH 8.5, the urea has been omitted and compared to the digestion solution containing 50 mM ammonium acetate and 2% ACN.

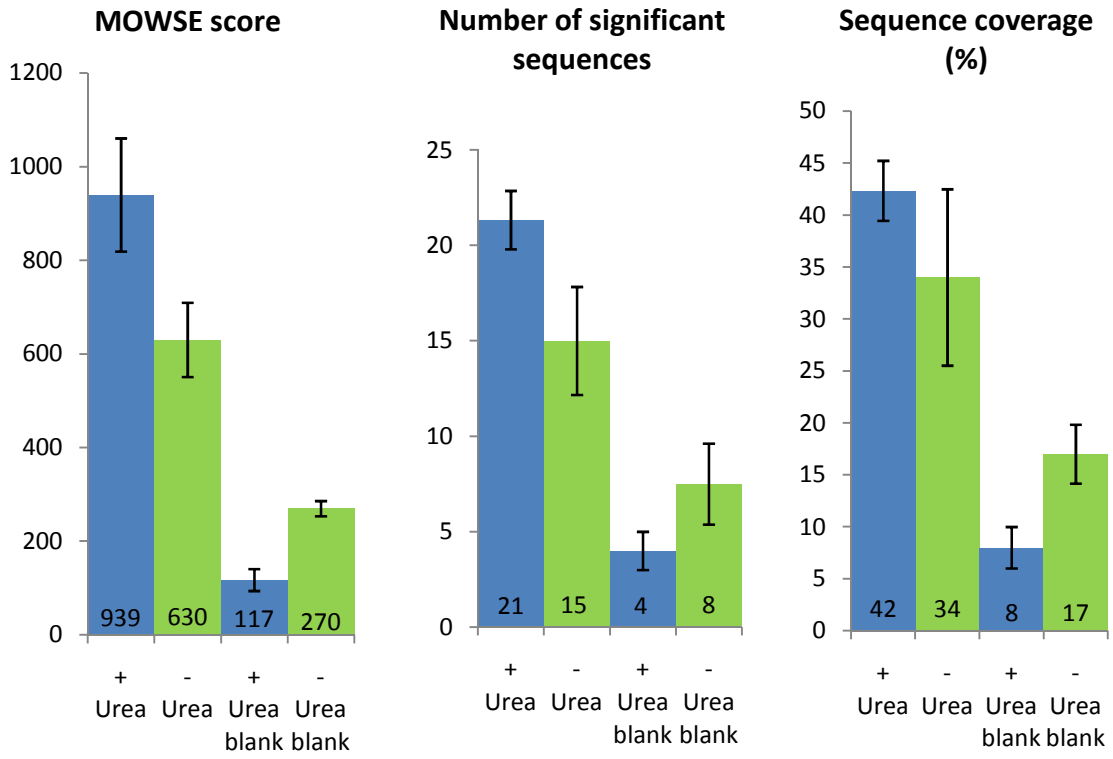

Figure 11: Proteomic identification parameters showing the impact of urea content in the digestion buffer. The blank runs are made with buffer injections instead of samples. The values are the average of three technical replicates.

Regarding the UV spectra, peaks are somewhat higher in the samples without urea, whereas the blanks cannot be distinguished. Using the MOWSE score as criteria for the digestion efficiency, this value is significantly higher for samples digested with the buffer containing urea, being 939 compared to a score of 630 for the samples digested without urea. The blanks after a sample with the urea buffer also performed better in terms of carry over, having an average score of only 117 or 12% of the sample score respectively, compared to a score of 270 or 43% of the average sample score. The higher digestion efficiency of the buffer containing urea can be understood by the chaotropic nature of urea which causes the protein being denatured and thus it adopts an extended conformation and exposing the cleavage sites to the protease. By enhancing protein solubility, urea also seems to minimize the attachment of the protein to the IMER matrix, improving the carry over issue. It seems

probable that the digestion efficiency of the IMER could still be increased using even higher urea concentrations. It has been reported that immobilized trypsin remains active even at urea concentrations of up to 8 M. However, a digestion solution containing 2 M urea already causes problems by crystallization in capillary systems and blocking the LC tubings. Thus we have refrained from using digestion solutions with higher concentrations of urea.

The digestion conditions established so far were applied on a six protein mix of equimolar concentration, consisting of BSA, ribonuclease A, cytochrome C, ubiquitin, lysozyme and carbonic anhydrase. Of the sample solution 15 and 100 pmol/µl of overall protein was injected, according to a concentration of 2.5 and 16.67 pmol respectively per single protein.

	15 pmol			**100 pmol**		
Identified protein	**MOWSE score**	**Nr. of sig. Peptides**	**Percentage of positive identification in replicates**	**MOWSE score**	**Nr. of sig. Peptides**	**Percentage of positive identification in replicates**
BSA	142.00	3.00	4/4	166.33	4.33	3/4
RNAS	237.00	2.50	4/4	475.75	5.00	4/4
CYC	166.25	3.25	4/4	215.25	4.00	4/4
UBIQ	56.75	1.25	4/4	159.50	2.75	4/4
LYSC	71.00	1.00	2/4	-----	-----	0/4
CAH	-----	-----	0/4	58.33	1.00	3/4

Table 4: Proteomic identification parameters of the digestion of 15 and 100 pmol of a six protein mixture, showing the average of four runs per injection volume

With both injection amounts, one protein could not be identified; carbonic anhydrase in the 15 pmol injection and lysozyme in the 100 pmol injection. Altogether, the results of the 100 pmol injections did not exceed the results of the 15 pmol injections to the expected rate.

3.1.4. Influence of digestion temperature

Protein digestion using trypsin is usually conducted at 37°C resembling physiological conditions. However, it has been reported that trypsin activity can be, under certain circumstances, increased to 60°C or even 72°C [166]. Therefore, we investigated the influence of an increase in digestion temperature to 60°C compared to 37°C. To conserve the immobilized trypsin from denaturation at 60°C, $CaCl_2$ was added to the buffer to a concentration of 20 mM. $CaCl_2$ is reported to protect the enzyme at higher temperatures and also to increase its activity [167]. As a control, $CaCl_2$ containing buffer was also used at 37°C.

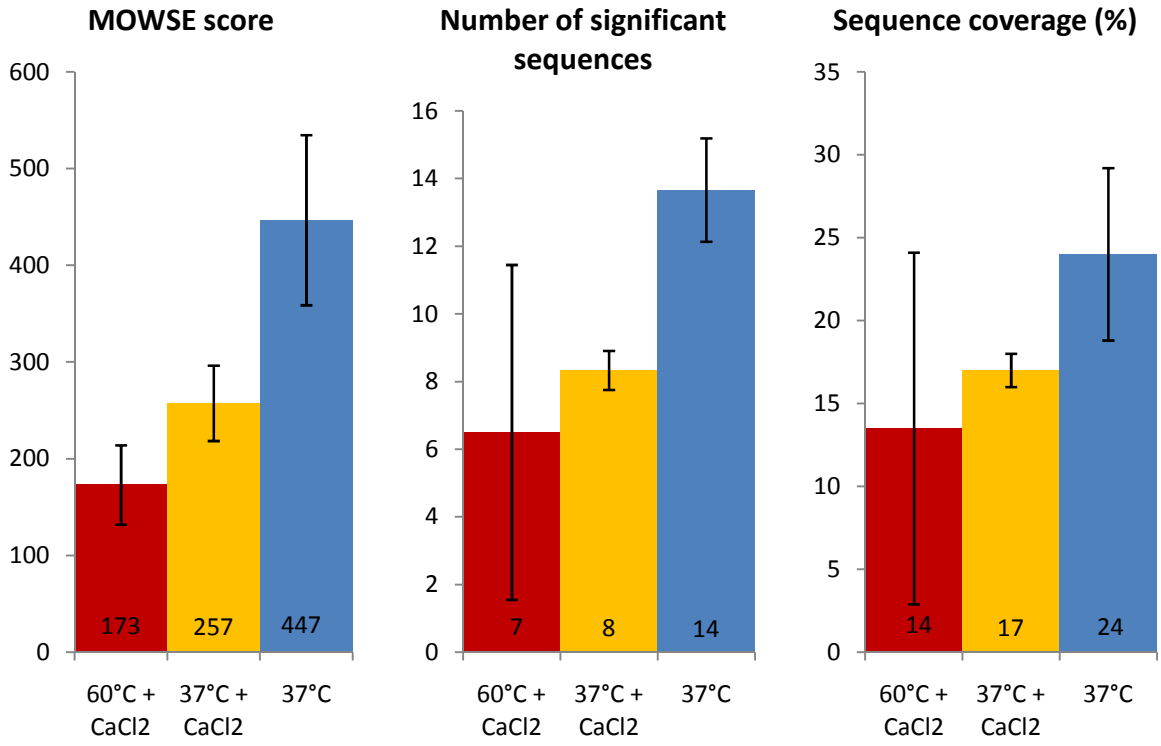

Figure 12: Comparison proteomic identification parameters of digestion temperatures.

The IMER has shown to have a much higher activity at 37°C compared to 60°C according to this experiment. $CaCl_2$ seems to have a negative impact on the activity, as the average score is considerably lower as to when no $CaCl_2$ is used.

3.2 Influence of flow rate on IMER activity

Digestion time is one of the decisive parameters in tryptic digestion experiments. In standard digestion protocols, it varies between two and 24 hours. For digestion in an IMER, less time is needed than for in solution digestion because of the much higher enzyme to sample ratio. It has been reported that reaction times of as low as a few seconds can already be sufficient to provide acceptable digestion of the sample [42]. The linear flow rate through the IMER is the parameter with which the digestion time in our system can be controlled. This linear flow rate depends on the volume flow rate (µl/sec) and the inner diameter of the monolithic IMER-bed. Two different aspects exist concerning the impact of flow rate on online digestion efficiency. On one hand, lower flow rates provide longer digestion times, on the other hand, higher mass transfer rate of enzyme-substrate complexes when using a higher flow rate, are supposed to enhance enzyme-substrate complex formation [168].

To investigate the effect of flow rate and thus digestion time on digestion efficiency, 3 pmol of reduced and alkylated BSA were applied on the IMER operated under different flow rates, and the protein identification parameters were compared. Based on a supposed IMER volume of 100 µl, the flow rates of 10 µl/min, 7 µl/min and 4 µl/min correspond to a digestion time of 10 minutes, 14 minutes and 25 minutes respectively.

