DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES

DER FAKULTÄT FÜR CHEMIE UND PHARMAZIE

DER LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Novel quadrupole time-of-flight mass spectrometry for shotgun proteomics

von

Scarlet Svenja Anna-Maria Beck

aus Tettnang

2016

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Matthias Mann betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 25.04.2017

Scarlet Beck

Dissertation eingereicht am 23.09.2016

1. Gutachter: Prof. Dr. Matthias Mann
2. Gutachter: Prof. Dr. Jürgen Cox

Mündliche Prüfung am 04.11.2016

ABSTRACT

Mass spectrometry (MS)-based proteomics has become a powerful technology for the identification and quantification of thousands of proteins. However, the coverage of complete proteomes is still very challenging due to the high sample complexity and the difference in protein concentrations. In data-dependent shotgun proteomics several peptides elute simultaneously from the column and are isolated by the quadrupole and fragmented by the collision cell one at a time. This method has two major disadvantages. On the one hand, a large number of eluting peptides cannot be targeted since the sequencing speeds of current instruments are too slow and on the other hand, peptides that only differ slightly in mass and elute together are co-isolated and co-fragmented, resulting in chimeric MS2 spectra. Therefore an urgent need for further developments and improvements of mass spectrometers remains.

The aim of this thesis was to co-develop, evaluate and improve novel quadrupole time-of-flight (QTOF) mass spectrometers. In my first project I have described the developments and improvements of the hardware of the high-resolution QTOF mass spectrometer, the impact II, and have shown that this instrument can be used for very deep coverage of diverse proteomes as well as for accurate and reproducible quantification. With this kind of instrument, I achieved the deepest proteome coverage reported with a QTOF instrument so far. In my second project, a QTOF mass spectrometer was additionally equipped with a trapped ion mobility spectrometry (TIMS) device up-front. With this instrument configuration a novel scanning method was developed, called parallel accumulation – serial fragmentation (PASEF), that yields an up to 10-fold increase in sequencing speed and that has the potential to overcome the challenges of MS-based proteomics. In the third project, I have addressed the production of chimeric MS2 spectra. My results show that TIMS in combination with MS can already reduce the proportion of chimeric MS2 spectra up to two-fold and therefore appears very promising for eliminating ratio compression in isobaric mass tagging experiments after further improvements.

CONTENTS

ABSTRACT……..……..**v**

ABBREVIATIONS…….……...**ix**

1. INTRODUCTION……..**1**

1.1. Omics studies: from the genome to the proteome……1

1.2. MS-based proteomics……4

1.3. Ion Mobility Spectrometry (IMS) – Mass Spectrometry (MS)……30

1.4. Applications of MS-based proteomics……37

1.5. Objectives……...40

2. RESULTS……**43**

2.1. A novel QTOF mass spectrometer for shotgun proteomics analyses……..43

2.2. Increase in sequencing speed and sensitivity by using Parallel Accumulation – Serial Fragmentation on a TIMS-QTOF instrument……...……..61

2.3. Improving isobaric mass tagging quantification by trapped ion mobility mass spectrometry……..……...73

3. CONCLUSION AND OUTLOOK……..**90**

4. REFERENCES……..**xiii**

5. ACKNOWLEDGEMENT……...……..…**xxiii**

ABBREVIATIONS

ADH	alcohol dehydrogenase
AP-MS	affinity purification-mass spectrometry
CCS	collision cross section
CID	collision-induced dissociation
CV	coefficient of variation
DC	direct current
DDA	data-dependent acquisition
DIA	data-independent acquisition
DNA	deoxyribonucleic acid
DTIMS	drift-time ion mobility spectrometry
ESI	electrospray ionization
ETD	electron-transfer dissociation
FA	formic acid
FAIMS	field asymmetric waveform ion mobility spectrometry
FDR	false discovery rate
FT	Fourier transformation
FT-ICR	Fourier-transform ion cyclotron resonance
HCD	higher-energy collisional dissociation
HILIC	hydrophilic interaction liquid chromatography
HPLC	high performance liquid chromatography
IA-MS	immunoaffinity-based mass spectrometry
ICAT	isotope-coded affinity tags
IMAC	immobilized metal affinity chromatography

IMS	ion mobility spectrometry
iTRAQ	isobaric tag for relative and absolute quantification
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LFQ	label-free quantification
LTQ	linear trap quadrupole
LOD	limit of detection
m/z	mass-to-charge ratio
MALDI	matrix-assisted laser desorption/ionization
MOAC	metal oxide affinity chromatography
MRM	multiple reaction monitoring
MS	mass spectrometry
MS1	mass spectrum
MS/MS	tandem mass spectrometry
PASEF	parallel accumulation – serial fragmentation
ppm	parts-per-million
PRM	parallel reaction monitoring
PSM	peptide spectrum match
PTM	post-translational modification
Q	quadrupole
QQQ	triple quadrupole
QTOF	quadrupole time-of-flight
RF	radio-frequency
RNA	ribonucleic acid
SAX	strong anion exchange

SCX	strong cation exchange
SILAC	stable isotope labeling with amino acids in cell culture
SRM	selected reaction monitoring
SWATH-MS	sequential windowed acquisition of all theoretical mass spectra
TFA	trifluoroacetic acid
TMT	tandem mass tag
TOF	time-of-flight
TWIMS	travelling wave ion mobility spectrometry
UHPLC	ultra high performance liquid chromatography

1 INTRODUCTION

1.1. Omics studies: from the genome to the proteome

Proteins play a central role in a variety of biological functions such as catalyzing chemical reactions, performing structural roles in cells and mediating cell communication. Proteins consist of amino acids linked by peptide bonds. The folded states of proteins are complex, although they can be simplified by conceptually breaking them down into different layers of structure. The linear amino acid sequence defines the primary structure and this primary sequence then adopts secondary structure, defined by the patterns of hydrogen bonds formed between amine hydrogen and carbonyl oxygen atoms contained in the backbone peptide bonds of the protein. These local patterns are then folded into specific three-dimensional arrangements, the tertiary structure. The sequence of each protein is encoded by genes, which in eukaryotic cells is transferred in two steps. First, genes are transcribed to generate single stranded messenger RNAs (mRNAs) that leave the nucleus and secondly, they are translated by the ribosomal machinery to generate proteins.

For a system-wide view of cellular processes one needs to capture the information encoded by the biomolecules that are involved, including DNA, lipids and proteins. Historically, this began with the genetic instruction encoded in the genome. In 1977, with the introduction of sequencing technologies [1-3] such as electrophoretic sequencing, the first genome was sequenced [4]. Later, by a collaboration of different laboratories the first drafts of the human genome was published [5, 6], a breakthrough in the era of genomics. Information obtained by using sequencing technologies allows scientists to identify changes in genes that are associated with diseases and phenotypes. Following the development of next-generation sequencing technologies, large-scale acquisition of individual genomes can now be performed within a matter of days. But with the information gained on the genome level alone, the description of biochemical processes that establish life is not complete.

In the gene to protein hierarchy, the next step is to characterize all mRNA molecules, the transcriptome, and their abundances in the cell. This has been achieved on a large-scale with DNA microarray technology [7, 8]. This kind of analysis provides an overview of which parts of the genome are actually expressed at a given time point and gives an indication on which genes are under- or overexpressed or absent between different biological states. Transcriptomes are now more commonly obtained in a high-throughput format via next-generation sequencing, a technology called RNA-Seq when applied to RNA [9]. Transcriptomics adds the first dynamic

information layer to the systems-view of the cell. However, the correlation between mRNA levels and protein levels can be low. Therefore it is essential to directly measure the executers of biological functions – the proteins.

In the broadest sense, the proteomes of biological systems (e.g. cell lines, tissues or body fluids) describe the entirety of all expressed proteins, including all protein isoforms and modifications, and interactions between individual proteins at a given point in time and under specific condition. The proteome is a dynamic entity whereas the genome does not change and therefore more challenging to determine. In comparison to the powerful methods to measure nucleic acids, methods for the large-scale and systematic measurement of proteins took longer to develop. First, starting in the 1970s, researcher used two-dimensional gel electrophoresis (separation by molecular charge and by mass) to visually observe changes in the proteome [10]. However, this technique had severe limitations, such as the fact that only the most abundant proteins could be detected and that the identification of those spots was difficult and slow. This situation somewhat improved with the advent of the soft ionization techniques electrospray ionization (ESI) [11] and matrix-assisted laser desorption/ionization (MALDI) [12], which made it possible to identify protein spots one by one from a two-dimensional gel. However, the full potential of these technologies, which for the first time allowed the gentle transmission of peptides or proteins into the mass spectrometer only came to the fore when electrospray was online coupled to liquid chromatography (LC) and algorithms were developed to identify peptides by their fragmentation spectra in sequence databases (a technique termed shotgun proteomics) [13]. Many further advances in mass spectrometry and adjacent fields were need to enable the large-scale studies of peptides and proteins and to create the field of mass spectrometry (MS)-based proteomics [14].

There are around 20,000 protein coding genes annotated in the human genome [15], however the complexity of the proteome caused by alternative splicing of the mRNAs, endogenous proteolysis and post-translational modifications (PTMs) means that the number of protein species is much greater (Figure 1). MS-based proteomics has now become the tool of choice to identify and quantify the proteome of an organism [14]. With the current state of MS technologies, as well as in the entire proteomics workflow, from sample preparation to enhanced LC systems and data analysis software, it has become possible to achieve a deep coverage of the proteome in a reasonable time [16]. However, the complete analysis of all expressed proteins in a complex biological system with their differences in abundance, modification state and dynamics remains extremely challenging with existing technical instrumentation. The coverage of the full dynamic range of protein expression represents one of the major difficulties. For example, proteins in blood

plasma differ by more than 10 orders of magnitude in abundance [17], yet only 4-6 orders of magnitude can be covered by current LC-MS. In any case, due to the high complexity of a biological sample and the limitation in LC separation and acquisition speed of existing mass spectrometers, an ultra-deep coverage of the proteome would be very time-consuming with current instrumentation configurations, since extensive fractionation of a sample is required to decreases its complexity and provide sufficient measuring time. Therefore, an unmet need for continued improvements and breakthroughs in proteomic technologies remains. The main aim of the work in this thesis was to co-develop, improve and evaluate novel MS technology to make MS-based proteomics more competitive with the other omics technologies.

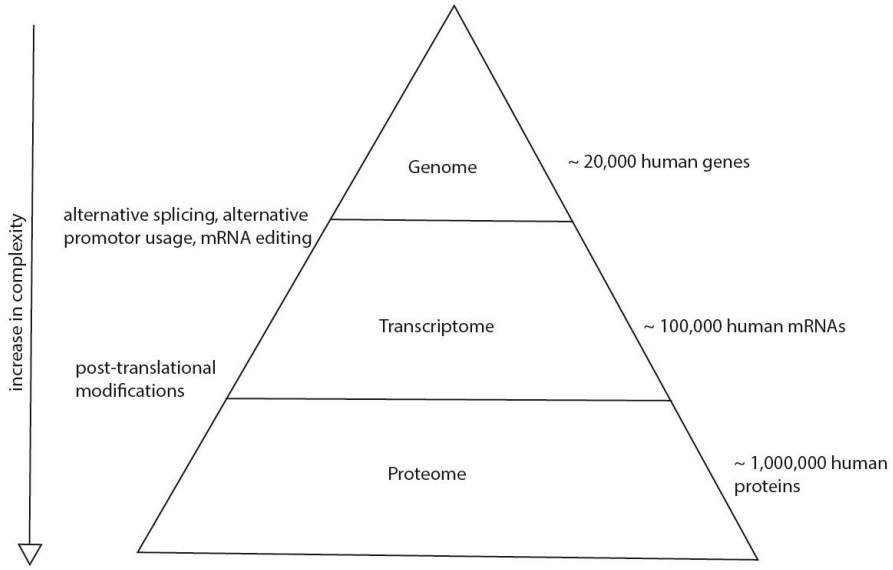


Figure 1: From the genome to the proteome. The size of the proteome increases by alternative splicing of transcripts and post-translational modifications of different proteins.

1.2. MS-based proteomics

In mass spectrometry, the analytes of interest need to be transferred into the gas-phase by an ion source. The ions can then enter the mass spectrometer where the mass-to-charge ratios (*m/z*) of ionized analytes are determined. MS has long been applied to many different areas in the screening of small molecules but could not analyze peptides and proteins due to the absence of techniques that would not destroy them during the ionization process. This changed with the introduction of ESI and laser desorption, for which the Nobel Prize in Chemistry was awarded in 2002. ESI became especially popular since it can easily be combined with high performance liquid chromatography (HPLC) by on-line coupling to MS. Thus, ESI is very suitable for the analysis of complex peptide mixtures, since peptides are already pre-separated before MS analysis. Different improvements and further developments in instrumentations, detailed below, have made mass spectrometry very attractive for proteomics, and have created the field of MS-based proteomics [14]. In addition to identifying and quantifying proteins, this technology can be applied to measure protein-protein interactions, post-translational modifications and even structural aspects of the proteome. It has become the method of choice for large-scale studies, due to its fast and high-throughput qualitative and quantitative analyses of the proteome of cells or organs in a hypothesis free and unbiased manner.

1.2.1. Sample preparation workflow for MS-based proteomics

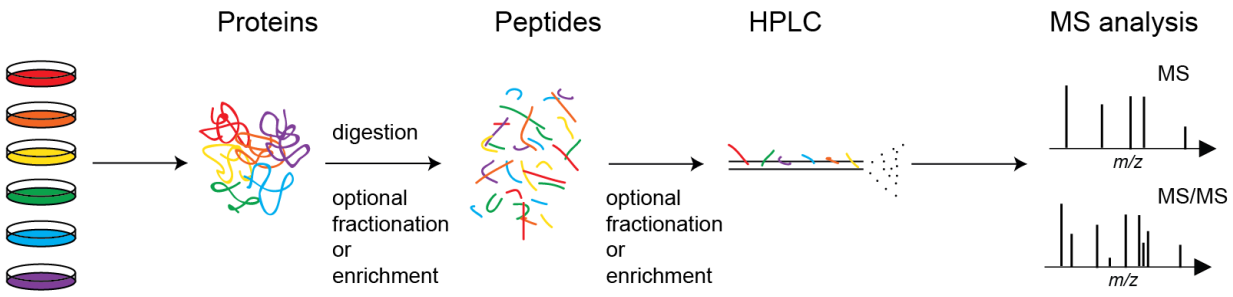


Figure 2: Overview of a typical shotgun proteomics workflow. Proteins derived from cells, tissues or body fluids are digested directly or enriched or affinity- purified before digestion. The resulting peptides are then separated chromatographically, which can be preceded by peptide fractionation or enrichment for specific peptides. Peptides eluting from the column are ionized by electrospray and analyzed in the mass spectrometer.

In the proteomic field there are two principal approaches: top-down and bottom-up analysis. In the top-down approach intact proteins and fragments produced from them are analyzed in the mass spectrometer [18-23]. With this method the whole protein sequence can in principle be covered and all different protein isoforms and modifications can be detected. However, the analysis of complex intact protein mixtures or very large proteins with LC-MS is extremely difficult because proteins are often not soluble, are difficult to separate chromatographically and have many more charge states. Therefore, the bottom-up approach, which relies on the analysis of peptides from a crude proteome digest, is much easier and by far the most-widely used method [14, 24-29] (Figure 2).