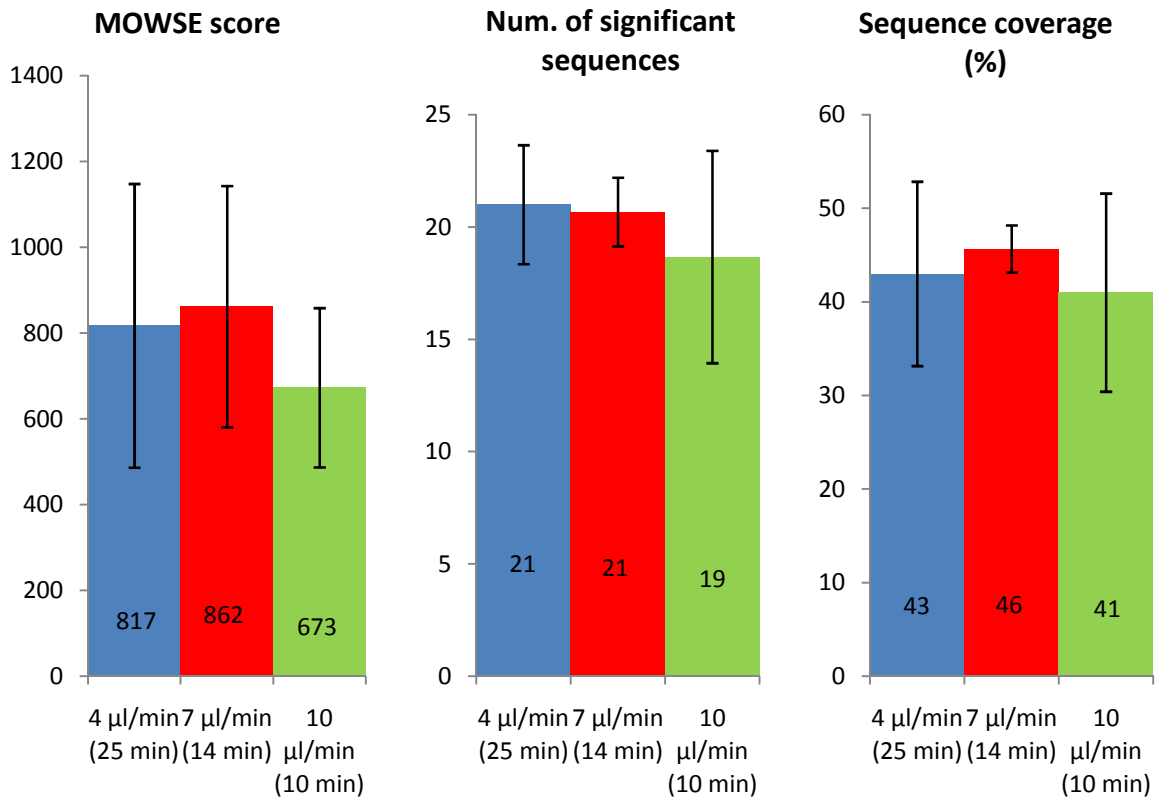

Figure 13: Effect of different flow rates/digestion time intervals on digestion efficiency

We anticipated that lower flow rates would enhance digestion efficiency due to longer contact between substrate and enzyme. In fact, with lower flow rates corresponding to digestion times of 14 minutes and below, slightly higher MOWSE scores and a higher number significant sequences were found. Overall, the effects were small and when using the higher digestion times than 14 minutes, no benefit was seen.

3.3 Minimizing Carry-over Effects

3.3.1. The sources of carryover

First the sources of carryover effects in the LC system independent of the IMER were investigated, using already digested BSA. Two different kinds of blank runs wre performed: one with injection of buffer and another without injection, circumventing the injection system of the autosampler. The blanks without injection were run through the separation column only and through the separation column and blank column. Blanks with buffer injection were run through the separation column and the trap column.

SAMPLE	MOWSE score	Num. of significant sequences	Sequ. Coverage (%)
1 pmol BSA peptides replicate 1	2602	34	60%
Blank separation column only (no injection)	181	7	14%
1 pmol BSA peptides replicate 2	1953	33	56%
Blank separation column + trap column (no injection)	192	7	16%
1 pmol BSA peptides replicate 3	2272	33	54%
Blank separation column + trap column + autosampler	516	17	39%
1 pmol BSA peptides replicate 4	2347	37	59%
Blank separation column (no injection)	291	8	15%

Table 5: Protein identification parameters in dependence on the accessibility of the incremented subsystems in order to locate the source of carryover when separating tryptic peptides from BSA

The MOWSE-score of the blank run accessing the separation column only averages 181, compared to 192 when the trap column was included. The trap column therefore appears to have no memory effect and the separation column is a minor source of carry over. When looking at the score of the blank including the injection device, i.e. 516, it is obvious that the autosampler is the major source of carryover in the system without IMER.

The purpose of the following experiment was to determine the nature of the carryover. In order to achieve this goal, samples of pre-digested BSA were run through the IMER, and the blanks were analyzed and then compared to our standard approach with protein sample. 3 pmol of both peptides and proteins were injected. To ensure that no carryover from previous experiments was measured, a new IMER was used.

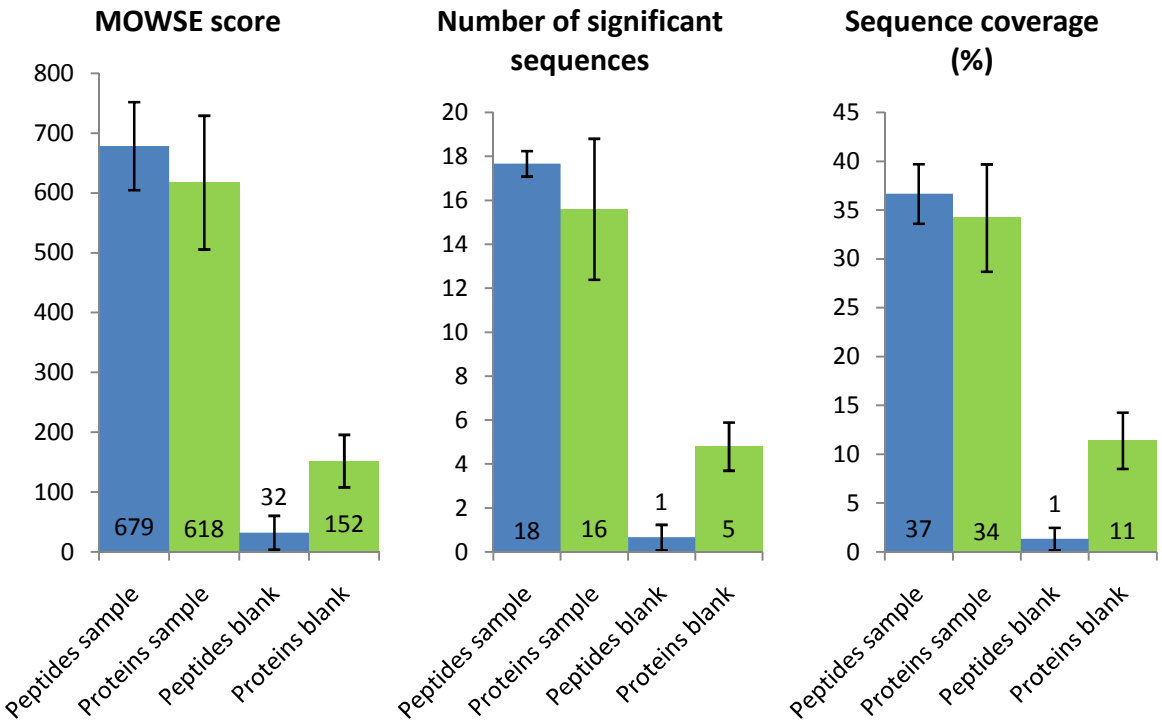

Figure 14: Carry over effects resulting from the IMER measured for pre-digested peptide samples and protein samples. Results constitute the average value of four technical replicates

The average blank score of the peptide samples is 32, being only 4.7% of the sample score, averaging 679. The blank score of the protein sample, on the other hand, averages 152 which are 24.6% of the overall sample score, 617. While achieving similar MOWSE scores with protein and blank samples, showing that online digestion is just as effective as the overnight approach, the carryover effect is obviously higher

with the protein samples. This finding supports the conclusion that most of the carryover is actually caused by proteins rather than peptides. This is supported also by the average number of peptides found, which is 5 for the protein samples and 1 for the peptide samples. If this conclusion is true, it would imply that the carryover could be reduced by further increasing the digestion efficiency as well as preventing adsorption of proteins on surfaces of the chromatographic system in front of the IMER or within the IMER.

3.3.2. Reducing Memory Effects

For finding appropriate washing conditions, an experiment has been conducted to compare the effect of washing the IMER with 2 M ammonium acetate solution, pH 4.8 to organic solvent consisting of 50% ACN; 30 % methanol and 20 % TFE.

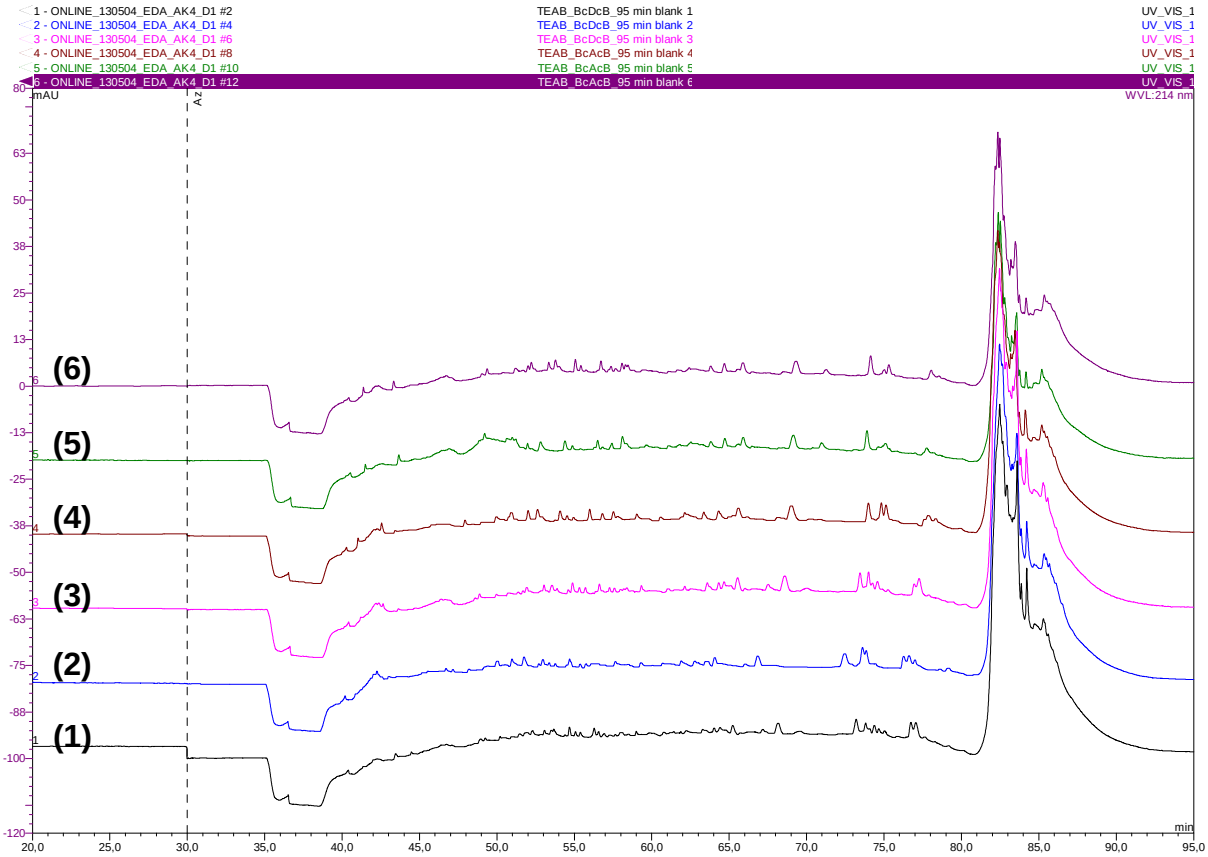

Figure 15: The chromatograms of the blanks from organic wash step (1-3) and salt wash step (4-6) compared.

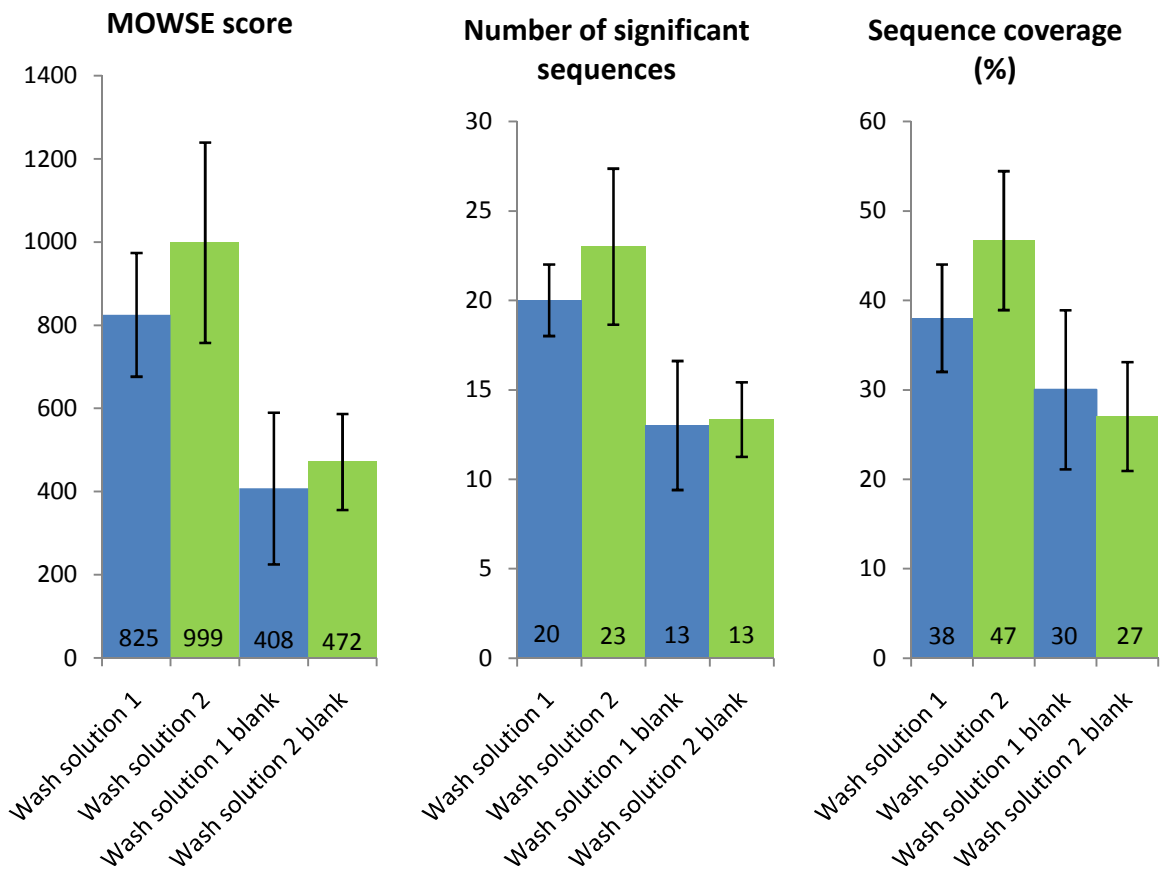

Figure 16: Proteomic identification parameters comparing the effectiveness of organic (wash solution 1; 50% ACN; 30 % MeOH, 20 % TFE) and salt wash solution (wash solution 2; 2 M ammonium acetate, pH 4.8) regarding the reduction of the memory effects. Results give average values of three technical replicates)

The outcome of the experiment suggests that neither washing with the chosen organic phase nor with salt solution could improve the carry over issue. The MOWSE score of the blanks obtained when the IMER was washed with organic solvent averages 407 or 49% of the average sample score compared to 471 or 47% respectively when ammonium acetate solution was used.

For testing our hypothesis that an improvement in digestion efficiency would enhance the carry-over behavior of the system, we investigated the issue at a later state of the experiment. To wash the IMER, only digestion solution (100 mM ammonium acetate, 2 M urea, 10% ACN, pH 8.5) was used.

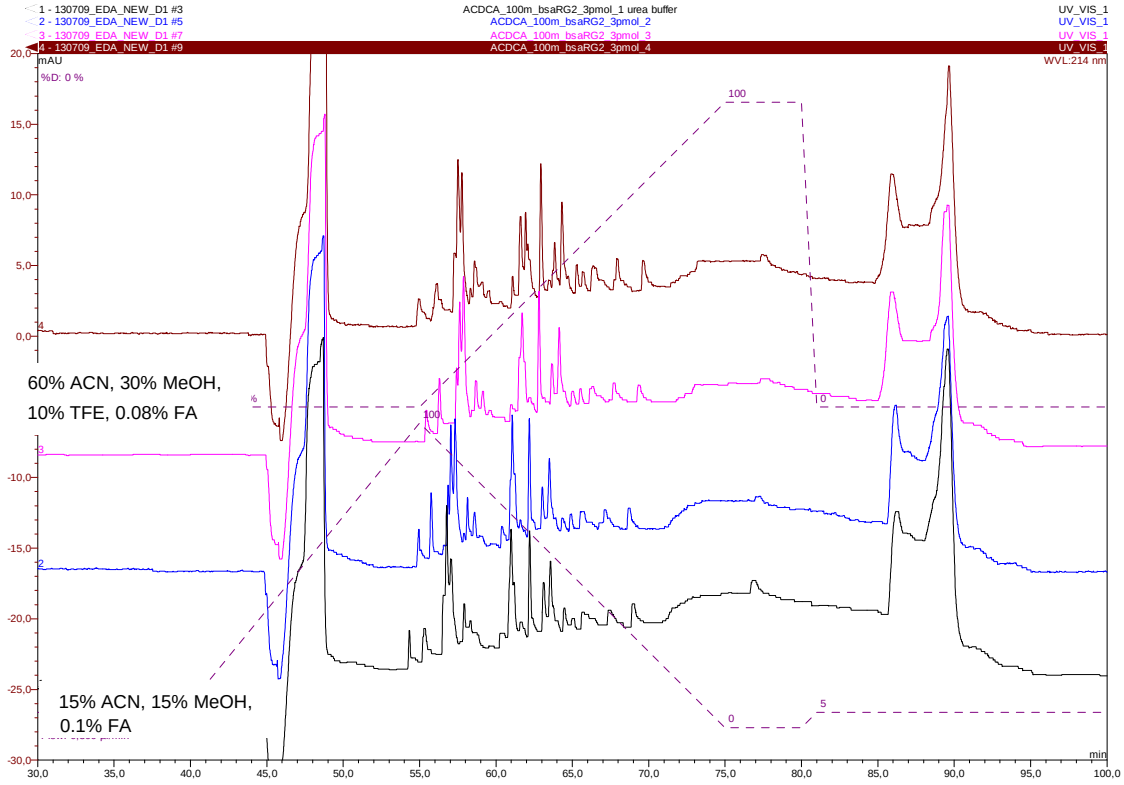

Figure 17: UV chromatograms of 3 pmol BSA online digest of four technical replicates