Having a robust and very reproducible workflow is absolutely necessary, especially for the comparison of different samples conditions. The standard sample preparation workflow in the bottom-up approach comprises three major steps: (1) cell- or tissue lysis together with protein denaturation, reduction of disulfide bonds and alkylation of cysteines, (2) enzymatic digestion(s) and (3) peptide clean-up. Depending on the biological question, a protein population of interest can be enriched by affinity enrichment techniques, e.g. protein extraction can be followed by immunoprecipitations. Moreover, different approaches have been developed for the enrichment of specific peptide classes, especially for low abundant peptides such as phosphopeptides. If very deep coverage of a proteome is of interest – both in terms of analyzing the proteins present and the coverage of the primary sequence by detected peptides, one can apply different fractionation techniques that have been developed and optimized over the last few years. After the sample preparation is completed, the peptide mixture is separated via LC which is on-line coupled to the MS analysis.

In the first step of a standard bottom-up MS-based proteomics workflow, proteins are extracted from a biological source such as cell lines, tissues or body fluids. In the case of cells or tissues this involves lysing the sample. Different protocols are available and which is applied is dependent on which protein condition, native or denatured, is desired. Disulfide bonds that stabilize the tertiary and quaternary protein structure are broken by a reducing agent (e.g. dithiothreitol, tris (2-carboxyethyl)phosphine, 2-mercaptoethanol)). The resulting free cysteine residues are irreversibly modified by alkylating reagents (e.g. iodo-acetamide or chloro-acetamide) before digestion of the protein mixture to avoid formation of new disulfide bonds.

Next, the protein mixture is digested to peptides with a sequence-specific enzyme. This is usually trypsin, which cleaves C-terminal to arginine or lysine, because of its robustness, low price, sequence specificity and the fact that the resulting peptides are generally easily detected and

sequenced in the mass spectrometer (mass range of 500 to 3,000 Da) [30]. One limitation of using trypsin is that more than 50% of the resulting peptides have less than six residues and therefore have little protein specific information [31]. To increase the proteome coverage, additional proteases such as LysC, GluC, LysN, AspN and chymotrypsin or protease combinations can be used [32-36]. Diverse protein digestion protocols have been developed to perform digestion either in-solution or in-gel. To remove remaining detergents that could interfere with high performance liquid chromatography (HPLC) separation or overwhelm the peptide signal in electrospray (ES), an additional clean-up step in the sample preparation workflow is generally applied in most protocols. Recently, our group has introduced the in-StageTip method, which was also used in this thesis, for performing sample preparation from cell lysis through digested and purified peptides in a single reaction vessel fashioned from a pipette tip [37].

The clean peptide mixtures are extremely complex and are therefore separated as a function of their hydrophobicity via reverse-phase, high performance liquid chromatography (HPLC) [38]. For proteomics, a chromatography column containing C_{18} material (reversed-phase C_{18} silica beads) is almost always used. Peptide mixtures are acidified by addition of amphiphilic acid (FA or TFA), which promotes their binding to the C_{18} material. They are eluted from the column by increasing the percentage of organic in the mobile phase, which is achieved by gradually changing the mixing ratio between aqueous and organic buffer solvent in the buffer. For the analysis of complex samples the highest resolution and peak capacity that can robustly achieved are advantageous for optimal proteome coverage. Therefore our laboratory employs long columns (around 50 cm) and small particle diameters (< 2 µm). Peptides that elute from the column are then ionized by electrospray ionization, which generally leaves them with at least two charges (on the basic N-terminus and terminal arginine or lysine residue) or more for larger peptides, before entering the mass spectrometer.

Very complex peptide mixtures greatly benefit from reducing the sample complexity by fractionation, resulting in more identifications or by enrichment using affinity purification (e.g. phosphopeptides enrichment) before LC-MS/MS analysis. Adding an extra chromatographic dimension of separation in addition to the reverse-phase separation results in an increase in peak capacity and a deeper proteome coverage. For this purpose different techniques for fractionation have been developed. Widely-used approaches include strong anion- or cation-exchange (SAX, SCX) columns [39, 40] as well as hydrophilic interaction liquid chromatography (HILIC) [41]. Recently, the combination of high-pH separation in the first dimension and the normal low-pH reversed-phase chromatography in the second dimension has become popular [42, 43]. In

comparison to SAX, SCX and HILIC, the high-pH reversed-phase method is only partially orthogonal, so it would be expected to be less beneficial. However, fractions can be pooled in a concatenated way (combining fractions that are sufficiently far apart in the first dimension), so that full orthogonality is achieved. If the focus is on a specific peptide population such as phosphopeptides, different enrichment methods can be applied, including immobilized metal affinity chromatography (IMAC) [44]. This technique is based on chelating agents that are immobilized on a polymer matrix. Added metal ions bind to chelating agents and also bind the negatively charged phosphopeptides in the sample. Alternatively, metal oxide affinity chromatography (MOAC) or titanium dioxide chromatography likewise have high affinity for phosphopeptides and can be used for the enrichment of phosphopeptides [45].

As the next step in the bottom-up MS-based proteomics workflow, the ions generated by ES enter the mass spectrometer where mass analysis takes place. For the analysis of complex peptide mixtures different mass spectrometer configurations and scan modes are available, which will be explained in more detail in the next sections.

1.2.2. Principles of mass spectrometric instrumentation used in proteomics

Gentle ionization methods

For the analysis of proteins or peptides in a mass spectrometer, it is necessary to convert them into gas-phase ions without destroying them during the ionization process. Over the years different ionization methods have been developed, but only two of them are suitable for the analysis of larger biopolymers, electrospray ionization (ESI) [11] and matrix-assisted laser desorption/ionization (MALDI) [12]. In the latter, the sample is embedded in a solid matrix and ions are produced by pulsed-laser irradiation and transferred to the mass analyzer. By its nature MALDI is therefore most easily coupled to time-of-flight (TOF) analyzers (see below). Despite different advantages of the MALDI method, such as the potential for very high-throughput, by far the most-widely used technique for the ionization of biomolecules in the proteomics field is ESI, which can be easily combined with liquid chromatography. ESI uses a small capillary at atmospheric pressure to which a high voltage is applied. From the needle, charged droplets containing solvent and analyte molecules are sprayed, which then evaporate further, ultimately generating multiple charged ions that enter the mass spectrometer.

Mass analyzers

For the analyses of (partially) LC separated and ionized peptides different mass analyzers can be used. They can be divided into two major groups based on their working principles: beam type analyzers that continuously scan ions (time-of-flight (TOF) and quadrupole (Q)) and trap-based analyzers that capture ions of interest for a specific time to acquire a mass spectrum (Fourier transformation ion cyclotron resonance (FT-ICR), ion trap and Orbitrap) (Figure 3). TOF instruments can be interfaced with a MALDI source for performing pulsed analysis. In contrast trapping mass analyzers are usually coupled to an ESI source. All instruments have unique properties and therefore vary in resolution, mass accuracy, scan speed, *m/z* range, dynamic range and sensitivity. All of them have undergone dramatic improvements with regards to these parameters over the last years.

The *resolution* (R) of a mass analyzer is calculated by the *m/z* value divided by the width of the peak at half height and gives an indication as to how well two different peaks with different *m/z* values can be distinguished. Mass analyzers with R > 10,000 are generally termed high-resolution mass spectrometers. Quadrupole and ion trap analyzers feature a low resolution (R < 1000), while TOF instruments can provide a resolution higher than 10,000. The highest resolution can be achieved with FT-ICR and Orbitraps (R > 100,000) as they measure frequencies of circulating ions with high accuracy. In proteomics, Orbitraps are currently the preferred instruments because they combine high resolution with less scanning time (scan speed) than FT-ICR instruments and because they are much smaller and more practical.

Mass accuracy is defined by how far the experimentally determined mass of an ion species differs from the actual (calculated) mass and is dependent on the resolution of an instrument. Moreover, both parameters are influenced by the signal-to-noise ratio. With an Orbitrap mass analyzer a mass accuracy of sub-parts-per-million (sub-ppm) can be achieved [46], while TOF instruments are usually in the range of several ppm.

TOF analyzers achieve a very high *scan speed* in the range of microseconds. But due to the limited number of ions collected in this time, multiple TOF spectra need to be combined, reducing the actual scan speed of the instrument to milliseconds (ms) [47]. High-resolution instrument such as FT-ICR suffer from the long transient duration and therefore are comparatively slow. In general there is an inverse correlation of scan speed and resolution. Orbitrap instruments currently need tens of milliseconds to achieve high resolution.

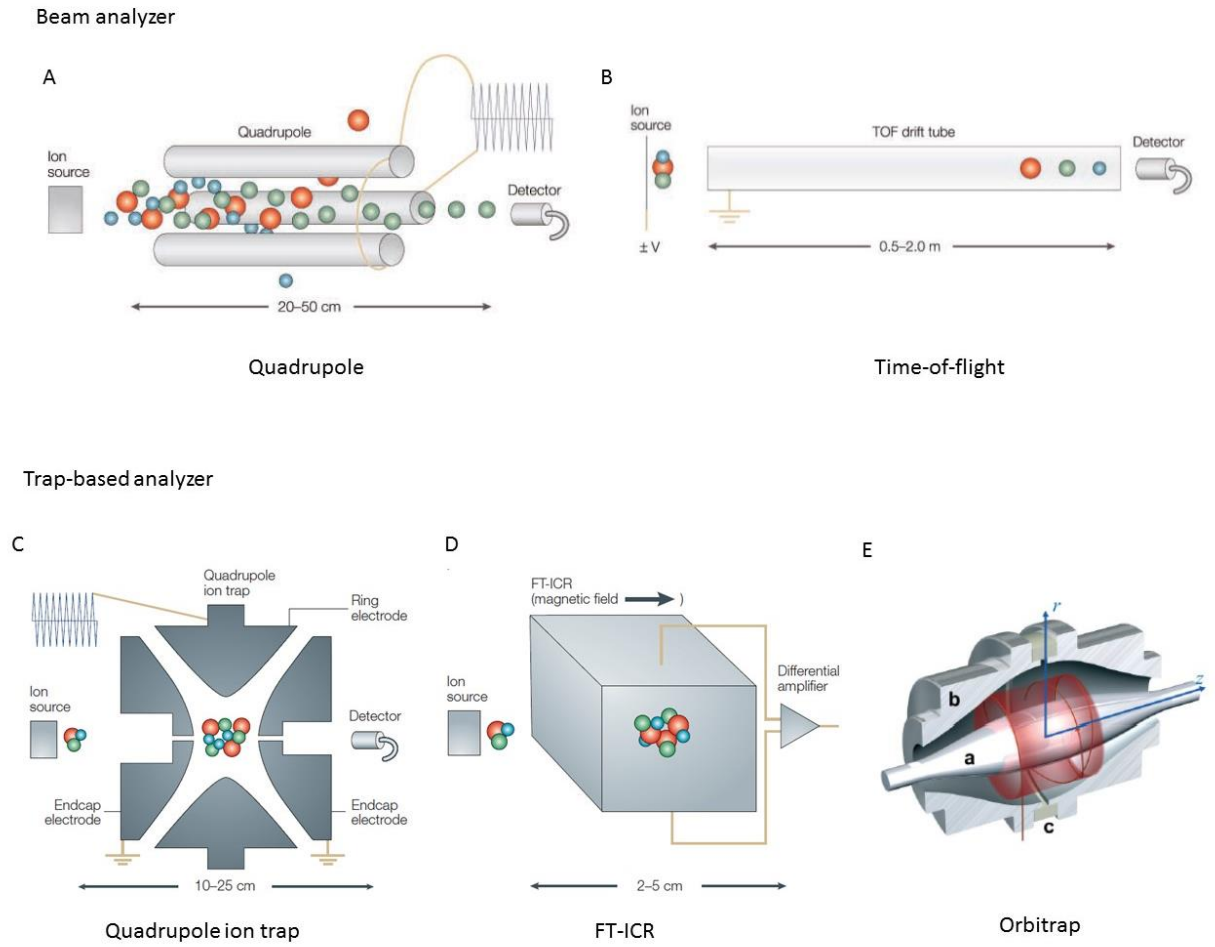


Figure 3: Different mass spectrometers used in MS-based proteomics. Dependent on the working principle they can be divided into beam type or trap-based analyzer. Adapted from [48, 49].

The *sensitivity* of a mass analyzer is dependent on its detection principle. There are currently two different methods available: Electron multiplier detectors, usually used by linear ion traps and TOF instruments and image current and Fourier transformation (FT)-based detectors, used by Orbitrap and FT-ICR analyzers. With the electron multiplier technology single ions can in principle be detected, while FT-based detection requires a few charges to achieve a detectable signal-to-noise ratio [50]. That said, improved electronics and thermal stability have enabled the detection of single ions with the Orbitrap mass analyzer, provided the ion is multiply charged and transient times are sufficiently long [51]. Currently, the limit of detection (LOD) of mass spectrometers is in the attomolar range, although this varies between simple and complex samples. Although TOF

mass analyzers in combination with sensitive detectors have the ability to detect a single ion, noise is also recorded with the same sensitivity and therefore the quality of the mass spectrum may be compromised. In practice, the current generation of Orbitrap instruments appear to have a better detection limit than TOF instruments [52].

Over the last decades, hybrid technologies, combining two or more mass analyzers in one platform have become popular in proteomics because this yields improved resolution, scan speed and sensitivity (Figure 4). Quadrupole mass analyzers with downstream TOF (QTOF) and Orbitrap instruments are widely used in the proteomics community and therefore these three mass analyzers are explained in more detail below. Dependent on the application a QTOF, triple quadrupole (QQQ) or the quadrupole Orbitrap may be the preferred hybrid instrument.

Quadrupole

The quadrupole mass analyzer is one of the oldest ones and consists of four parallel rods (Figure 3A). On each electrode pair a positive or negative direct current (DC) field is applied, which is superimposed by a time-dependent radio frequency (RF) field. A quadrupole technically works as a filter because ions entering the quadrupole are guided through the quadrupole longitudinal to the oscillating electric field caused by rapidly changing the RF field, which in combination with the DC field ensures that only ion species within a specific *m/z* range pass through the quadrupole. Switching off the DC field results in the transmission of the entire mass range. Quadrupoles feature low mass resolution but this can be improved somewhat by increasing the lengths of the rods, applying higher RF frequency, lower acceleration potential and slower scan speeds. A common configuration, especially in the targeted proteomics community, is the serial arrangement of three quadrupole analyzers. The first one is used for isolation, the next one for fragmentation and the last one for mass selection of fragments (Figure 4D).

The Orbitrap mass analyzer

The Orbitrap analyzer (Figure 3E) conceived and developed by Alexander Makarov was first described in 2000 and commercially introduced in 2005. This device consists of an outer barrel shaped and an inner spindle shaped electrode [53]. Ions entering the mass analyzer are trapped as they orbit around and along the central electrode. The oscillation along the field axis is specific for the *m/z* value of the ion and can be measured with high accuracy using image current detection on the segmented outer electrodes. The recorded image current is processed by Fourier

Transformation (FT), transforming the information into a mass spectrum. The principle of the Orbitrap is based on the Kingdon trap, which consisted of a wire and an enclosing electrode and was presented in 1923 [54], but only with the introduction of the Orbitrap shape and a C-trap [50] could ions be properly injected into the Orbitrap, making it possible to use it as a mass analyzer. Orbitrap mass analyzers feature high mass resolution and mass accuracy, which have been continuously improved during the last 10 years as well as reasonably high dynamic range [55-58]. Following the first hybrid Orbitrap mass spectrometer, the LTQ Orbitrap (Figure 4A) [50], Thermo Fisher introduced the Orbitrap Elite, which featured a dual linear ion trap and a higher resolution Orbitrap analyzer (Figure 4B) [57]. The quadrupole – Orbitrap combination was introduced with the Q Exactive [59] and the newest version of this configuration is called Q Exactive HF, featuring the high-resolution Orbitrap mass analyzer and an improved Fourier Transformation algorithm [58, 60]. Since hybrid Orbitrap mass spectrometers achieve a combination of high sequencing rate, together with very high resolution and mass accuracy, they have become the workhorse in the large majority of proteomics laboratories, especially for shotgun proteomics.