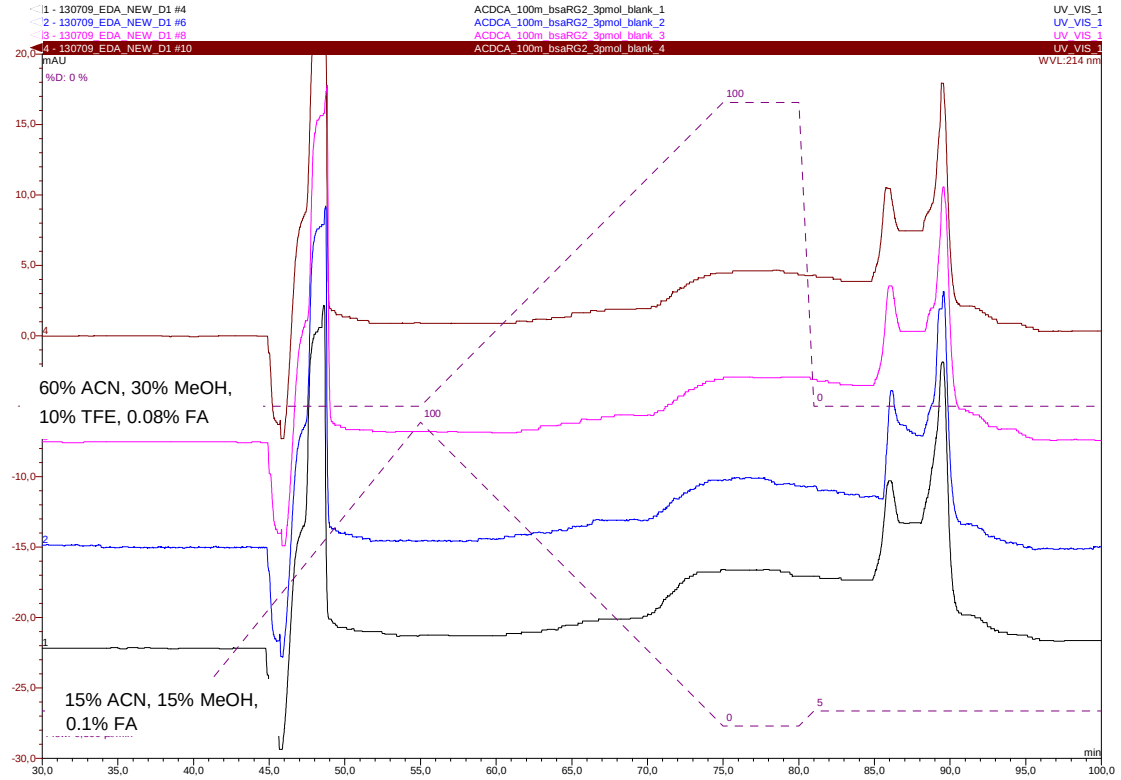

Figure 18: Blanks of 3 pmol BSA samples

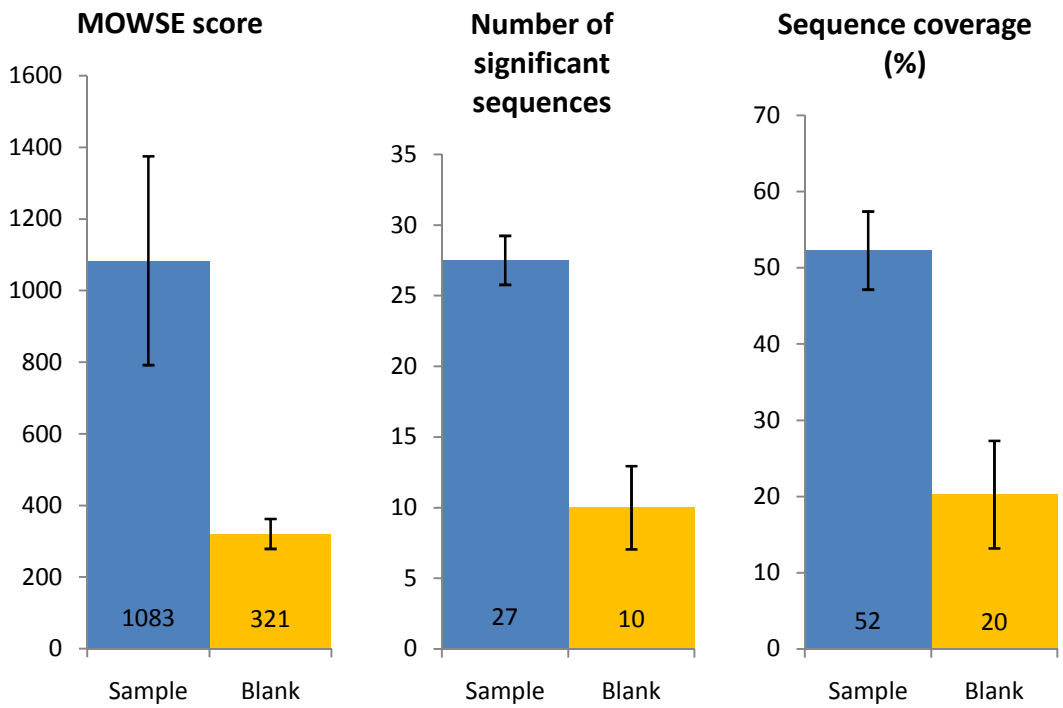

Figure 19: Mascot results of 3 pmol BSA samples and blanks when washing IMER with urea containing solution (equal to the digestion solution)

The average MOWSE score of the blanks represents 321 or 30% of the sample score which is 1081. When comparing to the previous experiment, the carryover effect is less pronounced, while the digestion efficiency has strongly increased. 3 pmol give a slightly higher MOWSE score than the previously used 12 pmol (1083 compared to 912), whereas the blank score is lower (312 or 30% compared to 471 or 47% with salt washing step and 407 or 49% with the organic washing step respectively).

3.4 Dynamic Range

The ability of the system to detect low abundant proteins was tested, using a mixture of eight proteins. In this mixture, the most abundant protein is about 4000 times more concentrated compared to the least abundant, constituting a dynamic range of three orders of magnitude. In a first step, an equimolar mixture of the proteins has been tested.

	CAH	β-Casein	Cyt C	Lyso-zyme	Sero-TF	RNAse	ADH	BSA
Molecular mass	29 kDa	25 kDa	12 kDa	14.3 kDa	75 kDa	13.7 kDa	145 kDa	66 kDa
Injected Amount (pmol)	10,91	13,13	13,37	17,23	13,43	13,82	13,13	13,09
Number of sig.Sequences	14	5	5	6	50	5	18	42
Sequence Coverage (%)	68	32	44	57	71	53	58	69

Table 6: Results of online digestion of an eight protein mix with equimolar concentration; values for sequence coverage and number of significant sequences are the average of three replicates

As shown in table 6, when injecting about 10 pmol per protein, all the different proteins could be identified with satisfactory sequence coverage. To address the issue of unequal amounts present, a mixture of the same proteins was tested, with injection amount varied between 0.02 and 85 pmol.

	CAH	β-Casein	Cyt C	Lyso-zyme	Sero-TF	RNAse	ADH	BSA
Molecular mass	29 kDa	25 kDa	12 kDa	14.3 kDa	75 kDa	13.7 kDa	145 kDa	66 kDa
Injected Amount (pmol)	0,02	0,09	0,17	6,32	2,67	7,79	0,07	84,75
Number of sig.Sequences	-----	-----	-----	3	29	5	9	64
Sequence Coverage (%)	-----	-----	-----	30	52	53	34	83

Table 7: Results of online digestion of an eight protein covering different protein concentrations; values for sequence coverage and number of significant sequences are the average of three replicates

The results of this experiment show that the method is not yet sensitive enough to cope with protein amounts below 1 pmol. In presence of high amounts of BSA, low abundant proteins with a concentration of 1:500 of BSA and lower are not detected. An exception is ADH, which was actually found at a concentration of about 1:1200 compared to BSA. This may be due to the fact that ADH (Mr ~ 145 kDa) is a rather large protein, yielding more tryptic peptides when digested compared to carbonic anhydrase (29 kDa), β-casein (25 kDa) and cytochrome C (12 kDa). Considering the fact that a high dynamic range in concentration is a key feature of clinical proteomic samples, the method needs to be improved in this respect.

3.5 Improvement of the LC Setup

The ultimate purpose of our LC-system is the fast, high throughput analysis of biological samples. In order to achieve this goal, we set up the instrument in a way which allows us to use two alternative IMERs, and two separation columns, each of which had a trap column connected upstream, as is depicted in Figure 20. In total, seven switching valves were implemented in the system. V1 was part of the autosampler unit. Via V2 the two IMERs could be used alternatively.V3 was used to connect/disconnect the IMER with the subsequent trapping and separating system. A T-piece is connected to V3, by which a 0.1% TFA solution is added to the digested sample by means of an additional pump. By V4, the sample was distributed to the two separate separation units. Using two separate nanopumps, it was possible to use the trap column and separation column connected to V5a to separate the sample, while the columns connected to V5b were equilibrated at the same time and vice versa. V6 was used to alternatively connect the digestion units 1 and 2 with the UV-detector and the MS.

Doing so, we were able to reduce the run time significantly and at the same time enhance the equilibration time for the trap columns and separation columns.

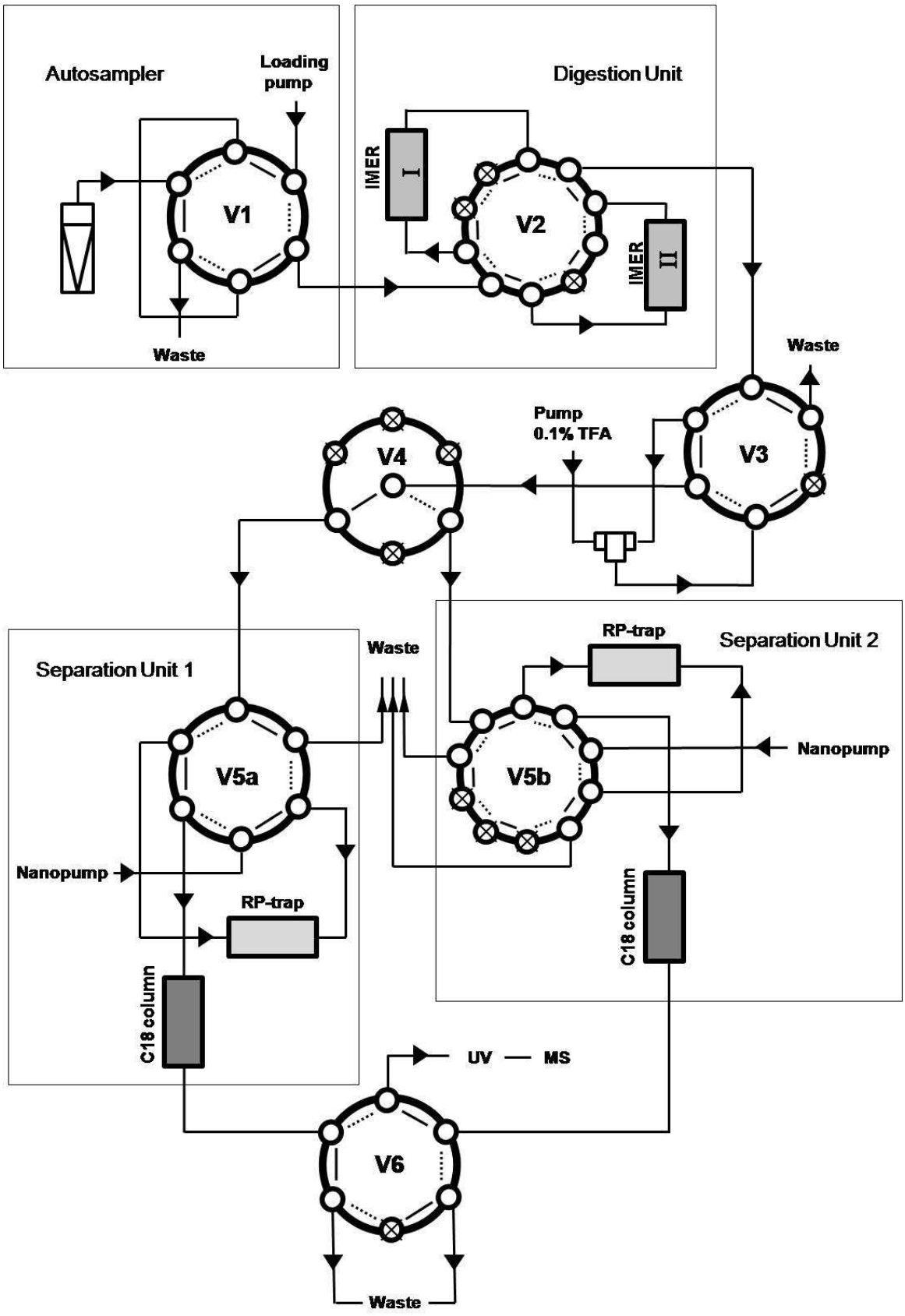

Figure 20: LC setup including two IMERs, two trap columns operated in back-flush mode and two separation columns in parallel.

3.6 Evaluation of the system setup by analyzing biological samples

3.6.1. Urine Samples

In order to test the system setup and the developed method for its suitability to analyse clinical samples, we compared its outcome to that of a classical offline approach using urine samples from cancer patients. Samples of 20µl were injected twice with an amount of 12 µg of protein for each run. The two approaches were performed on different LC-MS systems, so this variability should be considered when comparing the results. The purpose of this experiment was rather to get a first impression of the capability of our method.

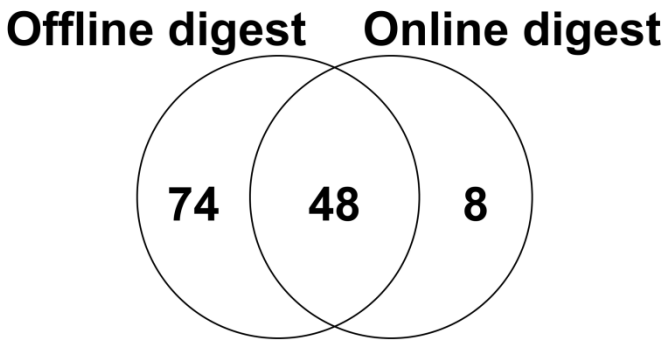

Figure 21: Number of proteins found in a urine sample with at least two identified peptides using online and offline digestion respectively.

Whereas 122 proteins could be found in the overnight digestion approach, 56 were identified in the on line digestion experiment. We conclude that digestion efficiency in the IMER and the sensitivity of the method still has to be enhanced in order to become competitive with the overnight digestion approach. Comparing the nature of the proteins detected, one sees that they are mostly congruent, since there are only eight proteins found solely by the online digest approach.

3.6.2. IVF Samples

In the initial experiment performed with IVF samples, four samples were analyzed. They were diluted to a total protein concentration of approximately 1 µg/µl. A volume of 6 µl, corresponding to about 6 µg of protein was injected for the online as well as the offline approach. The eluate fraction was not analyzed because no protein was found in the sample. The pre-digested samples were analyzed the same way as the undigested sample, with the exception that the sample was bypassing the IMER.

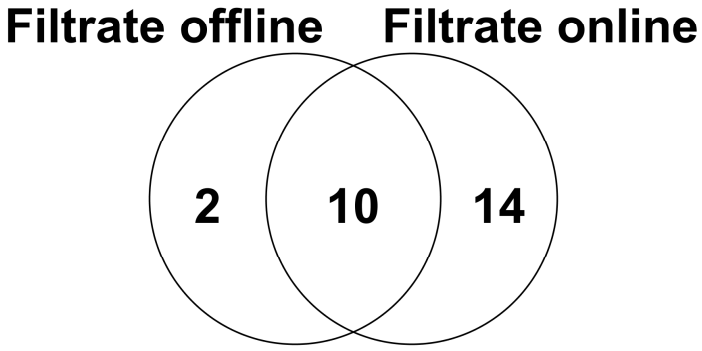

Figure 22: Venn diagram comparing online and offline results of the filtrate fraction of four samples. The numbers illustrate the count of proteins identified with two or more peptides

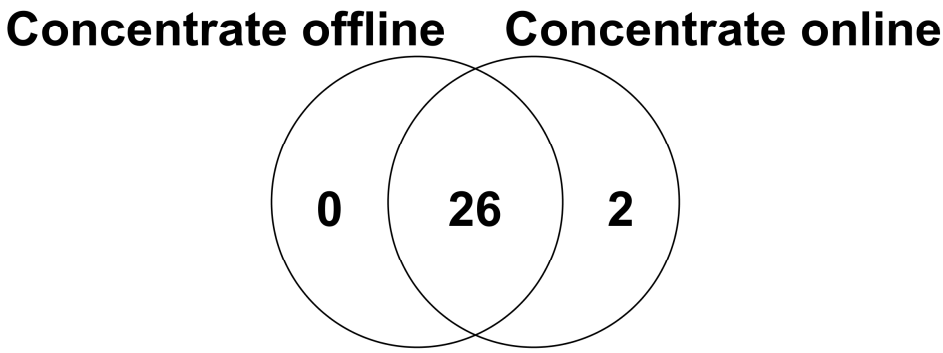

Figure 23: Venn diagram comparing online and offline results of the concentrate fraction of four technical replicates. The numbers illustrate the numbers of proteins identified with two or more peptides

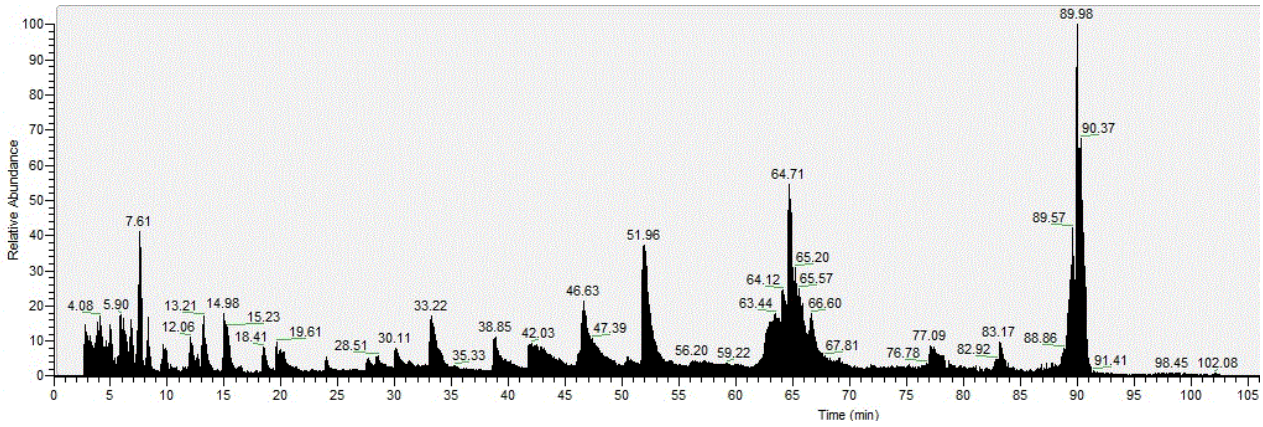

Figure 24: Total ion chromatogram of an online digest run of the concentrate fraction

In this instance, when using online digestion, the number of identified proteins was higher for both, concentrate and filtrate fraction. This is illustrated by the Venn diagrams in figure 22 and 23 which are based on four technical replicates. The data indicates that both methods could identify more or less the same proteins, with the exception of two proteins being identified with the online digestion only. The filtrate fraction, on the other hand, shows a different picture. Twice as many proteins were identified with the online method. Therefore we took a closer look at the single runs of the technical replicates, shown in figure 25.

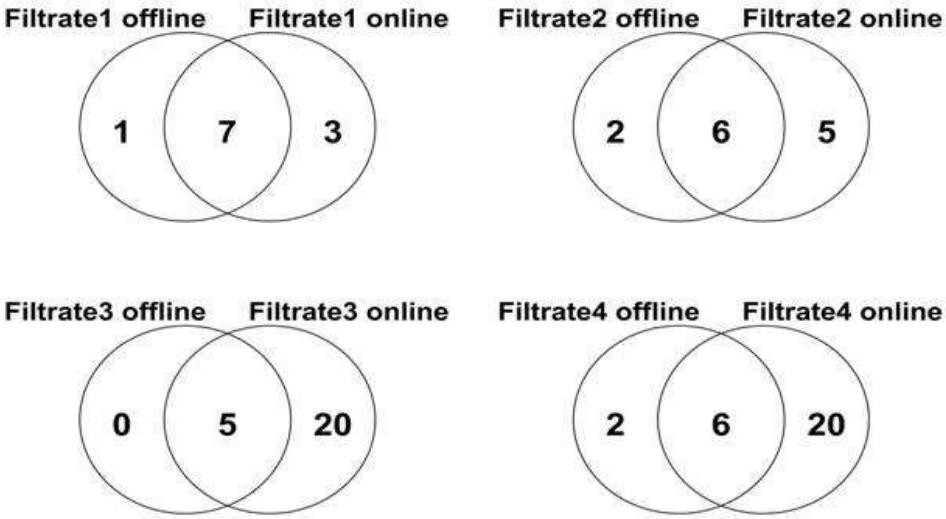

Figure 25: Venn diagrams showing the comparison of the single samples of the filtrate fraction.

While the number of proteins identified offline is constant, with online digestion the number shows a strong variation. While ten and eleven proteins were found in the first two samples, 25 and 26 respectively were identified in sample three and four. These discrepancies were not observed for the concentrate fraction. Excluding the possibility of variation within the sample, these findings are assumed to be caused by variations in the analytical method.

Considering the outcome of the concentrate fraction, the online digestion has shown to be on par with the overnight digestion approach. This of course has to be proven on a sound statistical basis, for which a great number of comparative runs are needed.

4 Curriculum Vitae

Philipp Obermayr BSc

Ausbildung

11/2012 - 11/2013	Masterarbeit am Klinischen Institut für Medizinische Labordiagnostik, Proteomics Core Facility, Medizinische Universität Wien
10/2011 – heute	Masterstudium Molekulare Biologie an der Universität Wien
07/2010	Erlangung des akademischen Grades „Bachelor of Science“
10/2004 – 07/2010	Studium der Mikrobiologie an der Universität Wien,
10/2001 – 06/2004	Studium der Pharmazie an der Universität Wien
06/2001	Matura am Bundesgymnasium BRG II

Berufserfahrung

02/2014 – heute	Fa. Chroma Pharma
08/2010 – 07/2012	Fa. AGES
08/2008 – 12/2009	Fa. Rösch & Handel

Zivildienst

10/2006 – 06/2007	Caritas Gruft, Obdachlosenbetreuung

5 Reference List

[1] Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D.*, et al.*, Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnology & genetic engineering reviews* 1996, *13*, 19-50.

[2] Finishing the euchromatic sequence of the human genome. *Nature* 2004, *431*, 931-945.

[3] Jensen, O. N., Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Current opinion in chemical biology* 2004, *8*, 33-41.

[4] Ayoubi, T. A., Van De Ven, W. J., Regulation of gene expression by alternative promoters. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 1996, *10*, 453-460.

[5] Reid, G. E., McLuckey, S. A., 'Top down' protein characterization via tandem mass spectrometry. *Journal of mass spectrometry : JMS* 2002, *37*, 663-675.

[6] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* 1989, *246*, 64-71.

[7] Armirotti, A., Damonte, G., Achievements and perspectives of top-down proteomics. *Proteomics* 2010, *10*, 3566-3576.

[8] Wu, C., Tran, J. C., Zamdborg, L., Durbin, K. R.*, et al.*, A protease for 'middle-down' proteomics. *Nature methods* 2012, *9*, 822-824.

[9] Kalli, A., Sweredoski, M. J., Hess, S., Data-dependent middle-down nano-liquid chromatography-electron capture dissociation-tandem mass spectrometry: an application for the analysis of unfractionated histones. *Analytical chemistry* 2013, *85*, 3501-3507.

[10] Mann, M., Hojrup, P., Roepstorff, P., Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biological mass spectrometry* 1993, *22*, 338-345.

[11] Henzel, W. J., Billeci, T. M., Stults, J. T., Wong, S. C.*, et al.*, Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proceedings of the National Academy of Sciences of the United States of America* 1993, *90*, 5011-5015.