Time-of-flight

In time-of-flight mass analyzers, the *m/z* ratios of ions are determined by the time needed to pass a field-free drift tube (Figure 3B). Ions enter the flight tube with the same kinetic energy meaning that ions with a smaller *m/z* ratio pass the flight tube faster than those with larger ones.

The concept of TOF instruments was introduced by Stephens in 1946 [61]. Almost 10 years later the first design became commercially available [62]. Due to different developments such as progress in electronics and ionization methods towards the end of the 1980s, the TOF instrument became increasingly competitive. TOF mass analyzers have several advantages compared to other instruments including their high scanning rate, absence of space charge limitations and coverage of a large mass range [63]. The major disadvantage of TOF mass analyzers has been their lower resolution compared to Orbitrap and FT-ICR instruments, which can be increased by a longer flight tube or lowering the acceleration voltage. However, lowering the acceleration voltage has the undesired effect of decreasing sensitivity. Nevertheless, developments in the electronics, digitizer, and stability of power supplies together with the introduction of delayed pulsed extraction and reflectron technology have helped to increase the resolution so that it can now be sufficient for proteomics applications. The reflectron technology was first proposed by Mamyrin [64] and is integrated in the flight tube to reduce the spatial and kinetic energy spread of ions, leading to higher resolution by focusing ions with the same *m/z*. Reflectrons usually consist

of a series of equally spaced grid electrodes, which are connected through a resistive network of equal value resistors. They are positioned behind the collision region and are arranged orthogonal to the ion path. A further increase in performance was achieved by a two-stage reflectron, which uses two homogeneous electric fields of different potential gradients.

With the introduction of orthogonal acceleration time-of-flight technology by O`Halloran et al. in 1964 [65] and reinvented by different groups in the beginning of 1990s [66, 67], it became possible to combine a TOF instrument with the ESI source. The bottleneck of this combination is that the fraction of ions used for analysis is between 5% - 50% because ions are generated continuously, but pushed out every few microseconds. Nevertheless with this technology a resolution greater than 10,000 as well as improved mass accuracy was achieved. The addition of an RF multipole ion guide has yielded a higher ion transmission and sensitivity. A typical QTOF configuration is shown in Figure 4C.

Tandem mass spectrometry

Tandem mass spectrometry (MS/MS) involves two stages of MS. First, intact peptides are analyzed in a full scan (MS1) and then ions of desired *m/z* are selected (termed parent ions or precursor ions). Second, isolated ions are fragmented to measure the mass of the generated peptide fragments (also termed product ions) in the MS/MS (MS2) spectrum. Tandem mass spectrometry can be classified into the two categories: “tandem in space” and “tandem in time” [68]. In “tandem in time” mass spectrometers the analysis of parent ions and product ions is performed consecutively in the same mass analyzer. In contrast “tandem in space” instruments record the mass of parent ions and product ions in different mass analyzers.

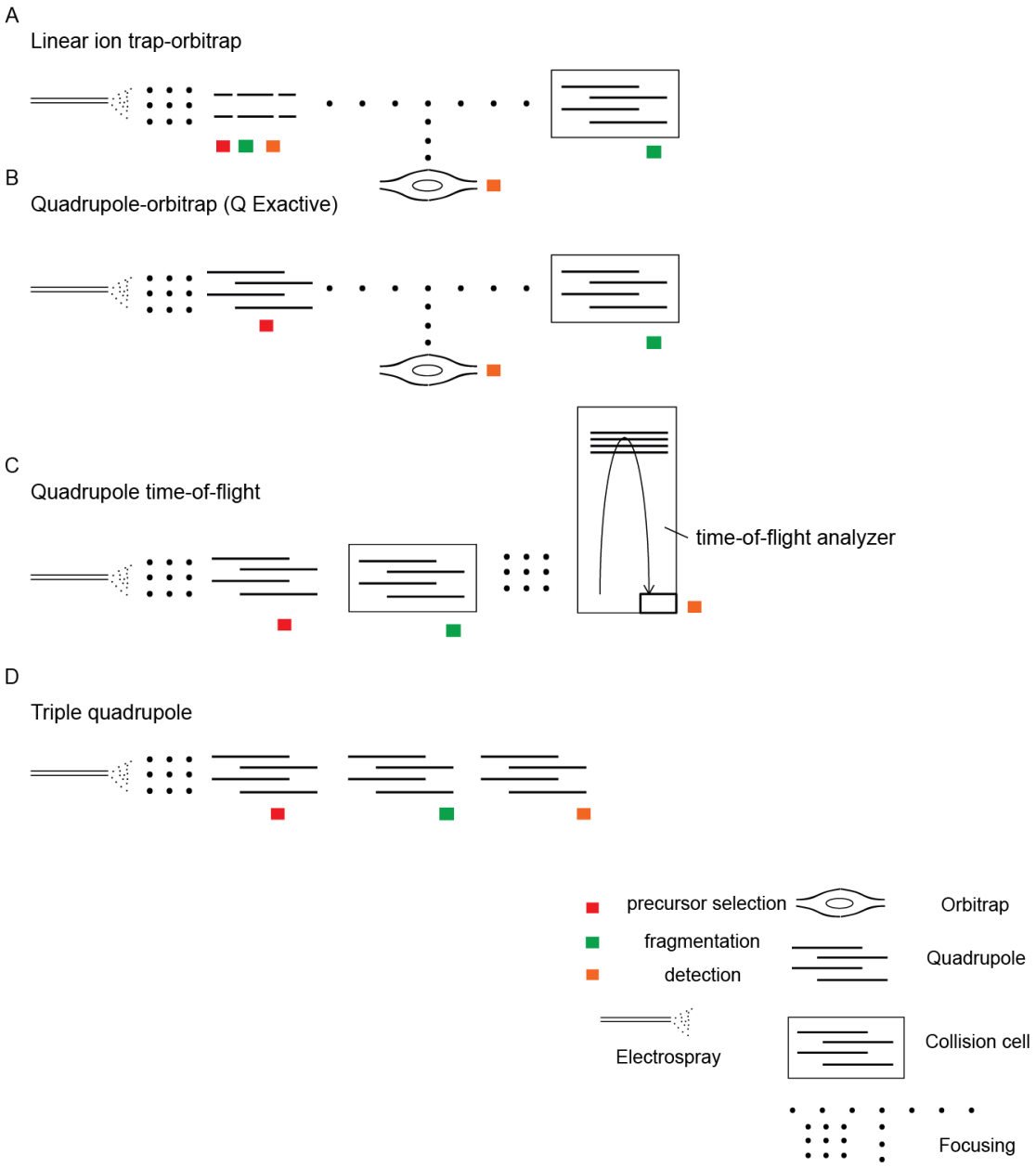


Figure 4: Overview of different mass spectrometer configurations used in MS-based proteomics. Colored squares indicate where precursor isolation (red), fragmentation (green) and detection (orange) is performed. Adapted from [69].

Fragmentation techniques

In tandem mass spectrometry product ions need to be generated by fragmentation to obtain sequence informative MS2 spectra. Several methods are used in MS-based proteomics, including collision-induced dissociation (CID), higher energy C-trap dissociation (HCD) and electron transfer dissociation (ETD) [70-73]. In CID as well as HCD collisions between peptides and inert gas molecules (e.g. He, Ar, N_2) result in bond breakages along the peptide backbone (Figure 5). Dependent on the terminus and cleavage site fragment ions are termed a-, b-, c- (when the N-terminus is intact) and x-, y-, z-ions (when the C-terminus is intact) (Figure 5) [74, 75]. When applying CID or HCD, predominantly b- and y-ions are generated. This leads to amino acid specific sequence information. HCD uses a higher energy for inducing peptide bond breakages in comparison to CID. The entire mass range of fragments is retained and transferred from the HCD cell to the mass analyzer (the Orbitrap in case of the Q Exactive). In contrast, when CID is performed in an ion trap, the low mass fragments fall outside the stability region and are lost (‘one third cut-off rule`) [76]. The low mass region, which is detectable with HCD, is especially important for reporter-based quantitation (see below) [77]. While CID and HCD is are very effective for the fragmentation of tryptic peptides and peptides with stable modifications, ETD may be preferable for intact proteins and longer peptides as well as for peptides carrying labile modifications like O-GlucNac [78]. In ETD, fragmentation is achieved by using an electron donor, such as anthracene or flouranthene anions, for the transfer of an electron to the analyte, resulting in an unstable transition state that instantaneously leads to fragmentation, mainly of the c- and z-type.

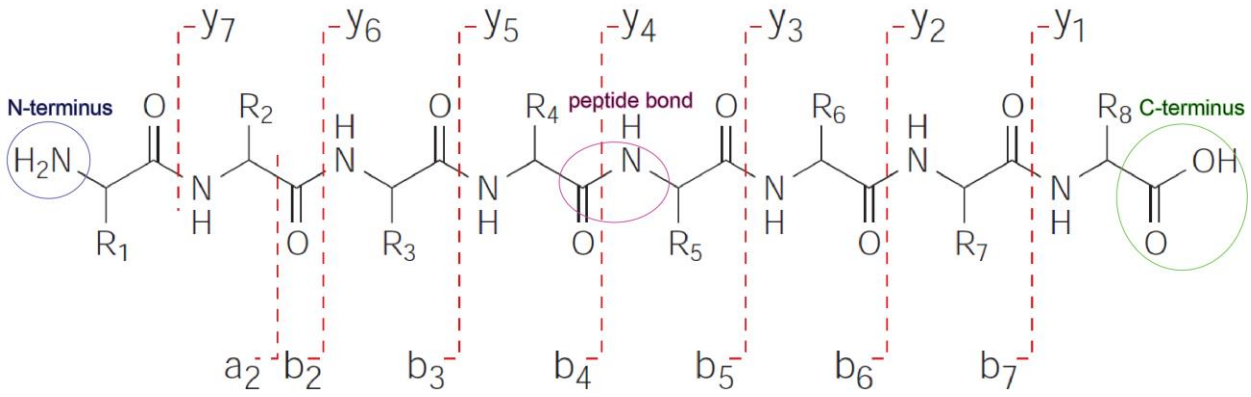


Figure 5: Peptide fragmentation. According to Roepstorff-Fohlmann-Biemann [74, 75], peptides generated depending on the remaining terminal residue a-, b- and c-ions (N-terminus) or x-, y-, and z-ions (C-terminus). Adapted from [79]

1.2.3. Data acquisition techniques in MS-based proteomics

In addition to the different instrument configurations for LC-MS/MS analysis, different MS acquisition modes and data query strategies (spectrum-centric or peptide-centric) have also been developed. There are three main operating modes of bottom-up proteomics: data-dependent acquisition (DDA, also known as shotgun proteomics), targeted data acquisition (performed mostly by selected reaction monitoring (SRM) or multiple reaction monitoring (MRM)) and data-independent acquisition (DIA). These strategies differ both in the way precursor and fragment data are recorded, and in how the acquired data are analyzed. The choice of which method to use depends on the biological question. For example, DDA is common in discovery proteomics, while SRM is preferably employed for very sensitive and accurate quantification of a small number of peptides in many samples.

Data-dependent acquisition

In DDA mode, the mass spectrometer scans the entire mass range (usually 300-1650 *m/z*) such that a mass spectrum (MS1) is obtained every few seconds (Figure 6). For identification, each peptide needs to be fragmented by tandem mass spectrometry (MS/MS, MS2). Generally, the 5-20 most abundant peptide peaks of each MS1 spectrum are successively isolated for fragmentation one at a time by a quadrupole or linear ion trap (topN cycle) (Figure 6). The decision of which precursor is selected for fragmentation is performed in real time by the software of the mass spectrometer. To avoid selecting the same peptide species multiple times during elution, a dynamic exclusion criterion is applied for a certain time interval. In addition singly charged ions can be excluded from the selection because they are usually contaminants and the tryptic peptides generally have at least two charges. The quadrupole isolation mass window is kept as small as possible to ensure that only the precursors of interest are selected for fragmentation. However, very small isolation windows limit transmission, so a common compromise is a selection window of about 1 to 2 *m/z* units (Th). Amide bonds are dissociated and an overlapping series of N-terminal (b-ions) and C-terminal (y-ions) fragments are generated. With the fragment information from the MS2 spectrum of one peptide, the peptide sequence can be determined by searching against a protein sequence database.

The shotgun or DDA approach is unbiased in that it fragments as many peptides as possible, which makes it suitable for discovery studies. It is typically performed on hybrid Orbitrap or quadrupole time-of-flight (QTOF) instruments (Figure 4A, B, C). Despite the relatively high effective scanning speed of both the QTOFs and Orbitraps, a challenge in DDA is that not all of

the eluting peptides may be selected for fragmentation, which is termed the undersampling or missing value problem.

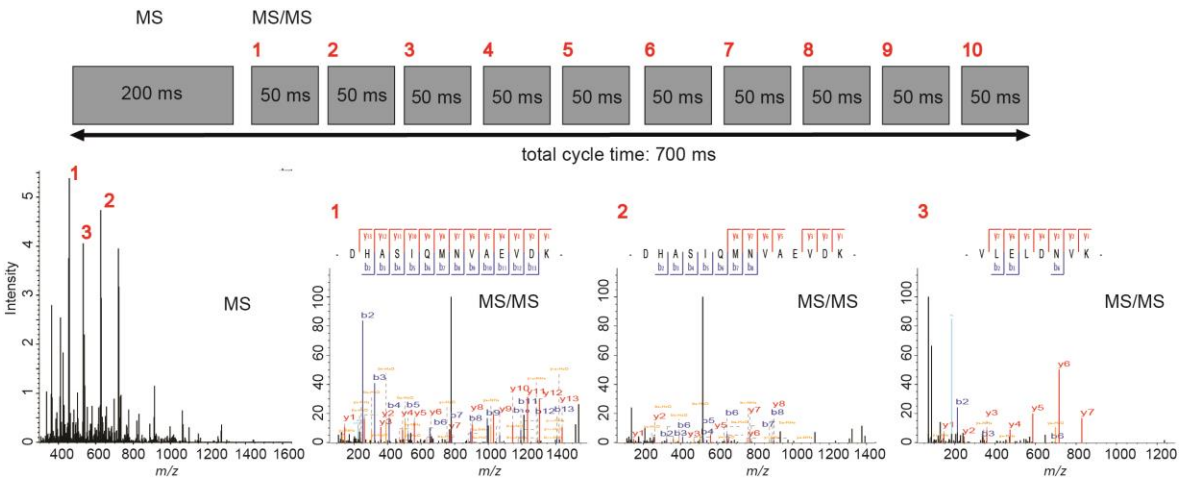


Figure 6: The data-dependent acquisition (DDA) mode. In DDA, the most abundant peptides from each full scan (MS1 scan) are selected one at a time by the quadrupole and fragmented in the collision cell. Information from the acquired MS2 spectra are used to match against a protein sequence database to identify the peptides present.

Targeted data acquisition

The selected reaction monitoring (SRM) or the closely related multiple reaction monitoring (MRM) approach are the most widely used methods in targeted proteomics. They determine the abundance of predefined peptides by repeatedly recording the transitions to fragment ions, which are also determined in advance, over the whole LC elution profile of the peptide. Fragment ion chromatographic signals of individual transitions are used to determine qualitative and quantitative information. This approach aims to ensure that peptides of interest and their fragments are measured in each sample at high sensitivity. Usually this method is carried out mostly on robust and sensitive triple quadrupole instruments (Figure 4D). Here, the first and third quadrupole act as mass filters, while the second one is used for fragmentation. However, due to the low resolution of the quadrupole, a main drawback of the method is its limited specificity which raises the issue of quantifying unrelated signal or noise instead of the intended target and requires extensive method optimization for each peptide. To improve specificity, parallel reaction

monitoring (PRM) or MS/MS^{All} has recently been introduced [80]. Conceptually, the third quadrupole of the triple quadrupole instruments is replaced with an Orbitrap mass analyzer, allowing the recording of high resolution MS2 signals with high accuracy over time. In the targeted approaches short gradients of about 30 min are commonly used and pre-fractionation is avoided for medium abundance peptides, which is important for high-throughput experiments. However, the development of peptide-specific assays is time consuming since the most promising peptides per protein and the most intense transitions need to be found. Moreover, all parameter settings need to be optimized for the instrument type used. Another disadvantage is that in practice generally less than 100 predefined peptides can be analyzed with MRM.