[12] Perkins, D. N., Pappin, D. J., Creasy, D. M., Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999, *20*, 3551-3567.

[13] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B.*, et al.*, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics : MCP* 2002, *1*, 376-386.

[14] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F.*, et al.*, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature biotechnology* 1999, *17*, 994-999.

[15] Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B.*, et al.*, Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics : MCP* 2004, *3*, 1154-1169.

[16] Ishihama, Y., Oda, Y., Tabata, T., Sato, T.*, et al.*, Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular & cellular proteomics : MCP* 2005, *4*, 1265-1272.

[17] Compton, P. D., Zamdborg, L., Thomas, P. M., Kelleher, N. L., On the scalability and requirements of whole protein mass spectrometry. *Analytical chemistry* 2011, *83*, 6868-6874.

[18] Brownridge, P., Beynon, R. J., The importance of the digest: proteolysis and absolute quantification in proteomics. *Methods* 2011, *54*, 351-360.

[19] Vandermarliere, E., Mueller, M., Martens, L., Getting intimate with trypsin, the leading protease in proteomics. *Mass spectrometry reviews* 2013, *32*, 453-465.

[20] Burkhart, J. M., Schumbrutzki, C., Wortelkamp, S., Sickmann, A., Zahedi, R. P., Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics. *Journal of proteomics* 2012, *75*, 1454-1462.

[21] Polgar, L., The catalytic triad of serine peptidases. *Cellular and molecular life sciences : CMLS* 2005, *62*, 2161-2172.

[22] Proc, J. L., Kuzyk, M. A., Hardie, D. B., Yang, J.*, et al.*, A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin. *Journal of proteome research* 2010, *9*, 5422-5437.

[23] Ahn, J., Cao, M. J., Yu, Y. Q., Engen, J. R., Accessing the reproducibility and specificity of pepsin and other aspartic proteases. *Biochimica et biophysica acta* 2013, *1834*, 1222-1229.

[24] Switzar, L., Giera, M., Niessen, W. M., Protein digestion: an overview of the available techniques and recent developments. *Journal of proteome research* 2013, *12*, 1067-1077.

[25] Swaney, D. L., Wenger, C. D., Coon, J. J., Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *Journal of proteome research* 2010, *9*, 1323-1329.

[26] Biringer, R. G., Amato, H., Harrington, M. G., Fonteh, A. N.*, et al.*, Enhanced sequence coverage of proteins in human cerebrospinal fluid using multiple enzymatic digestion and linear ion trap LC-MS/MS. *Briefings in functional genomics & proteomics* 2006, *5*, 144-153.

[27] Gauci, S., Helbig, A. O., Slijper, M., Krijgsveld, J.*, et al.*, Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. *Analytical chemistry* 2009, *81*, 4493-4501.

[28] Chen, R., Jiang, X., Sun, D., Han, G.*, et al.*, Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *Journal of proteome research* 2009, *8*, 651-661.

[29] Li, A., Sowder, R. C., Henderson, L. E., Moore, S. P.*, et al.*, Chemical cleavage at aspartyl residues for protein identification. *Analytical chemistry* 2001, *73*, 5395-5402.

[30] Swatkoski, S., Gutierrez, P., Ginter, J., Petrov, A.*, et al.*, Integration of residue-specific acid cleavage into proteomic workflows. *Journal of proteome research* 2007, *6*, 4525-4527.

[31] Lo, L. H., Huang, T. L., Shiea, J., Acid hydrolysis followed by matrix-assisted laser desorption/ionization mass spectrometry for the rapid diagnosis of serum protein biomarkers in patients with major depression. *Rapid communications in mass spectrometry : RCM* 2009, *23*, 589-598.

[32] Patwa, T. H., Wang, Y., Simeone, D. M., Lubman, D. M., Enhanced detection of autoantibodies on protein microarrays using a modified protein digestion technique. *Journal of proteome research* 2008, *7*, 2553-2561.

[33] Park, H.-B., Pyo, S.-H., Hong, S.-S., Kim, J.-H., Optimization of the hydroxylamine cleavage of an expressed fusion protein to produce a recombinant antimicrobial peptide. *Biotechnology letters* 2001, *23*, 637-641.

[34] Permentier, H. P., Bruins, A. P., Electrochemical oxidation and cleavage of proteins with on-line mass spectrometric detection: development of an instrumental alternative to enzymatic protein digestion. *Journal of the American Society for Mass Spectrometry* 2004, *15*, 1707-1716.

[35] Permentier, H. P., Jurva, U., Barroso, B., Bruins, A. P., Electrochemical oxidation and cleavage of peptides analyzed with on-line mass spectrometric detection. *Rapid communications in mass spectrometry : RCM* 2003, *17*, 1585-1592.

[36] Park, Z. Y., Russell, D. H., Thermal denaturation: a useful technique in peptide mass mapping. *Analytical chemistry* 2000, *72*, 2667-2670.

[37] Havlis, J., Thomas, H., Sebela, M., Shevchenko, A., Fast-response proteomics by accelerated in-gel digestion of proteins. *Analytical chemistry* 2003, *75*, 1300-1306.

[38] Bark, S. J., Muster, N., Yates, J. R., 3rd, Siuzdak, G., High-temperature protein mass mapping using a thermophilic protease. *Journal of the American Chemical Society* 2001, *123*, 1774-1775.

[39] Lopez-Ferrer, D., Petritis, K., Hixson, K. K., Heibeck, T. H.*, et al.*, Application of pressurized solvents for ultrafast trypsin hydrolysis in proteomics: proteomics on the fly. *Journal of proteome research* 2008, *7*, 3276-3281.

[40] Yang, H. J., Hong, J., Lee, S., Shin, S., Kim, J., Pressure-assisted tryptic digestion using a syringe. *Rapid communications in mass spectrometry : RCM* 2010, *24*, 901-908.

[41] Dycka, F., Bobal, P., Mazanec, K., Bobalova, J., Rapid and efficient protein enzymatic digestion: an experimental comparison. *Electrophoresis* 2012, *33*, 288-295.

[42] Freije, J. R., Mulder, P. P., Werkman, W., Rieux, L.*, et al.*, Chemically modified, immobilized trypsin reactor with improved digestion efficiency. *Journal of proteome research* 2005, *4*, 1805-1813.

[43] Segu, Z. M., Hammad, L. A., Mechref, Y., Rapid and efficient glycoprotein identification through microwave-assisted enzymatic digestion. *Rapid communications in mass spectrometry : RCM* 2010, *24*, 3461-3468.

[44] Lopez-Ferrer, D., Capelo, J. L., Vazquez, J., Ultra fast trypsin digestion of proteins by high intensity focused ultrasound. *Journal of proteome research* 2005, *4*, 1569-1574.

[45] Wang, S., Zhang, L., Yang, P., Chen, G., Infrared-assisted tryptic proteolysis for peptide mapping. *Proteomics* 2008, *8*, 2579-2582.

[46] Wang, S., Liu, T., Zhang, L., Chen, G., Efficient chymotryptic proteolysis enhanced by infrared radiation for peptide mapping. *Journal of proteome research* 2008, *7*, 5049-5054.

[47] Schaefer, H., Chamrad, D. C., Marcus, K., Reidegeld, K. A.*, et al.*, Tryptic transpeptidation products observed in proteome analysis by liquid chromatography-tandem mass spectrometry. *Proteomics* 2005, *5*, 846-852.

[48] Massolini, G., Calleri, E., Immobilized trypsin systems coupled on-line to separation methods: recent developments and analytical applications. *Journal of separation science* 2005, *28*, 7-21.

[49] Xie, S., Svec, F., Frechet, J. M., Design of reactive porous polymer supports for high throughput bioreactors: poly(2-vinyl-4,4-dimethylazlactone-co-acrylamide- co-ethylene dimethacrylate) monoliths. *Biotechnology and bioengineering* 1999, *62*, 30-35.

[50] Ma, J., Zhang, L., Liang, Z., Zhang, W., Zhang, Y., Recent advances in immobilized enzymatic reactors and their applications in proteome analysis. *Analytica chimica acta* 2009, *632*, 1-8.

[51] Ma, J., Zhang, L., Liang, Z., Zhang, W., Zhang, Y., Monolith-based immobilized enzyme reactors: recent developments and applications for proteome analysis. *Journal of separation science* 2007, *30*, 3050-3059.

[52] Dulay, M. T., Baca, Q. J., Zare, R. N., Enhanced proteolytic activity of covalently bound enzymes in photopolymerized sol gel. *Analytical chemistry* 2005, *77*, 4604-4610.

[53] Licklider, L., Kuhr, W. G., Characterization of reaction dynamics in a trypsin-modified capillary microreactor. *Analytical chemistry* 1998, *70*, 1902-1908.

[54] Tang, Z. X., Qian, J. Q., Shi, L. E., Preparation of chitosan nanoparticles as carrier for immobilized enzyme. *Applied biochemistry and biotechnology* 2007, *136*, 77-96.

[55] Palm, A. K., Novotny, M. V., Analytical characterization of a facile porous polymer monolithic trypsin microreactor enabling peptide mass mapping using mass spectrometry. *Rapid communications in mass spectrometry : RCM* 2004, *18*, 1374-1382.

[56] Turkova, J., Blaha, K., Malanikova, M., Vancurova, D.*, et al.*, Methacrylate gels with epoxide groups as supports for immobilization of enzymes in pH range 3-12. *Biochimica et biophysica acta* 1978, *524*, 162-169.

[57] Nicoli, R., Rudaz, S., Stella, C., Veuthey, J. L., Trypsin immobilization on an ethylenediamine-based monolithic minidisk for rapid on-line peptide mass fingerprinting studies. *Journal of chromatography. A* 2009, *1216*, 2695-2699.

[58] Nicoli, R., Gaud, N., Stella, C., Rudaz, S., Veuthey, J. L., Trypsin immobilization on three monolithic disks for on-line protein digestion. *Journal of pharmaceutical and biomedical analysis* 2008, *48*, 398-407.

[59] Rios, G. M., Belleville, M. P., Paolucci, D., Sanchez, J., Progress in enzymatic membrane reactors – a review. *Journal of Membrane Science* 2004, *242*, 189-196.

[60] Samskog, J., Bylund, D., Jacobsson, S. P., Markides, K. E., Miniaturized on-line proteolysis-capillary liquid chromatography-mass spectrometry for peptide mapping of lactate dehydrogenase. *Journal of chromatography. A* 2003, *998*, 83-91.

[61] Vlakh, E. G., Tennikova, T. B., Flow-through immobilized enzyme reactors based on monoliths: II. Kinetics study and application. *Journal of separation science* 2013, *36*, 1149-1167.

[62] Sheldon, R. A., Enzyme Immobilization: The Quest for Optimum Performance. *Advanced Synthesis & Catalysis* 2007, *349*, 1289-1307.

[63] Sakai-Kato, K., Kato, M., Toyo'oka, T., Creation of an on-chip enzyme reactor by encapsulating trypsin in sol-gel on a plastic microchip. *Analytical chemistry* 2003, *75*, 388-393.