Data-independent acquisition

The aim of the DIA method is to continuously acquire MS/MS spectra to cover the complete tryptic peptide mass range. This is achieved by using a much wider quadrupole isolation mass windows for further fragmentation since the instrument scan speed is too slow to cover the complete mass range with a small mass window – and this would also severely limit sensitivity. Due to the fact that multiple precursors are selected and fragmented, the resulting MS2 spectra comprise fragments from many different precursor ions. The acquired data are therefore analyzed using specialized database searching or spectral library matching strategies in which fragment and precursor ions clustered based on their chromatographic peak shapes. DIA can be performed on different hybrid mass spectrometer such as QTOF and Q Exactive (Figure 4B, C). A popular DIA approach was introduced in 2012 by AB Sciex and the Aebersold group and called sequential windowed acquisition of all theoretical mass spectra (SWATH-MS) [81]. This method was demonstrated on a QTOF instrument and data were generated by repeatedly cycling through 32 consecutive 25 Da quadrupole mass isolation windows within the 400-1200 *m/z* mass range. Data analysis was performed by using only the information of the chromatographic peak shapes of precursor and their fragment ions signals.

1.2.4. Computational proteomics

In MS-based proteomics very large datasets are generated. For instance, using our set-up the analysis of a complex peptide mixture within a 90 min gradient run, typically generates more than 7000 MS1 scans and 75,000 MS2 spectra. For the analysis and interpretation of such large datasets, specialized computational workflows have been developed. In our laboratory, Jürgen

Cox has developed the freely available MaxQuant package software, which has since become a standard in the community [82, 83]. In this thesis the MaxQuant environment was used exclusively and therefore it will be described in more detail. The analysis of shotgun proteomics data can be divided into four steps: (1) feature detection and processing, (2) peptide identification, (3) protein identification and (4) quantification. Each of these will be described here and quantification will be described in-depth in the next section.

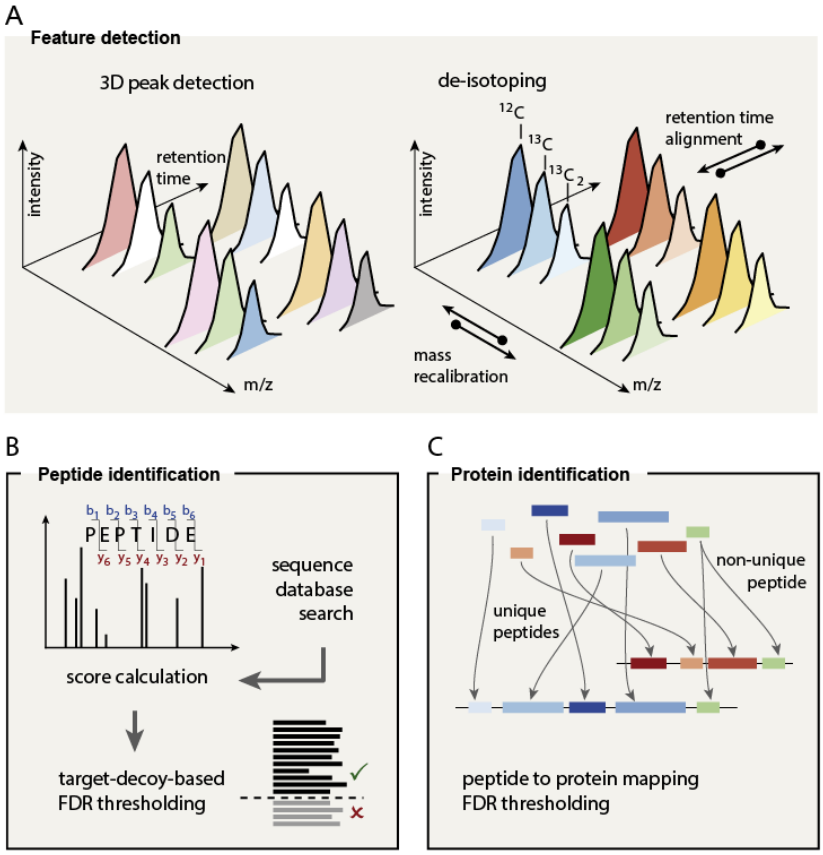


Figure 7: Computational proteomics: From feature detection to protein identification. (A) Peptide features in LC-MS/MS are plotted based on *m/z*, retention time and intensity. To achieve accurate measurements, recalibration of the mass and retention time can be performed. (B) Peptide identification based on database searching. (C) Assignment of peptides to their corresponding proteins. Adapted from [84].

Feature detection and processing: As a first step, peptide features are extracted from the raw data by detecting them in a three-dimensional area (Retention time, *m/z*, intensity) (Figure 7A)

and fitting them to peptide isotope patterns typical of that mass. To ensure very accurate measurements, MaxQuant applies algorithms to recalibrate the measured peptide masses on the basis of a subset of unambiguously identified peptides.

Peptide identification: Sequencing a peptide can either be accomplished *de novo* or by matching the MS/MS data to sequence databases. In the *de novo* approach the mass difference between each peak is calculated and assigned to a specific amino acid with that mass, ultimately leading a partial or total peptide sequence. This method can in principle even determine novel peptides and proteins that are not annotated or not present in the databases. However, *de novo* sequencing relies on complete fragment information, whereas database matching more easily deals with sparse information and is therefore more commonly used in shotgun proteomics. Precursor mass information from the full scans and MS2 spectra are used to compare them with theoretical spectra that are generated by an in-silico digest of the entire protein sequence repertoire of a specific organism (Figure 7B). The search engine is provided with digestion parameters such as protease specificity, number of allowed missed cleavages and fixed and variable (e.g. post-translational modifications) modifications. While it may seem advantageous to allow many such variations, this results in a combinatorial increase of the search space and negatively affects search specificity. Peptides are identified by scoring each obtained MS2 spectrum to a peptide in the database with a probability-based strategy. Peptides can be falsely assigned, most commonly caused by low quality MS2 spectra (few fragments and/or low ion intensity) or MS2 spectra containing co-fragmented precursors. To control for this, different techniques have been developed, including the target-decoy approach (reviewed in [85]). This method directly produces a false discovery rate (FDR) from the number of reverse peptide hits, which is generally set at 1%. Peptide identifications can be transferred by the so-called “match between runs” feature in MaxQuant, which transfers the identification of a sequenced peptide from one MS run to others where the peptide was only detected in the MS1 scan but not sequenced. This reduces the ‘missing value’ problem in shotgun proteomics.

Protein identification: After identifying peptides, they need to be assigned to the proteins that they originated from (Figure 7C). This ‘protein inference problem’ is non-trivial because not every peptide is unique for a specific protein but instead may occur in several proteins. In MaxQuant, proteins are grouped together into a ‘protein group’ when sharing one or multiple peptides together and no distinguishing peptides. Apart from the peptide FDR, a FDR on the protein level needs to be applied. The protein FDR works analogously to the peptide FDR: protein groups are ranked in order of the combined peptide spectrum match (PSM) scores and the ranked list of

protein groups is cut at the point where it contains 1% protein groups that match to the decoy database.

Protein quantification: After peptide and protein identification, the relative or absolute amount of each protein needs to be calculated and this is explained in the next section.

1.2.5. Quantitative proteomics

Overview of quantification techniques

In proteomics experiments, it is not only the presence or absence of thousands of proteins that is of interest, but even more their absolute and relative amounts, which is indispensable to understand their functional role in biological systems. In relative quantification the amount of proteins is compared between several samples, and in absolute quantification, the total amount of proteins in a sample is measured (for instance in femtomoles or in ng/mL in a body fluid). Over the years different approaches for relative and absolute quantitation have been developed and they can be categorized into label-free quantification and label-based quantification. Depending on the quantification approach, samples are combined at different stages of the sample preparation processing (Figure 8), resulting in differences in the accuracy and robustness of quantification. Moreover, depending on the strategy, the quantitative information is extracted either at the MS1 or MS2 (MS/MS) level (Figure 9).

Label-free quantification

In label-free quantification (LFQ), samples are not mixed and each sample is measured in separate MS runs (Figure 8). Due to the fact that each sample is treated separately until data analysis, artifactual quantitative differences can arise during sample processing and therefore the label-free strategy can be less accurate in comparison to the label-based techniques. Ensuring highly reproducible sample preparation and LC-MS measurements together with effective data analysis tools can nevertheless enable quite accurate quantification, which is also evident in this thesis. Peptide abundance can be calculated in different ways. The most common is the intensity-based LFQ approach. To enable accurate quantification, peptide peaks have to be distinguishable on the elution time and mass dimensions and therefore performance is best if data were acquired on high resolution mass spectrometers together with high resolution LC separation. The LFQ method is based on the assumption that the peak intensities of peptides correlate in a linear manner with the peptide concentration over a wide range of concentrations [86]. Label-free

strategies have existed in various forms since the beginning of electrospray ionization but have been refined over time (see for instance Bondarenko et al. [86] and Cox et al. [87]). In MaxQuant, the peptide features are detected in all MS measurements within a project, before aligning all runs on the basis of their retention times to make them comparable between each other (Figure 9). In this way even non-linear shifts in chromatography can be handled. In the next step the *matching between runs* feature is applied in MaxQuant, transferring peptide identifications from a run in which the peptide was identified to others in which the peptide was present put not sequenced [88]. This feature helps to overcome the limitation of shotgun proteomics where only the most abundant peptides are selected for fragmentation and can therefore be identified (the undersampling or missing value problem mentioned above). Moreover, this strategy increases the number of peptides that can be quantified. Importantly, the *matching between runs* feature can only be applied effectively if accurate mass measurement and corrected retention times are available. The raw intensities are then normalized on a global scale to avoid differences due to potential differences in overall sample amount, i.e. due to pipetting errors. This is done by assuming that the overall changes between all proteins must be zero. By considering all available pairwise peptide ratios, the protein intensity profile of each protein can be determined from the peptide intensities over all samples.

The main advantages of label-free quantification is that no extra sample handling is necessary for quantification and that therefore the technique can be applied to any sample. The absence of labeling also makes sample preparation very economical. Furthermore, in principle, the number of samples that can be analyzed and compared is unlimited.

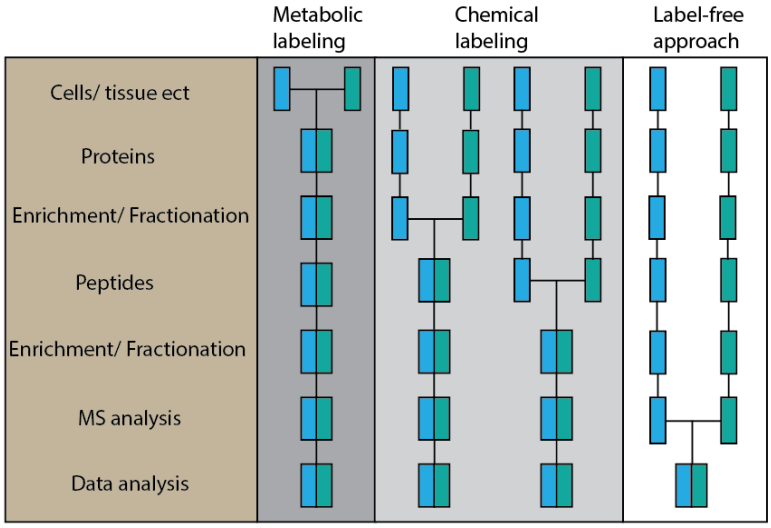


Figure 8: Approaches for relative quantification and their corresponding sample preparation processing strategies. The two differently colored boxes (blue and green) indicate two different samples to be compared. In metabolic labeling, samples are combined directly after obtaining them. All sample treatment steps are applied equally, resulting in very accurate quantification. In chemical labeling, samples are combined at a later stage of the workflow, potentially leading to less accurate quantitation. In the label-free approach samples are not combined for MS analysis at all which requires a very reproducible and robust sample preparation workflow. Whereas metabolic labeling can be multiplexed to some degree and chemical labeling can be highly multiplexed, label-free quantification requires separate and even repeated analyses per sample. Adapted from [89, 90].

Label-based quantification

In label-based quantification approaches, different stable heavy isotopes such as ^{13}C, ^{15}N, ^{2}H (D) are used. Usually all samples are heavy-labeled while using a light control. Introduction of the stable isotopes causes peptides to differ in their mass while their other physicochemical properties remain unchanged. After labeling, all samples can be combined and are treated equally during further sample processing and are analyzed together in the same LC-MS/MS runs. Depending on the labeling strategy samples are combined at early or later stage in the sample preparation workflow (Figure 8). The number of samples that can be combined and analyzed together, termed multiplexing, is dependent on the labeling technique. It typically ranges between 2 and 10 but analysis of 54 samples has also been shown by combining metabolic and chemical labeling [91],

potentially resulting in an enormous reduction of MS measurement time. Label-based quantification approaches can be divided into three main categories: metabolic labeling, chemical labeling and spike-in of labeled standards.

In *metabolic labeling*, the metabolism of living cells or organisms is used to introduce stable isotopes by feeding them with heavy-isotope-modified amino acids. In *stable isotope labeling with amino acids in cell culture (SILAC)* heavy-isotope substituted forms of arginine and/or lysine are used [92]. SILAC can also be employed with light, medium and heavy labeling, enabling direct comparison of three states. After labeling, samples are mixed at the earliest possible stage of the sample preparation workflow (Figure 8). As there are no artificial changes between samples during further sample preparation, highly accurate quantification results can be obtained. The samples are distinguishable on the MS1 level (Figure 9) because after digestion with trypsin peptides with at least a single labeled amino acid are generated – the C-terminal arginine or lysine. The mass spectrometer then detects each peptide in its light, medium (if applicable) and heavy version, which reflects the relative amount of each protein. Although considered the gold standard in quantitative proteomics, there are several drawbacks using this technique. Firstly not all sample types can be SILAC labeled, such as clinical samples and organisms that can readily synthesize their own lysine and arginine. Secondly, only two to three sample can be analyzed together. Thirdly, the combination of multiple samples can lower the number of identifications due to the increase in complexity introduced by the different isotopic peaks at the MS1 level. Additionally, as in any multiplexing technique there is a dilution effect as the total sample amount loaded onto the columns comes from three different samples and therefore the ion intensities are distributed between several isotopic peaks. Some of these challenges have been addressed by technologies such as super-SILAC [93] or neutron-encoded mass signatures for multiplexed proteome quantification (NeuCode) [94].

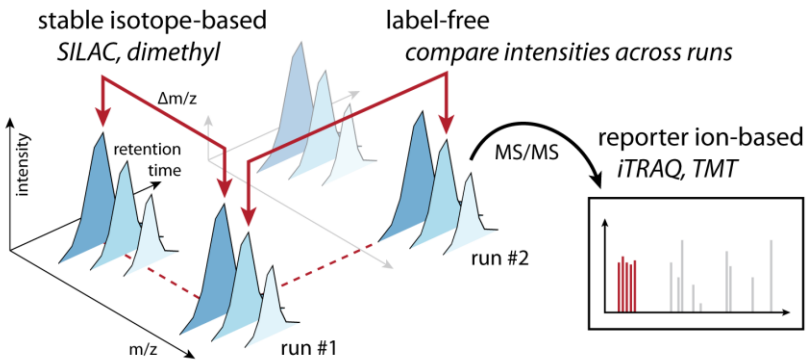


Figure 9: Extraction of information in different quantitative approaches. Depending on the approach, quantification is performed on the MS1 or MS2 (MS/MS) level. In the label-free approach peptide intensities are compared across runs at the MS1 level, and in stable isotope-labeling such as SILAC or dimethyl labeling peptide intensities are compared between the differentially labeled peptides within a run, also at the MS1 level. In contrast, in the reporter ion-based approaches quantification is performed on the reporter ion intensities at the MS2 level. Adapted from [84].