[64] Vlakh, E. G., Tennikova, T. B., Flow-through immobilized enzyme reactors based on monoliths: I. Preparation of heterogeneous biocatalysts. *Journal of separation science* 2013, *36*, 110-127.

[65] Celis, J. E., Gromov, P., 2D protein electrophoresis: can it be perfected? *Current opinion in biotechnology* 1999, *10*, 16-21.

[66] Anderson, N. L., Anderson, N. G., The human plasma proteome: history, character, and diagnostic prospects. *Molecular & cellular proteomics : MCP* 2002, *1*, 845-867.

[67] O'Farrell, P. H., High resolution two-dimensional electrophoresis of proteins. *The Journal of biological chemistry* 1975, *250*, 4007-4021.

[68] Braun, R. J., Kinkl, N., Beer, M., Ueffing, M., Two-dimensional electrophoresis of membrane proteins. *Analytical and bioanalytical chemistry* 2007, *389*, 1033-1045.

[69] Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., Aebersold, R., Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proceedings of the National Academy of Sciences of the United States of America* 2000, *97*, 9390-9395.

[70] Gao, M., Deng, C., Yu, W., Zhang, Y*., et al.*, Large scale depletion of the high-abundance proteins and analysis of middle- and low-abundance proteins in human liver proteome by multidimensional liquid chromatography. *Proteomics* 2008, *8*, 939-947.

[71] Alaoui-Jamali, M. A., Xu, Y. J., Proteomic technology for biomarker profiling in cancer: an update. *Journal of Zhejiang University. Science. B* 2006, *7*, 411-420.

[72] Issaq, H. J., The role of separation science in proteomics research. *Electrophoresis* 2001, *22*, 3629-3638.

[73] Boysen, R. I., Hearn, M. T., HPLC of peptides and proteins. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* 2001, *Chapter 8*, Unit8 7.

[74] Molnar, I., Horvath, C., Separation of amino acids and peptides on non-polar stationary phases by high-performance liquid chromatography. *Journal of chromatography* 1977, *142*, 623-640.

[75] Rafferty, J. L., Siepmann, J. I., Schure, M. R., Understanding the retention mechanism in reversed-phase liquid chromatography: insights from molecular simulation. *Advances in chromatography* 2010, *48*, 1-55.

[76] Mant, C. T., Hodges, R. S., Analysis of peptides by high-performance liquid chromatography. *Methods in enzymology* 1996, *271*, 3-50.

[77] Fekete, S., Veuthey, J. L., Guillarme, D., New trends in reversed-phase liquid chromatographic separations of therapeutic peptides and proteins: theory and applications. *Journal of pharmaceutical and biomedical analysis* 2012, *69*, 9-27.

[78] Guo, D. C., Mant, C. T., Hodges, R. S., Effects of ion-pairing reagents on the prediction of peptide retention in reversed-phase high-performance liquid chromatography. *Journal of chromatography* 1987, *386*, 205-222.

[79] Apffel, A., Fischer, S., Goldberg, G., Goodley, P. C., Kuhlmann, F. E., Enhanced sensitivity for peptide mapping with electrospray liquid chromatography-mass spectrometry in the presence of signal suppression due to trifluoroacetic acid-containing mobile phases. *Journal of chromatography. A* 1995, *712*, 177-190.

[80] Garcia, M. C., The effect of the mobile phase additives on sensitivity in the analysis of peptides and proteins by high-performance liquid chromatography-electrospray mass spectrometry. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* 2005, *825*, 111-123.

[81] Dolan, J. W., Snyder, L. R., Blanc, T., Selectivity differences for C18 and C8 reversed-phase columns as a function of temperature and gradient steepness. II. Minimizing column reproducibility problems. *Journal of chromatography. A* 2000, *897*, 51-63.

[82] Walter, T. H., Iraneta, P., Capparella, M., Mechanism of retention loss when C8 and C18 HPLC columns are used with highly aqueous mobile phases. *Journal of chromatography. A* 2005, *1075*, 177-183.

[83] Hamada, T., Tanaka, H., Izumine, H., Ohira, M., Evaluation of an embedded polar C4 phase for hydrophobic protein analysis by reversed-phase liquid chromatography. *Journal of chromatography. A* 2004, *1043*, 27-32.

[84] Kayillo, S., Dennis, G. R., Shalliker, R. A., An assessment of the retention behaviour of polycyclic aromatic hydrocarbons on reversed phase stationary phases: selectivity and retention on C18 and phenyl-type surfaces. *Journal of chromatography. A* 2006, *1126*, 283-297.

[85] Kaczmarski, K., Prus, W., Kowalska, T., Adsorption/partition model of liquid chromatography for chemically bonded stationary phases of the aliphatic cyano, reversed-phase C8 and reversed-phase C18 types. *Journal of chromatography. A* 2000, *869*, 57-64.

[86] Welinder, B. S., Sorensen, H. H., Hansen, B., Reversed-phase high-performance liquid chromatography of insulin. Resolution and recovery in relation to column geometry and buffer components. *Journal of chromatography* 1986, *361*, 357-367.

[87] Bakry, R., Huck, C. W., Bonn, G. K., Recent applications of organic monoliths in capillary liquid chromatographic separation of biomolecules. *Journal of chromatographic science* 2009, *47*, 418-431.

[88] Liu, K., Aggarwal, P., Lawson, J. S., Tolley, H. D., Lee, M. L., Organic monoliths for high-performance reversed-phase liquid chromatography. *Journal of separation science* 2013.

[89] Bisjak, C. P., Lubbad, S. H., Trojer, L., Bonn, G. K., Novel monolithic poly(phenyl acrylate-co-1,4-phenylene diacrylate) capillary columns for biopolymer chromatography. *Journal of chromatography. A* 2007, *1147*, 46-52.

[90] Ivanov, A. R., Zang, L., Karger, B. L., Low-attomole electrospray ionization MS and MS/MS analysis of protein tryptic digests using 20-microm-i.d. polystyrene-divinylbenzene monolithic capillary columns. *Analytical chemistry* 2003, *75*, 5306-5316.

[91] Gusev, I., Huang, X., Horvath, C., Capillary columns with in situ formed porous monolithic packing for micro high-performance liquid chromatography and capillary electrochromatography. *Journal of chromatography. A* 1999, *855*, 273-290.

[92] Premstaller, A., Oberacher, H., Walcher, W., Timperio, A. M.*, et al.*, High-performance liquid chromatography-electrospray ionization mass spectrometry using monolithic capillary columns for proteomic studies. *Analytical chemistry* 2001, *73*, 2390-2396.

[93] Eeltink, S., Dolman, S., Detobel, F., Swart, R.*, et al.*, High-efficiency liquid chromatography-mass spectrometry separations with 50 mm, 250 mm, and 1 m long polymer-based monolithic capillary columns for the characterization of complex proteolytic digests. *Journal of chromatography. A* 2010, *1217*, 6610-6615.

[94] Nischang, I., Teasdale, I., Bruggemann, O., Towards porous polymer monoliths for the efficient, retention-independent performance in the isocratic separation of small molecules by means of nano-liquid chromatography. *Journal of chromatography. A* 2010, *1217*, 7514-7522.

[95] Wu, R., Hu, L., Wang, F., Ye, M., Zou, H., Recent development of monolithic stationary phases with emphasis on microscale chromatographic separation. *Journal of chromatography. A* 2008, *1184*, 369-392.

[96] Fekete, S., Olah, E., Fekete, J., Fast liquid chromatography: the domination of core-shell and very fine particles. *Journal of chromatography. A* 2012, *1228*, 57-71.

[97] Horvath, C. G., Preiss, B. A., Lipsky, S. R., Fast liquid chromatography: an investigation of operating parameters and the separation of nucleotides on pellicular ion exchangers. *Analytical chemistry* 1967, *39*, 1422-1428.

[98] Guiochon, G., Gritti, F., Shell particles, trials, tribulations and triumphs. *Journal of chromatography. A* 2011, *1218*, 1915-1938.

[99] Chervet, J. P., Ursem, M., Salzmann, J. P., Instrumental requirements for nanoscale liquid chromatography. *Analytical chemistry* 1996, *68*, 1507-1512.

[100] Belov, M. E., Gorshkov, M. V., Udseth, H. R., Anderson, G. A., Smith, R. D., Zeptomole-sensitivity electrospray ionization--Fourier transform ion cyclotron resonance mass spectrometry of proteins. *Analytical chemistry* 2000, *72*, 2271-2279.

[101] Mitulovic, G., Smoluch, M., Chervet, J. P., Steinmacher, I.*, et al.*, An improved method for tracking and reducing the void volume in nano HPLC-MS with micro trapping columns. *Analytical and bioanalytical chemistry* 2003, *376*, 946-951.

[102] Schaefer, H., Chervet, J. P., Bunse, C., Joppich, C.*, et al.*, A peptide preconcentration approach for nano-high-performance liquid chromatography to diminish memory effects. *Proteomics* 2004, *4*, 2541-2544.

[103] Zenobi, R., Knochenmuss, R., Ion formation in MALDI mass spectrometry. *Mass spectrometry reviews* 1998, *17*, 337-366.

[104] Strupat, K., Karas, M., Hillenkamp, F., 2,5-Dihydroxybenzoic acid: a new matrix for laser desorption—ionization mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes* 1991, *111*, 89-102.

[105] Beavis, R. C., Chait, B. T., Matrix-assisted laser-desorption mass spectrometry using 355 nm radiation. *Rapid communications in mass spectrometry : RCM* 1989, *3*, 436-439.

[106] Beavis, R. C., Chait, B. T., Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins. *Rapid communications in mass spectrometry : RCM* 1989, *3*, 432-435.

[107] Beavis, R. C., Chaudhary, T., Chait, B. T., α-Cyano-4-hydroxycinnamic acid as a matrix for matrixassisted laser desorption mass spectromtry. *Organic Mass Spectrometry* 1992, *27*, 156-158.

[108] Armstrong, D. W., Zhang, L. K., He, L., Gross, M. L., Ionic liquids as matrixes for matrix-assisted laser desorption/ionization mass spectrometry. *Analytical chemistry* 2001, *73*, 3679-3686.

[109] Stump, M. J., Fleming, R. C., Gong, W.-H., Jaber, A. J.*, et al.*, MATRIX-ASSISTED LASER DESORPTION MASS SPECTROMETRY. *Applied Spectroscopy Reviews* 2002, *37*, 275-303.

[110] Knochenmuss, R., Ion formation mechanisms in UV-MALDI. *The Analyst* 2006, *131*, 966-986.

[111] Ehring, H., Karas, M., Hillenkamp, F., Role of photoionization and photochemistry in ionization processes of organic molecules and relevance for matrix-assisted laser desorption Ionization mass spectrometry. *Organic Mass Spectrometry* 1992, *27*, 472-480.