Chemical labeling

An alternative approach to metabolic labeling is *chemical labeling*, which can be applied to any sample. In chemical labeling approaches, the stable isotopes are introduced at a later stage of the sample preparation workflow in a chemical reaction (Figure 8). This can happen either at the protein or peptide level. Different chemical labeling strategies have been developed, including isotope-coded affinity tags (ICAT) [95], isobaric tag for relative and absolute quantification (iTRAQ) [96] and tandem mass tag (TMT) [90]. TMT is described in more detail since it plays a major role in this thesis. Isobaric tagging, including iTRAQ and TMT, utilize molecular tags that are not discriminable at the MS1 level because each of the labels confers the same overall mass addition. Fragmentation yields two types of product ions at the MS2 level: reporter ions specific to the label and peptide backbone fragment ions (Figure 9). Quantification is based on the relative intensities of the reporter ions. Advantages of chemical labeling are the degree of multiplexing, high precision in the reporter ion pattern and universal applicability. Disadvantages include expensive reagents, ensuring complete reaction without side reactions and the dilution of the sample between the combined samples.

Tandem mass tag (TMT)

Different TMT reagents are available such as TMTzero, TMTduplex, TMT 6-plex and TMT 10-plex. These share the same structure but differ in the number and positioning of ^{13}C and ^{15}N isotopes in the reporter region (Figure 10). In TMT 6-plex, the six tags generate their specific reporter ion at *m/z* 126, 127, 128, 129, 130 and 131 (Figure 10) and both QTOFs and Orbitraps can easily detect ions in the low *m/z* range with this unit mass spacing. In the higher multiplexed versions of TMT, the mass difference between the reporter ions is very small and mass spectrometers with very high resolution are needed. Currently only Orbitrap instruments can achieve the resolution required to resolve the 6 mDa differences necessary to measure all 10 channels of TMT 10-plex [97]. Advantages of isobaric quantification strategies at the MS2 level include that there is no dilution effect at the MS1 level due to multiple samples, no redundancy in MS/MS scanning events due to multiple precursor ion species [98] and it circumvents the C-trap capacity limit by not storing the entire mass range for quantification [99-101]. A major drawback of reporter ion-based quantification (e.g. iTRAQ and TMT), which has prevented more general adoption is the fact that quantification accuracy is severely limited by the so-called ratio compression problem [102-104], which will be described next.

Ratio compression: Accurate quantification of TMT-labeled peptides can only be achieved when a single precursor ion is selected for fragmentation. Co-eluting peptides within the same isolation window used for selection followed by fragmentation of specific ions results in under- or overestimation of the true peptide ratios in the measured samples (Figure 11). This is because co-fragmented peptides will all contribute to the same reporter ions, distorting the measured ratios. Another reason for inaccurate quantification is that artefactual spectral peaks can be present and interfere [105]. To overcome those challenges different strategies have been developed.

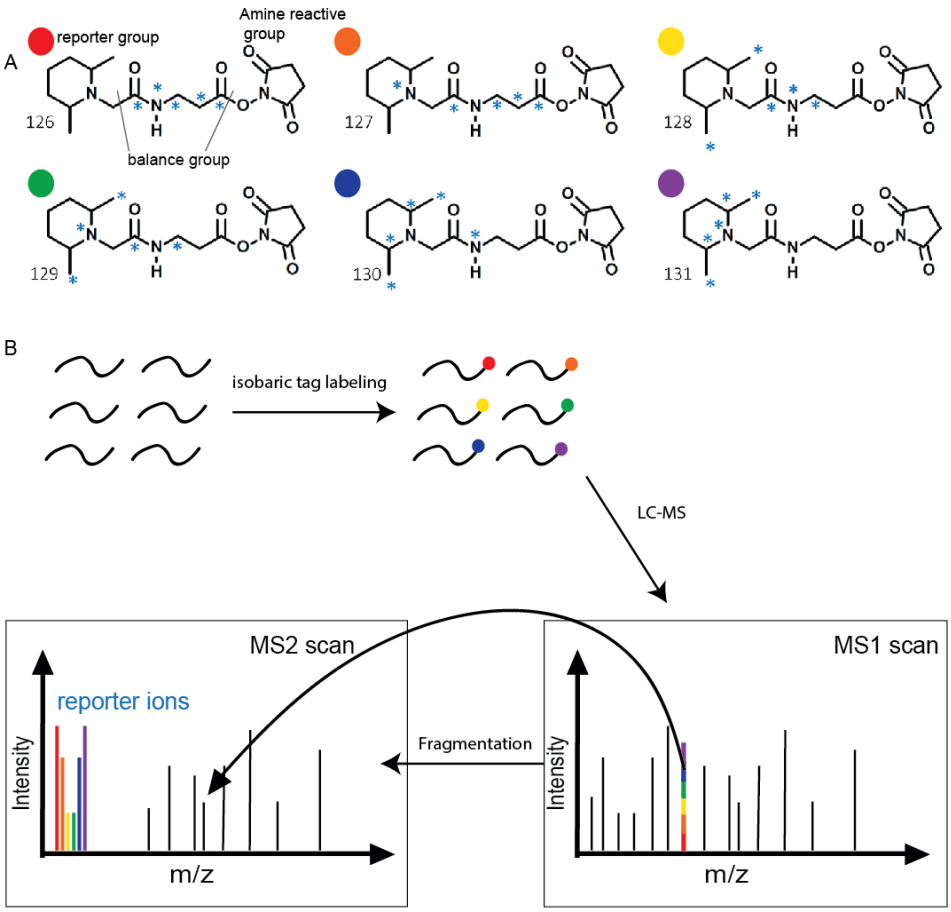


Figure 10: Tandem mass tag (TMT). (A) Chemical structure of the TMT 6-plex reagents. Blue asterisks indicate heavy isotope positions of ^{13}C and ^{15}N. Tags feature different distributions of isotopes between the reporter and balancer groups. (B) Workflow of a typical isobaric tagging experiment. Here, TMT 6-plex is used to label peptides from 6 different samples. Those are combined after labeling and analyzed together in the mass spectrometer. In the full MS scan (MS1) differentially labeled peptide ions are not distinguishable but after isolation and fragmentation of a specific peptide species quantification can be performed on the reporter ion intensities. Adapted from [106].

Overcoming ratio compression: Multiplexing strategies promise high-throughput, hence there is a great interest in solving the ratio compression issue. The higher the complexity of the sample, the more frequently co-elution plays a role, therefore ratio compression should be partly ameliorated by fractionation [107], however, in practice further fractionation actually worsens the problem perhaps because more low level peptide species contribute. Very narrow MS/MS isolation width settings also decrease interference from other ion species to some extent [103]. Moreover, the

selection of peptides followed by fragmentation at the apex of the LC peaks reduces co-fragmentation [103]. Although these methods can reduce the ratio compression problem, they do not solve it entirely. An approaches that more successfully reduces interfering ions in complex mixtures is the so-called MS3 method [108] and its refinement, the synchronous precursor selection (SPS) method [109]. In MS3, precursors are selected and fragmented like usual and then peptide fragments are selected and fragmented to generate MS3 spectra. The major disadvantage of applying MS3 is that the cycle times become longer than when performing MS2 only, sensitivity is reduced and currently, selection of fragmentation products from different, co-selected precursors can still lead to ratio compression. The SPS method can only be performed on linear ion trap - Orbitrap-based instruments and this combination is currently exclusively available on the very high end FUSION instrument.

It has been suggested that QTOF instruments equipped with ion mobility have the potential to eliminate interference of co-eluting peptides as IMS-MS separates ions based on shape, *m/z* and charge [110], and this is a major aim of this thesis.

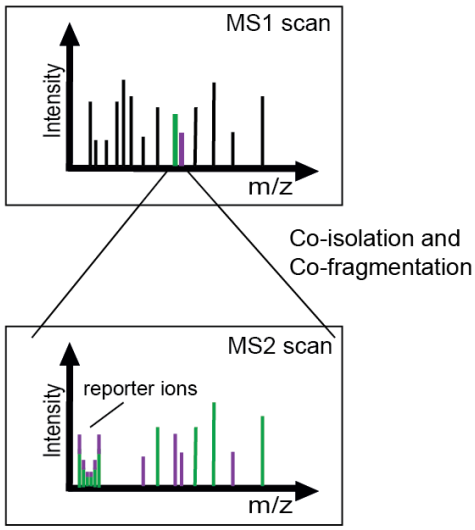


Figure 11: The ratio compression problem. Peptides that elute at the same time from the column and are selected by the quadrupole in the same isolation mass window are co-fragmented in the collision cell, resulting in chimeric MS/MS (MS2) spectra, with contributions of all of them. The reporter ion signals do not derive only from the targeted peptide species, leading to quantification errors.

1.2.6 Challenges in shotgun proteomics

Recent advances in proteomic technologies have made MS-based proteomics a central research tool. However, truly complete proteomes are still elusive, mainly due to the high complexity and dynamic range of biological samples with their range of post-translational modifications. More than 10,000 proteins are typically present in each biological sample at one specific condition and after digestion the complexity of analytes increases by at least one order of magnitude, because each protein generates tens to hundreds of peptides. The tremendous challenge of the dynamic range is best illustrated by plasma where protein concentration differs by more than 10 orders of magnitude, i.e. between albumin and low level cytokines [17], while LC-MS typically covers a dynamic range of 4-6 orders of magnitude [29]. Moreover in blood plasma the 22 most abundant proteins constitute about 99% of the total protein mass. Depletion of the high abundant proteins, fractionation techniques and enrichment approaches that target a specific sub-proteome (e.g. phosphopeptides) together with improvements in high resolution liquid phase separations and MS technologies have made it possible to cover the proteome in greater dynamic range and depth, but often introduce their own challenges. For instance, while extensive fractionation of a sample reduces complexity and results in a better proteome coverage [42], it also requires more measurement time, which reduces throughput. Besides reducing the complexity of the biological sample by fractionation or enrichment, high resolution in on-line LC separation helps to address these challenges. Very high pressure, long columns (50 cm) with small porous particles (<75 µg) and long gradients have improved the number of identifiable proteins even in single runs without any depletion or fractionation [111-113]. However, these developments still do not allow complete proteome characterization, and a premise of this thesis is that this could be achieved by a different approach – essential the introduction of third dimension of separation in a compact and efficient way (see below). In the following, the challenges of shotgun proteomics that could be addressed by such a novel approach are described.

The above mentioned ‘missing value’ or ‘undersampling’ problem in shotgun proteomics stems from the fact that a finite number of the most abundant precursor can be selected for fragmentation at any time and that this process is partially stochastic. An increase in sequencing speed may help to an extent but does not solve the issue completely since the complexity of the sample is too high and since the peptide separation via HPLC alone is insufficient so that multiple peptides at different concentrations co-elute from the column at the same time. Illustrating this, our group showed that more than 100,000 isotope features could be detected with our mass spectrometric configurations, but that only 16% of them were actually selected for fragmentation

and even less identified [114]. While an increase in sequencing speed or an increase in ion current would help to successfully target more peptide features, co-eluting peptides with a similar *m/z* or a small mass difference would still be isolated together by the quadrupole and fragmented in the collision cell, resulting in chimeric MS2 spectra (Figure 11). This decreases the potential to identify the peptide with high confidence. Already now only about 50% of all acquired MS2 spectra can be identified. This is partly a result of chimeric MS2 spectra but also due to low quality MS2 spectra. Although several search engines can now handle chimeric MS2 spectra and can identify multiple peptides from one MS2 spectrum, this strategy loses effectiveness when the intended peptide is of low abundance compared to the co-fragmented one. Small mass isolation windows (currently down to 0.4 Th) also help to some extent but there is a limit to the size of the window because making it too small would lead to less ion transmission efficiency and hence reduced sensitivity. In summary, co-elution and co-fragmentation of several peptides result in a decrease in peptide and protein identification (in addition to the problem of ratio compression explained above).

Another challenge in MS-based proteomics is the identification of peptide variants and post-translational modifications. This is currently either not part of the standard workflow at all or achieved by PTM-specific workflows. However, improvements in MS-based proteomics technology could ideally allow detection of such variants in a standard and generic way.

For all the reasons given in this section, further improvements on the existing technologies are absolutely necessary. In particular, we were searching for a way to increase fragmentation speed without losing sensitivity as well as a way to reduce the extent of chimeric MS2 spectra and ratio compression. For this purpose I have co-developed, improved and evaluated a novel mass spectrometer for quantitative proteomics.

1.3. Ion Mobility Spectrometry (IMS) – Mass Spectrometry (MS)

1.3.1 General principles

Ion mobility spectrometry (IMS) separates ions in the gaseous phase based on their size and shape using an electric field. Classical studies of the famous physicist Paul Langevin are a basis for this separation technique [115, 116]. Over the past several decades, IMS instrumentation has been applied to very diverse areas such as chemical weapons monitoring [117], detection of dangerous and illegal substances [118], food quality analysis [119], drug detection [120] and to a limited degree, biological analysis [121]. There are different ion mobility principles, which can be categorized into dispersive ones that capture the entire ion mobility range and selective ones that pass a particular ion mobility range. The dispersive techniques include drift-time ion mobility spectrometry (DTIMS), travelling wave ion mobility spectrometry (TWIMS) and trapped ion mobility spectrometry (TIMS) while the selective techniques are represented by field asymmetric waveform ion mobility spectrometry (FAIMS). Ion mobility separation typically happens in milliseconds, which readily fits between the LC separation time (seconds to minutes) and the mass spectrometer scan time (operating on the microsecond time scale). In combination with MS it can be used to determine structural information, gain insights into the conformational dynamics of a protein or protein complex and resolve isomers of the same chemical compound that are not distinguishable by mass spectrometric measurement alone [122]. For proteomics applications, the potential increase in peak capacity, increased dynamic range and improved signal-to-noise ratio are most attractive [123].

Hybrid IMS-MS instruments were first described over 50 years ago [124-126]. Due to improvements on the MS as well as the IMS side, IMS-MS has shown great potential for the analysis of complex mixtures in proteomics [127-131], glycomics [132-134], metabolomics [135-137] and petroleomics [138, 139]. Therefore four different IMS techniques that can be combined with MS are described in more detail in the following.

In IMS ions are separated based on their structure and shape, which together determine the collision cross section (CCS). In more detail, the cross section of an ion Ω is related to the average shape of the ion and is determined by the collision rate with a buffer gas. Due to fewer collisions between compact ions and the buffer gas in comparison to elongated ions (e.g. planar structures, helices, etc.), compact species have higher ion mobilities than the ions with a more open structure.

The mobility of an ion is determined by [140]:

K = v/E

Where K is the mobility of an ion, v the speed an ion moves and E the electric field to which it is subjected.

The mobility is dependent on the experimental temperature and pressure. Therefore a reduced mobility K_0 (normalized to standard temperature and pressure) is used. According to the Mason-Schamp equation [141], Ω can be calculated for a classical linear drift tube as follows:

$$\Omega = \frac{3ze}{16N}\left(\frac{2\pi}{k_b T}\right)^{\frac{1}{2}}\frac{1}{K_0}$$

where z is the charge state of the ion, e is the elementary charge, N is the number density of the drift gas, k_b is the Boltzmann constant, and T is the gas temperature. The low field limit is the range in which this relationship between Ω and K_0 is valid. The determined CCS provides characteristic information of the structure of each ion. This data can be used to compare it to other structural techniques such as nuclear magnetic resonance spectroscopy. CCSs can be precisely measured, usually to a few percent [142-144], Note, however, that CCS values need not be constant even within the low field limit as they will depend on the gas, and as ions may start to unfold and change structure as they are heated.

1.3.2. Drift-time Ion Mobility Spectrometry – Mass Spectrometry

DTIMS is the oldest IMS technique and is implemented as a series of stacked-ring electrodes to which an electric field is applied in the direction of the drift tube axis. Ions in a carrier gas enter the drift tube, which is filled with a buffer gas - usually helium (but nitrogen or argon can also be used), and are guided through the drift tube by applying a static and uniform electric field (usually 5-100 V) (Figure 12A). The time needed for an ion to pass through the tube depends linearly on its CCS. The measured drift time of each ion type is then used for the calculation of the CCS according to the Mason-Schamp equation given above [145].

One issue of early developed DTIMS instruments was their low duty cycle (percentage of ions that could be analyzed compared to total ions produced). This limitation has been partially addressed by different improvements such as ion funnel traps [146-150] and using higher

pressure. Additionally, applying a radio frequency for ion confinement has led to better transmission of ions [149, 151-153].

Clemmer and coworkers developed the first dispersive IMS-MS instrument [154]. Nowadays, different DTIMS instruments are commercially available from TOFWERK, Excellims and Agilent, for example, and they usually address low molecular weight applications. DTIMS-MS instruments consist of a short drift tube into which ions are injected and where – after separation - they are detected by a compact mass spectrometer. In case of a TOF instrument, the time the ions need to pass from the entrance of the drift tube to the TOF pulser determines the drift time. Ideally all ions are then guided into the TOF. With DTIMS-MS a resolving power up to around 250 [122] can be achieved but usually it is in the range of around 60-80 [143]. Using a longer drift tube improves resolving power.

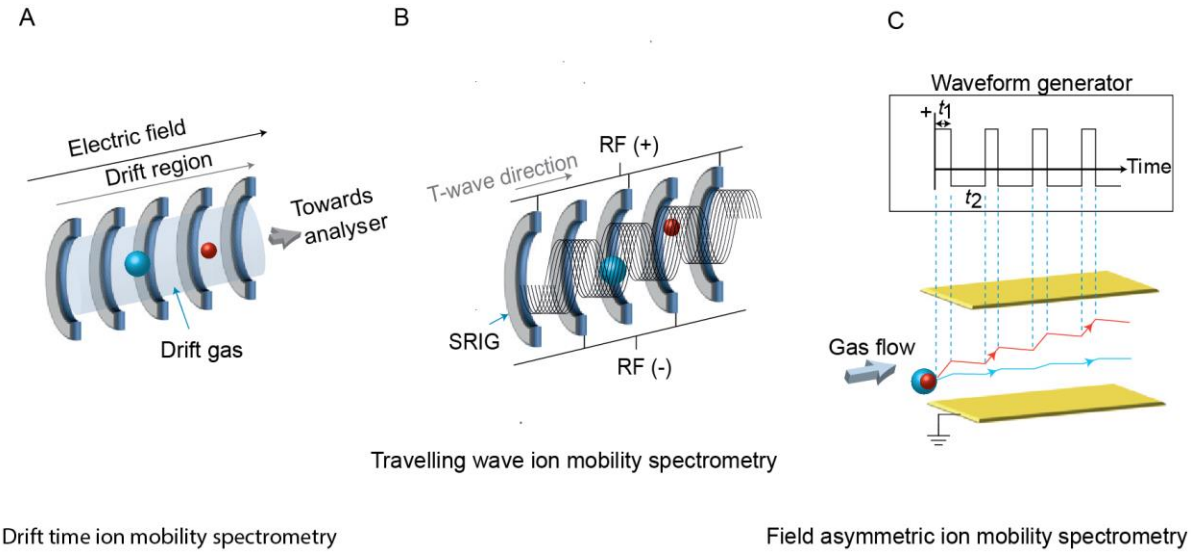


Figure 12: Schematic of different commercially available ion mobility devices. (A) In drift-time ion mobility spectrometry (DTIMS) a linear electric field is applied to a drift tube containing a buffer gas. Separation is achieved because depending on their size, ions have greater or fewer collisions with the buffer gas. (B) Travelling wave ion mobility spectrometry (TWIMS) uses a travelling voltage wave to push ions through the device. Separation is achieved because higher-mobility ions tend to be carried by the wave, whereas ions with lower mobility roll over the wave. (C) Field asymmetric ion mobility spectrometry (FAIMS) uses an alternating asymmetric electric field so that ions drift through the electrodes at different rates. By applying a compensation voltage, ions of only one specific cross section are repelled and refocused and are prevented from colliding with the electrodes. Small ions are in red, large ions in blue. Adapted from [155].

1.3.3 Travelling Wave Ion Mobility Spectrometry – Mass Spectrometry

Travelling wave ion mobility spectrometry (TWIMS) was introduced a decade ago by Waters Corporation [156-159]. The TWIMS device consists of a series of planar electrodes, termed stacked ring ion guides that are arranged orthogonally to the transmission axis (Figure 12B, 13). In contrast to the drift tube, TWIMS applies a dynamic electric field in a travelling wave manner [156]. Ions are radially confined by RF voltages at the electrodes, while a superimposed DC voltage creates the travelling wave. A wave moves from the entrance of the T-wave ion guide until the exit and pushes ions toward the detector. Dependent on the size and shape ions are picked up by the wave and slowed down by this action. Here, the measured mobility is dependent on how often a wave passes the ion and how many waves are needed for an ion to traverse the T-wave ion guide, which is filled with gas [156, 160]. Ions with higher mobility are carried by the wave and therefore pass the TWIMS device faster than low mobility ions, which roll over some of the waves. This technique has a higher ion transmission efficiency compared to other devices. For the determination of CCS with a TWIMS device, calibration of the drift time under defined conditions is absolutely necessary because the changing electric field means the relationship between Ω and K_0 is not valid in the Mason-Schamp equation. For this reason analytes with known CCS are used for calibration.

The T-wave device is available in combination with a TOF instrument, as the Synapt-G2 mass spectrometer from Waters Corporation (Figure 13). Here, the ion mobility device is located downstream of the analytical quadrupole, which is different to the other wide-spread IMS-MS instruments in which ions are first accumulated before they are injected into the ion mobility device followed by MS analysis. In those instruments ion mobility separation is performed in the stacked ring ion guides described above and they can also be used for fragmentation before mobility separation. One advantage of TWIMS-MS instruments is that ion separation is dependent on the height and the velocity of the traveling wave. By altering these parameters, specific ions can be targeted with high mobility resolution. Synapt-G2 instruments are used in different fields [161, 162], however the currently achievable resolving power of around 45 is relatively low [158].

Scientists at Waters have previously described a DIA method termed MS^E in which the entire mass range is fragmented and peptide are identified by correlating precursor mass elution profiles with those of multiplexed fragments [163, 164]. An ion mobility based refinement, $HDMS^E$ is used on TWIMS-MS instruments. Here, parallel fragmentation of multiple precursor ions is performed and ion mobility arrival times of parent ions and fragment ions are aligned, followed by database

searching. With this method over 50% more peptides and proteins could be detected in comparison to the MS^E method alone [165], indicating the potential of adding ion mobility information. Instead of using a fixed collision energy for the fragmentation of all present precursors, an ion mobility-dependent collision energy can be used, resulting in a significant increase in fragmentation efficiency and peptide and protein identification rate [130].

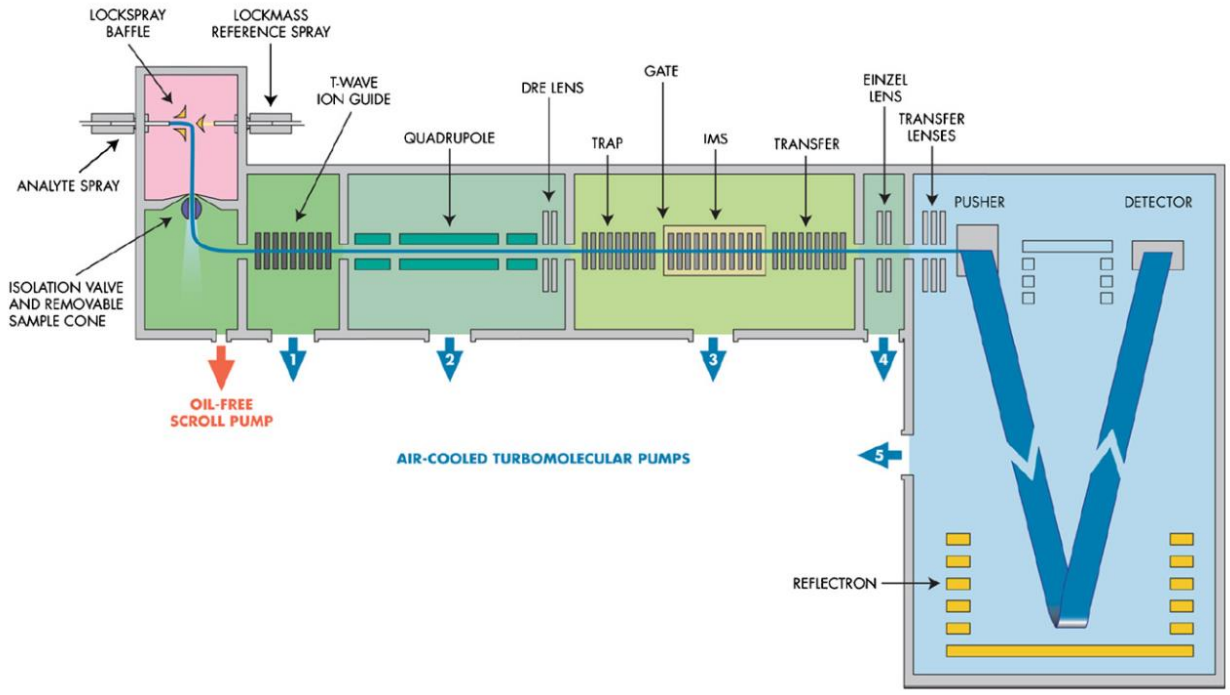


Figure 13: Hybrid quadrupole/TWIMS/TOF mass spectrometer (Synapt HDMS mass spectrometer from Waters). Adapted from [159].

1.3.4 Field asymmetric Ion Mobility Spectrometry – Mass Spectrometry

The field asymmetric ion mobility spectrometry (FAIMS) device consists of two electrodes, across which an electric field is applied (Figure 12C). The two main electrode configurations are planar [166] and cylindrical [167]. In the planar configuration ions are separated in high electric fields (> 7500 V/cm). Ions enter the device perpendicular to the electric field and collinearly to the drift gas. Separation is achieved by dispersion and compensation voltages. By using alternating high and low electric fields, ions radially drift as they traverse the electrodes. If a dispersion voltage was

applied exclusively, ions would collide with the electrodes. Therefore a compensation voltage refocuses the ions and this voltage is set to the properties of a desired ion type. Depending on the compensation voltage, specific ions are repelled in such a way as to traverse the device while other ions are removed. In this way, the FAIMS device is mostly used as a mobility filter, removing unwanted interfering ions, which enhances selectivity in combination with MS [168] due to their orthogonal separation principles. Note that CCSs cannot be determined by FAIMS because the electric field used exceeds the low-field limit.

FAIMS can be combined with any type of mass spectrometer and is commercially available from Thermo Scientific and AB Sciex, amongst others.

Table 1: Principle characteristics of the most wide-spread IMS devices (adapted from [169] and [155])

	DTIMS	**TWIMS**	**FAIMS**
Drift gas	yes	yes	yes
Pressure	Ambient (1 bar)	0.025-3 mbar	Ambient (1 bar)
Temperature	Ambient (300 K)	360 K	Ambient (300 K)
Electric Field E	Uniform low	Dynamic and non-uniform low	Alternating asymmetric high/low
Advantages	+ CCS can be determined + high resolution	+ CCS can be determined + separation of precursors and fragments	+ high resolution + combination with diverse mass spectrometer
Disadvantages of IMS-MS instruments	+ separation only possible after ionization + ion losses during transfer from atmospheric pressure to reduced pressure	+ Calibration necessary + low resolving power + ion heating	+ CCS cannot be determined + separation only possible directly after ionization + low usage of generated ions/only applicable for selected ions or ion mobility ranges

1.3.5. Trapped Ion Mobility Spectrometry – Mass Spectrometry

One of the latest developments in ion mobility spectrometry is trapped ion mobility spectrometry (TIMS), introduced by Park and co-workers at Bruker [123]. The TIMS analyzer consists of pairs of stacked electrodes and is divided into three regions: entrance funnel, TIMS tunnel, and exit funnel (Figure 14A). Due to the division of the TIMS analyzer into quadrants, it is possible to apply RF potentials independently to each quadrant [170, 171]. This creates a dipole field (for trapping

and focusing ions) in the entrance and exit funnel and a quadrupole field (for confining ions) in the TIMS tunnel. In addition, a DC field is superimposed on the RF field of the funnel and tunnel plates. This leads to an axial electric field gradient in the TIMS tunnel, which is set via a resistor divider chain. During TIMS analysis, a fixed DC potential is applied to the exit of the tunnel and a scanable voltage to the entrance of the tunnel. Therefore the mobility range, resolution and analysis speed is dependent on the settings of the entrance potential. In the entrance and exit funnels, the DC potential is responsible for pushing the ions downstream.

During operation, ions are introduced to the TIMS analyzer through a glass capillary and they are then deflected into the entrance funnel. Then ions are guided through the very compact (around 10 cm long) TIMS tunnel via a flow of gas, which drags them along in the presence of a counteracting electric field (Figures 14A and 14B). Depending on their mobility, ions rest at a position where the two forces are equal. Ions with larger CCS are close to the exit of the TIMS tunnel, whereas ions of smaller CCS are close to the entrance, the opposite of the separation order of DTIMS and TWIMS [123]. Ions are accumulated for a user-defined length of time in the TIMS tunnel, and they are then released or eluted by decreasing the electric field strength, in a manner that is also adjustable by the user [172].

One major advantage of the TIMS device is that the mobility resolution is not dependent on the length of the tunnel and is easily adjustable [172]. The mobility resolution is controlled by the bath gas velocity and the electric field ramp speed and Park and co-workers showed that a resolution of > 200 in principle can be achieved [173]. Another very attractive feature of TIMS device is that due to its very compact configuration, it can be easily incorporated into different mass spectrometers such as the micrOTOF-Q instrument [123] or FT-ICR mass spectrometer [174] (both from Bruker Daltonics). Moreover the user can decide to measure with or without TIMS mode and can also adjust the duty cycle [172].