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

[113] Karas, M., Gluckmann, M., Schafer, J., Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *Journal of mass spectrometry : JMS* 2000, *35*, 1-12.

[114] Taylor, G. I., McEwan, A. D., The stability of a horizontal fluid interface in a vertical electric field. *Journal of Fluid Mechanics* 1965, *22*, 1-15.

[115] Fernández de la Mora, J., The Fluid Dynamics of Taylor Cones. *Annual Review of Fluid Mechanics* 2007, *39*, 217-243.

[116] Wu, X., Oleschuk, R. D., Cann, N. M., Characterization of microstructured fibre emitters: in pursuit of improved nano electrospray ionization performance. *The Analyst* 2012, *137*, 4150-4161.

[117] Covey, T. R., Thomson, B. A., Schneider, B. B., Atmospheric pressure ion sources. *Mass spectrometry reviews* 2009, *28*, 870-897.

[118] Mora, J. F., Van Berkel, G. J., Enke, C. G., Cole, R. B.*, et al.*, Electrochemical processes in electrospray ionization mass spectrometry. *Journal of mass spectrometry : JMS* 2000, *35*, 939-952.

[119] Last, I., Levy, Y., Jortner, J., Beyond the Rayleigh instability limit for multicharged finite systems: from fission to Coulomb explosion. *Proceedings of the National Academy of Sciences of the United States of America* 2002, *99*, 9107-9112.

[120] Fenn, J. B., Electrospray Wings for Molecular Elephants (Nobel Lecture). *Angewandte Chemie International Edition* 2003, *42*, 3871-3894.

[121] Kebarle, P., Verkerk, U. H., Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass spectrometry reviews* 2009, *28*, 898-917.

[122] Nguyen, S., Fenn, J. B., Gas-phase ions of solute species from charged droplets of solutions. *Proceedings of the National Academy of Sciences* 2007, *104*, 1111-1117.

[123] Lomeli, S. H., Yin, S., Ogorzalek Loo, R. R., Loo, J. A., Increasing charge while preserving noncovalent protein complexes for ESI-MS. *Journal of the American Society for Mass Spectrometry* 2009, *20*, 593-596.

[124] Chapeaurouge, A., Bigler, L., Schäfer, A., Bienz, S., Correlation of stereoselectivity and ion response in electrospray mass-spectrometry. Electrospray ionization-mass spectrometry as a tool to predict chemical behavior? *Journal of the American Society for Mass Spectrometry* 1995, *6*, 207-211.

[125] Ahadi, E., Konermann, L., Modeling the behavior of coarse-grained polymer chains in charged water droplets: implications for the mechanism of electrospray ionization. *The journal of physical chemistry. B* 2012, *116*, 104-112.

[126] Konermann, L., Rodriguez, A. D., Liu, J., On the formation of highly charged gaseous ions from unfolded proteins by electrospray ionization. *Analytical chemistry* 2012, *84*, 6798-6804.

[127] Wilm, M., Principles of Electrospray Ionization. *Molecular & Cellular Proteomics* 2011, *10*.

[128] Wilm, M. S., Mann, M., Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *International Journal of Mass Spectrometry and Ion Processes* 1994, *136*, 167-180.

[129] Schmidt, A., Karas, M., Dulcks, T., Effect of different solution flow rates on analyte ion signals in nano-ESI MS, or: when does ESI turn into nano-ESI? *Journal of the American Society for Mass Spectrometry* 2003, *14*, 492-500.

[130] Yost, R. A., Enke, C. G., Selected ion fragmentation with a tandem quadrupole mass spectrometer. *Journal of the American Chemical Society* 1978, *100*, 2274-2275.

[131] Picotti, P., Aebersold, R., Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nature methods* 2012, *9*, 555-566.

[132] Schwartz, J. C., Jardine, I., Quadrupole ion trap mass spectrometry. *Methods in enzymology* 1996, *270*, 552-586.

[133] Douglas, D. J., Frank, A. J., Mao, D., Linear ion traps in mass spectrometry. *Mass spectrometry reviews* 2005, *24*, 1-29.

[134] Guilhaus, M., Mlynski, V., Selby, D., Perfect Timing: Time-of-flight Mass Spectrometry†. *Rapid Communications in Mass Spectrometry* 1997, *11*, 951-962.

[135] Mamyrin, B. A., Time-of-flight mass spectrometry (concepts, achievements, and prospects). *International Journal of Mass Spectrometry* 2001, *206*, 251-266.

[136] Mamyrin, B. A., Karataev, V. I., Shmikk, D. V., Zagulin, V. A., The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *TrAC, Trends in Anal. Chem.* 1973, *37*, 45-48.

[137] Makarov, A., Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Analytical chemistry* 2000, *72*, 1156-1162.

[138] Hu, Q., Noll, R. J., Li, H., Makarov, A.*, et al.*, The Orbitrap: a new mass spectrometer. *Journal of mass spectrometry : JMS* 2005, *40*, 430-443.

[139] Makarov, A., Denisov, E., Kholomeev, A., Balschun, W.*, et al.*, Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Analytical chemistry* 2006, *78*, 2113-2120.

[140] Makarov, A., Denisov, E., Lange, O., Horning, S., Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer. *Journal of the American Society for Mass Spectrometry* 2006, *17*, 977-982.

[141] de Hoffmann, E., Tandem mass spectrometry: A primer. *Journal of Mass Spectrometry* 1996, *31*, 129-137.

[142] Chernushevich, I. V., Loboda, A. V., Thomson, B. A., An introduction to quadrupole–time-of-flight mass spectrometry. *Journal of Mass Spectrometry* 2001, *36*, 849-865.

[143] Glish, G. L., Burinsky, D. J., Hybrid mass spectrometers for tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry* 2008, *19*, 161-172.

[144] Vestal, M. L., Campbell, J. M., Tandem time-of-flight mass spectrometry. *Methods in enzymology* 2005, *402*, 79-108.

[145] Wells, J. M., McLuckey, S. A., Collision-induced dissociation (CID) of peptides and proteins. *Methods in enzymology* 2005, *402*, 148-185.

[146] Zubarev, R. A., Kelleher, N. L., McLafferty, F. W., Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *Journal of the American Chemical Society* 1998, *120*, 3265-3266.

[147] McLafferty, F. W., Horn, D. M., Breuker, K., Ge, Y.*, et al.*, Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance. *Journal of the American Society for Mass Spectrometry* 2001, *12*, 245-249.

[148] Mikesh, L. M., Ueberheide, B., Chi, A., Coon, J. J.*, et al.*, The utility of ETD mass spectrometry in proteomic analysis. *Biochimica et biophysica acta* 2006, *1764*, 1811-1822.

[149] Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* 2004, *101*, 9528-9533.

[150] Chi, A., Huttenhower, C., Geer, L. Y., Coon, J. J.*, et al.*, Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* 2007, *104*, 2193-2198.

[151] Bakhtiar, R., Guan, Z., Electron capture dissociation mass spectrometry in characterization of peptides and proteins. *Biotechnology letters* 2006, *28*, 1047-1059.

[152] La Marca, A., Sunkara, S. K., Individualization of controlled ovarian stimulation in IVF using ovarian reserve markers: from theory to practice. *Human reproduction update* 2014, *20*, 124-140.

[153] French, D. B., Sabanegh, E. S., Jr., Goldfarb, J., Desai, N., Does severe teratozoospermia affect blastocyst formation, live birth rate, and other clinical outcome parameters in ICSI cycles? *Fertility and sterility* 2010, *93*, 1097-1103.

[154] Mantikou, E., Youssef, M. A., van Wely, M., van der Veen, F.*, et al.*, Embryo culture media and IVF/ICSI success rates: a systematic review. *Human reproduction update* 2013, *19*, 210-220.

[155] Papanikolaou, E. G., Camus, M., Kolibianakis, E. M., Van Landuyt, L.*, et al.*, In vitro fertilization with single blastocyst-stage versus single cleavage-stage embryos. *The New England journal of medicine* 2006, *354*, 1139-1146.

[156] Rienzi, L., Ubaldi, F., Iacobelli, M., Romano, S.*, et al.*, Significance of morphological attributes of the early embryo. *Reproductive biomedicine online* 2005, *10*, 669-681.

[157] Meseguer, M., Rubio, I., Cruz, M., Basile, N.*, et al.*, Embryo incubation and selection in a time-lapse monitoring system improves pregnancy outcome compared with a standard incubator: a retrospective cohort study. *Fertility and sterility* 2012, *98*, 1481-1489 e1410.

[158] Fauser, B. C., Diedrich, K., Bouchard, P., Dominguez, F.*, et al.*, Contemporary genetic technologies and female reproduction. *Human reproduction update* 2011, *17*, 829-847.

[159] Cortezzi, S. S., Garcia, J. S., Ferreira, C. R., Braga, D. P.*, et al.*, Secretome of the preimplantation human embryo by bottom-up label-free proteomics. *Analytical and bioanalytical chemistry* 2011, *401*, 1331-1339.

[160] Katz-Jaffe, M. G., Gardner, D. K., Schoolcraft, W. B., Proteomic analysis of individual human embryos to identify novel biomarkers of development and viability. *Fertility and sterility* 2006, *85*, 101-107.

[161] Katz-Jaffe, M. G., McReynolds, S., Gardner, D. K., Schoolcraft, W. B., The role of proteomics in defining the human embryonic secretome. *Molecular human reproduction* 2009, *15*, 271-277.

[162] Katz-Jaffe, M. G., Schoolcraft, W. B., Gardner, D. K., Analysis of protein expression (secretome) by human and mouse preimplantation embryos. *Fertility and sterility* 2006, *86*, 678-685.

[163] Saborowski, R., Sahling, G., Navarette del Toro, M. A., Walter, I., García-Carreño, F. L., Stability and effects of organic solvents on endopeptidases from the gastric fluid of the marine crab Cancer pagurus. *Journal of Molecular Catalysis B: Enzymatic* 2004, *30*, 109-118.

[164] Spross, J., Sinz, A., A capillary monolithic trypsin reactor for efficient protein digestion in online and offline coupling to ESI and MALDI mass spectrometry. *Analytical chemistry* 2010, *82*, 1434-1443.

[165] Slysz, G. W., Schriemer, D. C., On-column digestion of proteins in aqueous-organic solvents. *Rapid communications in mass spectrometry : RCM* 2003, *17*, 1044-1050.

[166] Turapov, O. A., Mukamolova, G. V., Bottrill, A. R., Pangburn, M. K., Digestion of native proteins for proteomics using a thermocycler. *Analytical chemistry* 2008, *80*, 6093-6099.

[167] Sipos, T., Merkel, J. R., An effect of calcium ions on the activity, heat stability, and structure of trypsin. *Biochemistry* 1970, *9*, 2766-2775.

[168] Ponomareva, E. A., Kartuzova, V. E., Vlakh, E. G., Tennikova, T. B., Monolithic bioreactors: effect of chymotrypsin immobilization on its biocatalytic properties. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* 2010, *878*, 567-574.