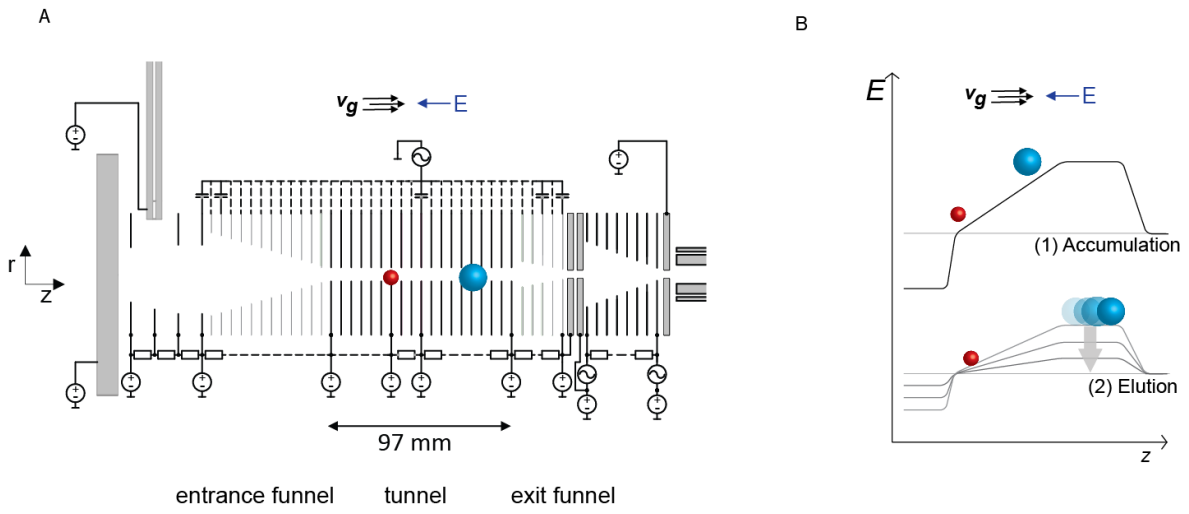


Figure 14: Principle of trapped ion mobility spectrometry (TIMS). (A) Schematic of the TIMS device. Ions coming from the capillary inlet are deflected by the deflection plate and enter the very compact (around 10 cm in length) TIMS tunnel through the entrance funnel. Ions are stored in the TIMS tunnel and once released, enter the mass spectrometer via the exit funnel. (B) General TIMS operation mode, including ion accumulation and serial elution of ion mobility separated ions from the TIMS tunnel by decreasing the electric field gradient. Adapted from [175].

1.4. Applications of MS-based proteomics

MS-based proteomics can be applied amongst others for the analysis of almost complete proteomes of model organisms such as yeast [176], for the analysis of post-translation modifications (PTMs) as well as for the investigation of protein-protein interactions. In addition to the global analysis of protein composition, MS-based proteomics can also quantify them. The detection of all existing proteins is difficult since a large number of proteins are only expressed at specific conditions and time points and vary from cell type to cell type. Dependent on the biological question, the analysis of a sub-proteome may be preferred. For example, for the investigation of protein-protein interactions, only the proteins of interest and their interacting partners are used from a crude protein mixture and for the study of PTMs only the corresponding sub-proteome is used. Another application of MS-based proteomics besides interaction proteomics and PTM-related proteomics is clinical proteomics. Here, protein mixtures derived from patient material are used. The aim of clinical proteomics is to contribute to diagnosis and treatment of various diseases.

Mass spectrometry-based interaction proteomics: Identifying all molecular interactions of a biological system has been one goal in systems biology for a long time. Different approaches beside mass spectrometry have been developed to investigate protein-protein interactions [177, 178]. A well-known workflow in MS-based interaction proteomics is the so-called affinity purification-mass spectrometry (AP-MS) approach [179]. Here, protein-protein interactions are investigated by immobilizing the protein of interest via an antibody or genetically encoded tag to an affinity matrix, before using it as a bait to capture its interacting proteins from the whole proteome. After several washing steps to eliminate unspecific interaction partners, the proteins that interact with the protein of interest are eluted from the matrix and analyzed in the mass spectrometer. AP-MS had some limitations such as the loss or under-representation of weak binding partners due to the washing steps in the sample preparation workflow or the misinterpretation of results since specific and unspecific binding interactors were not directly distinguishable. With the introduction of quantitative proteomics those limitation could have been largely solved [180, 181]. For example, isotopic labeling approaches made it possible to distinguish true interacting partners from unspecific binders. However, isotopic labeling is still difficult to perform on a large-scale. Therefore, protein-protein interactions are mostly performed by label-free quantification [182, 183]. In addition to protein-protein interactions MS-based interaction proteomics can investigate protein interactions with DNA and RNA. Moreover, it can be used to analyze the topology and stoichiometry of complexes by using chemical cross-linkers [184, 185].

Large-scale analysis of post-translational modifications (PTMs): Post-translational modifications play a major role in signal transduction. Mostly this involves not only the PTM change of one protein, but rather a cascade of PTM alterations. To understand the complete signaling network, global analysis of PTMs is performed. PTMs can easily be detected with mass spectrometry because if a PTM is present, it leads to a shift in the mass of the modified peptide. A large number of modifications are known, such as phosphorylation [186, 187], lysine acetylation [188, 189] or methylation [190]. But the analysis of PTMs is also very challenging, since modified peptides are generally very low abundant, making MS/MS spectra analysis more complicated and because database search becomes much larger when allowing potential modifications. For the global analysis of PTMs modified peptides are usually enriched to reduce the sample complexity, increase peptide amounts, increasing the probability for detection by LC-MS analysis. PTMs on peptides can be localized at different positions. Therefore, to localize PTMs with high confidence, high quality MS2 spectra are necessary to detect the relevant fragment ion. Post-data processing then provides a localization score for each PTM, which reflects the confidence of localizing the

PTM to a specific site. With current MS technology more than 10,000 phosphopeptides can be identified in a typical experiment [191].

Mass spectrometry-based clinical proteomics: Over the last years improvements in MS instrumentation as well as in the entire proteomics workflow and in computational proteomics have made it possible to apply mass spectrometry to clinical proteomics. For example, MS-based proteomics could in principle be used to classify patients into treatment resistant and non-resistant groups. Moreover MS-based proteomics could in the future potentially be used for the measurement of disease relevant biomarkers [192], including troponin I that indicates a myocardial infarction [193]. In biomarker discovery blood, plasma and serum are the sources of choice because they are easily accessible from patients and apart from the typical plasma proteins they contain proteins from all tissues in the body since it has contact to them. In clinical practice immunoassays are used for routine quantitative analyses. However, immunoassays have limitations such that multiplexing is very limited or not possible and that isoforms cannot necessarily be distinguished. Therefore, MS-based proteomics is very attractive because it has the potential to address those limitations. Nevertheless, the analysis of the plasma proteome is also challenging for MS, for instance due to the high dynamic range of proteins in the blood plasma. Moreover, to use MS-based proteomics for clinical samples, a very reproducible, robust and high-throughput sample preparation workflow is necessary. Mass spectrometry in combination with immunoaffinity-based approaches, termed immunoaffinity-based-MS (IA-MS) is already applied for the quantification of biomarkers [194, 195]. Here, proteins of interest are first enriched by specific antibodies and then analyzed by targeted proteomic approaches. Like enzyme-linked immunosorbent assays, IA-MS only analyzes proteins of interest. With both techniques new biomarkers cannot be identified. This is in contrast to discovery proteomics, which has the potential to quantify known biomarkers but also identify and quantify new biomarkers [192]. Besides blood plasma, other body fluids or tissues are investigated in biomarker discovery. For example, urine is used for the diagnosis of diabetes or bladder cancer, or lavage fluid for lung diseases. Those have the analytical advantage that the difference in protein concentration is lower in comparison to blood plasma.

In summary, MS-based proteomics is already successfully applied in several areas. It can be used to compare protein-protein interactions, detect changes in PTMs between different samples as well as classify patients dependent on their disease status. In the future it may have a great impact on personalized medicine, including better diagnose and treatment of patients.

1.5. Objectives

Fundamental challenges in MS-based proteomics include the high sample complexity and the difference in protein concentration in a biological sample. Developments in mass spectrometers towards higher sensitivity, sequencing speed and resolution partly address these challenges, however, overcoming them remains a major goal of the field of MS-based proteomics and therefore continued improvements in MS technology are highly desirable.

Quadrupole time-of-flight (QTOF) mass spectrometers are one of the two major instrument types used in proteomics. QTOFs are applied in different areas including metabolomics and the screening of small molecules. Today, for very deep proteome analyses the Orbitrap mass analyzers combined with a linear ion trap or quadrupole has become the workhorse in most laboratories due to their very high resolution and mass accuracy [55]. However, TOF technology has fundamental attractions, such as the very high scanning speed and the absence of space charge effects (which is a limitation of all trapping instruments) and therefore there is a great interest in the community to use QTOFs for deep shotgun proteomics analyses. To date, QTOF instrument have not achieved the performance of hybrid Orbitrap instruments. In this thesis, I have developed, improved and evaluated novel QTOF technologies for deep shotgun proteomics analyses.

I started working on a novel benchtop quadrupole time-of flight (QTOF) mass spectrometer, the impact II from Bruker, for shotgun proteomics applications. This instrument features several new hardware developments such as a novel collision cell, reflectron and detector to improve the ion transmission efficiency and resolving power. In this thesis, I first describe and evaluate these new technologies and conclude that the increase in ion extraction and resolution was enormous compared to previous TOF instruments. The next goal was to investigate how reproducibly and deeply the proteome can be covered in a single run format with this instrument. For this purpose I measured peptide mixtures generated from a human cancer cell line (HeLa) and from yeast cells, standard samples in the proteomics community in a single run format. After establishing optimal instrument settings, I investigated the depth of the proteome achievable from HeLa after high-pH reversed-phase fractionation, a particularly effective way of pre-fractionation peptide mixtures [42, 43]. Tissues are more challenging to analyze by MS-based proteomics because they are composed of different cell types as well as extracellular matrix and I evaluated the instrument on a complex peptide mixture derived from murine cerebellum. Here, I achieved the highest proteome coverage reported with a QTOF instrument for any proteomics system so far.

In proteomics it is not just important to identify proteins but also to quantify them accurately. Therefore, I next investigated the reproducibility and accuracy of label-free quantification in diverse experiments. For instance, I determined the coefficients of variation (CVs) between three technical replicates, which yielded very low CVs, and analyzed the fold changes in a spike-in experiment, which likewise resulted in results very close to those expected from the mixing ratios. Finally, I analyzed haploid against diploid yeast and characterized overall proteome differences between diverse mouse cell lines, which demonstrated potential application areas of this instrument.

In my second project, the QTOF from the first project was additionally equipped with trapped ion mobility spectrometry (TIMS) in front of the QTOF mass analyzer. As explained above, this separates ions based on their size-to-charge ratio in the TIMS device before separating them depending on their mass-to-charge ratio in the mass spectrometer. TIMS in combination with mass spectrometry has several potential advantages: improved signal-to-noise ratio as signal is concentrated while noise is distributed and improved precursor ion fraction values since precursors in the same quadrupole isolation window are separated by their differences in ion mobility. In a TIMS-MS/MS analysis, the quadrupole isolates a specific ion population within a complete TIMS scan (usually every 50 ms) for further fragmentation. However, this is not optimal because each different ion population elutes from the TIMS device within a few ms. Therefore, the aim of this project was to investigate if multiple precursors could be selected in a single TIMS scan, which would result in an increase in sequencing speed. First, the hardware requirements were established, including parameters that ensure high ion mobility resolution, efficient storage of precursor ions and extremely rapid switching of the quadrupole. In experiments involving direct infusion of a complex peptide mixture, four precursors instead of one could be selected in every TIMS scan, resulting in a four times higher sequencing speed. Finally, the effects of such an improvement were modelled for shotgun proteomics by using the HeLa data-set from the first project. We showed that this method, termed parallel accumulation – serial fragmentation (PASEF), could represent a large improvement in shotgun proteomics as 70% of the detectable peptide features could be targeted instead of around 20% without applying the PASEF method.

In my third project, I investigated if TIMS combined with mass spectrometry could be applied for reducing the compression problem in isobaric tagging experiments described above. Clearly, ion mobility separation has the potential to separate the precursors present in the same isolation window by their ion mobility. In this project, I first have investigated which ion mobility resolution can be achieved by varying the release time from the TIMS device. Next, I modelled how often

co-isolation and co-fragmentation of peptides appear with and without applying TIMS-MS. Moreover, I investigated in a real TIMS-MS analyses of TMT 6-plex labeled complex peptide mixtures if a median resolution of 78 is sufficient for accurate quantification. This resolution already reduces the number of interfering peptides up to two-fold and additional improvements could help to further reduce the number of chimeric MS2 spectra.

2. RESULTS

2.1. A novel QTOF mass spectrometer for shotgun proteomics analyses

Scarlet Beck, Annette Michalski, Oliver Raether, Markus Lubeck, Stephanie Kaspar, Niels Goedecke, Carsten Baessmann, Daniel Hornburg, Florian Meier, Igor Paron, Nils A. Kulak, Juergen Cox, and Matthias Mann

The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics

Molecular & Cellular Proteomics, 2015 July; 14(7)

Fundamental challenges of shotgun proteomics include the very large numbers of peptides that elute from the column simultaneously and peptide concentrations that vary by many orders of magnitude. To address these critical challenges, developments in instrumentation towards higher sensitivity, sequencing speed and resolving power have been made in recent decades.

In MS-based proteomics two major types of instrumentation are used, the hybrid Orbitrap and QTOF mass spectrometers. TOF mass analyzers possess several advantages including higher sequencing speed and absence of space charge effects in comparison to trapping mass analyzers. However, for deep shotgun proteomics analyses hybrid Orbitrap mass spectrometers are used in the majority of laboratories worldwide due to their superior resolution and mass accuracy.

In 2014, our laboratory started a close collaboration with Bruker Daltonics Inc. with the aim of co-developing, improving and evaluating several hardware and software features of a novel QTOF mass spectrometer, termed impact II, for shotgun proteomics. This instrument features a new collision cell, reflectron and detector, which are described and evaluated in the following paper. With the improved collision cell twice as many ions can be extracted and developments on the reflectron and detector technologies yield 80% greater resolving power (40,000 at *m/z* 1222) compared to the previous model.

The ion transmission efficiency in a QTOF instrument was previously unknown. Therefore we set out to measure the number of ions that successfully pass through the instrument – starting from where they enter the vacuum system to where they reach the detector. By direct infusion of a bovine serum albumin (BSA) or blank solution, the ion current difference between these two

conditions could be measured which established that a total of 10% of the BSA ions were detected.

Motivated by this excellent proportion of detectable ions, I investigated the performance of the impact II for single shot analysis by analyzing complex peptide mixtures derived from a mammalian cell line and yeast, standard samples in the proteomic community. With optimized instrument settings and changes in the heated liquid chromatography system, I achieved high reproducibility ($R^2 > 0.99$) and good proteome coverage (4800 proteins with a 90 min gradient). To evaluate the impact II for very deep proteome coverage, high-pH reverse-phase pre-fractionation was performed. I identified more than 11,200 proteins in cerebellum, the deepest proteome coverage reported with a QTOF mass spectrometer so far.

In MS-based proteomics it is important to quantify each identified protein. For this purpose, I also evaluated the instrument for label-free quantification in different experiments. After showing that very reproducible and accurate quantitation can be achieved, I next applied the QTOF-based workflow to two proteomic experimental questions. I analyzed diploid and haploid S. cerevisiae to detect proteome changes between them as well as different murine cell lines to perform global proteomic comparisons.

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics*[S]

Scarlet Beck‡, Annette Michalski§, Oliver Raether§, Markus Lubeck§, Stephanie Kaspar§, Niels Goedecke§, Carsten Baessmann§, Daniel Hornburg‡, Florian Meier‡, Igor Paron‡, Nils A. Kulak‡, Juergen Cox¶, and Matthias Mann‡‖

Hybrid quadrupole time-of-flight (QTOF) mass spectrometry is one of the two major principles used in proteomics. Although based on simple fundamentals, it has over the last decades greatly evolved in terms of achievable resolution, mass accuracy, and dynamic range. The Bruker impact platform of QTOF instruments takes advantage of these developments and here we develop and evaluate the impact II for shotgun proteomics applications. Adaption of our heated liquid chromatography system achieved very narrow peptide elution peaks. The impact II is equipped with a new collision cell with both axial and radial ion ejection, more than doubling ion extraction at high tandem MS frequencies. The new reflectron and detector improve resolving power compared with the previous model up to 80%, *i.e.* to 40,000 at *m/z* 1222. We analyzed the ion current from the inlet capillary and found very high transmission (>80%) up to the collision cell. Simulation and measurement indicated 60% transfer into the flight tube. We adapted MaxQuant for QTOF data, improving absolute average mass deviations to better than 1.45 ppm. More than 4800 proteins can be identified in a single run of HeLa digest in a 90 min gradient. The workflow achieved high technical reproducibility (R2 > 0.99) and accurate fold change determination in spike-in experiments in complex mixtures. Using label-free quantification we rapidly quantified haploid against diploid yeast and characterized overall proteome differences in mouse cell lines originating from different tissues. Finally, after high pH reversed-phase fractionation we identified 9515 proteins in a triplicate measurement of HeLa peptide mixture and 11,257 proteins in single measurements of cerebellum—the highest proteome coverage reported with a QTOF instrument so far. ***Molecular & Cellular Proteomics 14: 10.1074/mcp.M114.047407, 2014–2029, 2015.***

Building on the fundamental advance of the soft ionization techniques electrospray ionization and matrix-assisted laser desorption/ionization (1, 2), MS-based proteomics has advanced tremendously over the last two decades (3–6). Bottom-up, shotgun proteomics is usually performed in a liquid chromatography-tandem MS (LC-MS/MS)[1] format, where nanoscale liquid chromatography is coupled through electrospray ionization to an instrument capable of measuring a mass spectrum and fragmenting the recognized precursor peaks on the chromatographic time scale. Fundamental challenges of shotgun proteomics include the very large numbers of peptides that elute over relatively short periods and peptide abundances that vary by many orders of magnitude. Developments in mass spectrometers toward higher sensitivity, sequencing speed, and resolution were needed and helped to address these critical challenges (7, 8). Especially the introduction of the Orbitrap mass analyzers has advanced the state of the art of the field because of their very high resolution and mass accuracy (9, 10). A popular configuration couples a quadrupole mass filter for precursor selection to the Orbitrap analyzer in a compact benchtop format (11–13).

In addition to the improvements in MS instrumentation, there have been key advances in the entire proteomics workflow, from sample preparation through improved LC systems and in computational proteomics (14–16). Together, such

From the ‡Proteomics and Signal Transduction, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany; §Bruker Daltonik GmbH, Fahrenheitstr. 4, 28359 Bremen, Germany; ¶Computational Systems Biochemistry, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

Received December 12, 2014, and in revised form, April 21, 2015
Published, MCP Papers in Press, May 19, 2015, DOI 10.1074/mcp.M114.047407

Author contributions: S.B., A.M., O.R., M.L., C.B., J.C., and M.M. designed research; S.B., A.M., O.R., M.L., S.K., N.G., F.M., I.P., N.A.K., and J.C. performed research; S.B., O.R., N.G., D.H., N.A.K., and J.C. contributed new reagents or analytic tools; S.B., A.M., O.R., M.L., S.K., F.M., and M.M. analyzed data; S.B., A.M., O.R., M.L., J.C., and M.M. wrote the paper.

[1] The abbreviations used are: LC-MS/MS, liquid chromatography-tandem MS; CAA, chloroacetamide; ES, electrospray; FDR, false discovery rate; Hepa 1–6, mouse hepatoma; ID, inner diameter; LFQ, label-free quantification; MCP, multichannel plate; MEFs, mouse embryonic fibroblasts; MS/MS, tandem mass spectrometry; NSC-34, spinal cord neuron-neuroblastoma; QTOF, quadrupole time-of-flight; TOF, time-of-flight; UPS, universal protein standard.

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

advances are making shotgun proteomics increasingly comprehensive and deep analyses can now be performed in a reasonable time (13, 17–19). Nevertheless, complete analysis of all expressed proteins in a complex system remains extremely challenging and complete measurement of all the peptides produced in shotgun proteomics may not even be possible in principle (20, 21). Therefore, an urgent need for continued improvements in proteomics technology remains.

Besides the Orbitrap analyzer and other ion trap technologies, the main alternative MS technology is time-of-flight, a technology that has been used for many decades in diverse fields. The configuration employed in proteomics laboratories combines a quadrupole mass filter via a collision cell and orthogonal acceleration unit to a reflectron and a multichannel plate (MCP) detector (22). TOF scans are generated in much less than a millisecond (ms), and a number of these "pulses" are added to obtain an MS or MS/MS spectrum with the desired signal to noise ratio. Our own laboratory has used such a quadrupole time-of-flight (QTOF) instrument as the main workhorse in proteomics for many years, but then switched to high-resolution trapping instruments because of their superior resolution and mass accuracy. However, TOF technology has fundamental attractions, such as the extremely high scan speed and the absence of space charge, which limits the number of usable ions in all trapping instruments. In principle, the high spectra rate makes TOF instruments capable of making use of the majority of ions, thus promising optimal sensitivity, dynamic range and hence quantification. It also means that TOF can naturally be interfaced with ion mobility devices, which typically separate ions on the ms time scale. Data independent analysis strategies such as MS^E, in which all precursors are fragmented simultaneously (23, 24) or SWATH, in which the precursor ion window is rapidly cycled through the entire mass range (25), also make use of the high scanning speed offered by QTOF instruments. It appears that QTOFs are set to make a comeback in proteomics with recent examples showing impressive depth of coverage of complex proteomes. For instance, using a variant of the MS^E method, identification of 5468 proteins was reported in HeLa cells in single shots and small sample amounts (26). In another report, employing ion mobility for better transmission of fragment ions to the detector led to the identification of up to 7548 proteins in human ovary tissue (27).

In this paper, we describe the impact II™, a benchtop QTOF instrument from Bruker Daltonics, and its use in shotgun proteomics. This QTOF instrument is a member of an instrument family first introduced in 2008, which consists of the compact, the impact, and the maXis. The original impact was introduced in 2011 and was followed by the impact HD, which was equipped with a better digitizer, expanding the dynamic range of the detector. With the impact II, which became commercially available in 2014, we aimed to achieve a resolution and sequencing speed adequate for demanding shotgun proteomics experiments. To achieve this we developed an improved collision cell, orthogonal accelerator scheme, reflectron, and detector. Here we measure ion transmission characteristics of this instrument and the actually realized resolution and mass accuracy in typical proteomics experiments. Furthermore, we investigated the attainable proteome coverage in single shot analysis and we ask if QTOF performance is now sufficient for very deep characterization of complex cell line and tissue proteomes.

EXPERIMENTAL PROCEDURES

Preparation of HeLa Lysates—HeLa cells (ATCC, S3 subclone) were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum, 20 mM glutamine and 1% penicillin-streptomycin (all from PAA Laboratories, Freiburg, Germany). Cells were collected by centrifugation at 200 × *g* for 10 min, washed once with cold phosphate buffered saline (PBS) and centrifuged again. Supernatant was carefully discarded and the cell pellet shock frozen in liquid nitrogen and stored at −80 °C until further use. A pellet containing 5×10^7 cells was resuspended in 1.5 ml of ice cold Milli-Q water, then an equal volume of trifluoroethanol (Sigma-Aldrich, Taufkirchen, Germany) was added. The cell suspension was kept on ice for 10 min, vortexed for 1 min and sonicated for 2 min at 20% duty cycle and output control 3 (Branson Ultrasonics sonifier, Danbury, CT; model 250). After the addition of 200 µl Tris (pH 8.5, final concentration: 100 mM), 400 µl TCEP (final concentration: 10 mM) and 400 µl 2-chloroacetamide (CAA) (final concentration: 40 mM) the lysate was incubated for 10 min at 95 °C. Then the sample was diluted to 15 ml with 50 mM ammonium bicarbonate. The mixture was digested by adding LysC (Wako Chemicals GmbH, Neuss, Germany; ratio 1 µg LysC:100 µg sample protein) for 2 h at 37 °C, followed by adding trypsin (ratio 1 µg trypsin:75 µg sample protein, Promega GmbH, Mannheim, Germany) at 37 °C overnight. After a further digestion with trypsin (ratio 1:125) for 5 h at 37 °C, the digested peptides with an estimated concentration of 1 µg/µl were diluted 1:4 with water and acidified by adding formic acid (FA) (final concentration: 0.2%) and purified on Sep-Pak tC18 cartridges (Waters, Milford, MA) according to manufacturer's instructions. Peptide concentration was determined using a NanoDrop spectrophotometer (Thermo Scientific, Wilmington, DE).

Preparation of Yeast Lysates—Saccharomyces cerevisiae strains BY4742 and BY4743 (EUROSCARF) were grown at 30 °C in yeast extract peptone dextrose (YPD) media (10 g/l BactoYeast extract, 20 g/l Bacto™ peptone (BD), 2% w/v glucose). Cells were grown to log phase (OD_{600} of 0.6), harvested by centrifugation at 1600 × *g* for 10 min at 4 °C, washed with cold Milli-Q water and then collected again by centrifugation at 10,000 × *g* for 5 min at 4 °C. Cells were lysed in 1% sodium deoxycholate, 10 mM TCEP, 40 mM CAA in 100 mM Tris pH 8.5, boiled for 10 min at 95 °C and sonicated for 3 min at 30% duty cycle and output control 3 (Branson Ultrasonics sonifier; model 250). Protein concentrations were determined by tryptophan fluorescence emission assay. Cell lysates were diluted 1:2 with Milli-Q water and digested by adding LysC (Wako Chemicals GmbH, ratio 1 µg LysC:50 µg sample protein) for 4 h at 37 °C, followed by adding again LysC (ratio 1:50) overnight at 37 °C. An equal volume of ethyl acetate acidified with 1% TFA was added to the solution, samples were vortexed for 2 min and digested peptides were purified with SDB-RPS StageTips as described in Kulak *et al.* (19). Peptide concentrations were determined using a NanoDrop spectrophotometer.

Preparation of MEFs, Hepa, and NSC Cell Line Lysates—Spinal cord neuron-neuroblastoma (NSC-34) (CED-CLU140, Biozol, Eching, Germany), mouse embryonic fibroblasts (MEFs) (American Type Culture Collection, Manassas, VA), and mouse hepatoma (liver cancer, Hepa 1–6) (CRL-1830, American Type Culture Collection) cell lines

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

were cultured and proteins prepared as previously described (28). Briefly, the cells were lysed in lysis buffer (4% SDS, 10 mM Hepes, pH 8.0) during sonication for 15 min (level 5, Bioruptor; Diagenode, Seraing (Ougrée) - Belgium). Cell lysis was followed by reduction of disulfide bonds with 10 mM DTT for 30 min and subsequent alkylation with 55 mM IAA for 45 min. To remove the detergent, cold acetone (−20 °C) was added to 100 μg of proteins to a final concentration of 80% v/v, and proteins were precipitated for at least 2 h at −20 °C. The suspension was centrifuged for 15 min (4 °C, 16,000 × *g*) and the precipitate was washed with 80% acetone (−20 °C) prior to re-suspension in 50 μl of 6 M urea/2 M thiourea, 10 mM Hepes, pH 8.0. An initial digestion step (3 h) was carried out after the addition of 1 μg of LysC, followed by dilution with four volumes of 50 mM ammonium bicarbonate and the final digestion with 1 μg of trypsin overnight at room temperature. The resulting peptide mixtures were desalted on SDB-RPS StageTips (29) and subjected to single shot LC-MS/MS analysis.

Preparation of Cerebellum Lysates—Cerebellum from a single mouse (strain: C57Bl6) was homogenized in 4% SDS in 100 mM Tris pH 7.6 using a FastPrep 24 homogenizer (MP Biomedicals, Eschwege, Germany), incubated for 10 min at 95 °C and sonicated for 3 min at 30% duty cycle and output control 3 (Branson Ultrasonics sonifier; model 250). To remove the detergent, acetone (−20 °C) was added to a final concentration of 80% v/v and proteins were precipitated overnight at −20 °C. Supernatants were carefully discarded after centrifugation at 1600 × *g* for 20 min at 4 °C, and the pellets were washed with 80% acetone (−20 °C). The protein pellets were dissolved in 8 M Urea in 10 mM Hepes and protein concentrations were determined by the tryptophan fluorescence emission at 350 nm using an excitation wavelength of 295 nm. Proteins were reduced with 10 mM DTT for 30 min and alkylated with 55 mM iodoacetamide for 20 min. After addition of thiourea to a final concentration of 0.1 M, samples were digested by adding LysC (Wako Chemicals, ratio 1 μg LysC:100 μg sample protein) for 3 h at RT, diluted with four volumes of 50 mM ammonium bicarbonate, and further digested with trypsin (ratio 1 μg trypsin:100 μg sample protein, Promega) at RT overnight. After a further digestion with LysC and trypsin (ratio 1:100) for 8 h at RT, digested peptides were acidified by adding TFA (final concentration: 0.5%) and purified on Sep-Pak tC18 cartridges (Waters) according to manufacturer's instructions. Peptide concentrations were determined using a NanoDrop spectrophotometer.

Sample Preparation for Quantification—Universal Proteomics Standard (UPS-1, Sigma-Aldrich) and Proteomics Dynamic Range Standard (UPS-2, Sigma-Aldrich), both containing 48 human proteins, either at equimolar concentrations (UPS-1) or formulated into a dynamic range of concentrations, covering five orders of magnitude (UPS-2), were prepared according to ref (30). Predigested yeast sample (Promega) was re-suspended in 0.1% trifluoroacetic acid to a final concentration of 500 ng/μl. Digested UPS-2 sample was spiked in two different amounts of 250 fmol to 2.5 amol peptide amount for sample 1 and 500 fmol to 5 amol for sample 2 into 500 ng yeast background, thereby creating two samples with a theoretical ratio 1:1 for the yeast proteome and 1:2 for the UPS peptides. In another sample, digested UPS-1 sample (25 fmol for all components) was spiked into 500 ng yeast.

High-pH Reverse-Phase Fractionation—We performed high-pH reversed-phase peptide prefractionation with fraction concatenation on 175 μg HeLa or cerebellum peptides on a 2.1 × 300 mm Acquity UPLC Peptide BEH column packed with 130 Å pore, 1.7 μm particle size C_{18} beads (Part No. 186005792, Waters). A gradient of basic reversed-phase buffers (Buffer A: 0.1% formic acid, ammonium hydroxide pH 10; Buffer B: 0.1% formic acid, 80% acetonitrile, ammonium hydroxide pH 10) was run on a Prominence HPLC system (Shimadzu, Duisburg, Germany) at a flow rate of 150 μl/min at 60 °C. The LC run lasted for 240 min with a starting concentration of 5% buffer B increasing to 30% over the initial 120 min and a further increase in concentration to 60% over 70 min. This elution gradient was followed by a 95% wash and re-equilibration. Fraction collection started after 0.2 ml elution and fractions were collected every 140 s resulting in 72 fractions used for concatenation into 24 fractions as described previously (31).

Inlet Capillary and CaptiveSpray—In our instrument, in contrast to many other commercial ion source designs, the high voltage for the electrospray (ES) process is applied to the vacuum capillary inlet, whereas the sprayer is kept at ground, which allows for a simpler source design (supplemental Fig. S1*A*). To electrically decouple the ES voltage and the electrical potential of the vacuum section, we use an inlet capillary made from high resistive glass (~1GOhm). Positioning the ES voltage at the capillary entrance means that the ions are transported opposite to the electrical gradient by the gas flow (32). In this configuration, charged molecules travel somewhat slower than the surrounding gas. According to Bernoulli's law, ions are then focused toward the area of highest gas velocity along the center axis of the capillary. The set-up tends to reduce the contamination of the inner capillary walls (33).

The Bruker CaptiveSpray nanoflow ES source is directly attached to the vacuum inlet capillary via a short capillary extension that can be heated using the instrument's drying gas (Supplemental Fig. S1*A*). The spray tip is automatically mechanically aligned on axis with the capillary inlet without the need for any adjustments. The principle of the CaptiveSpray is a vortex gas (usually air) that sweeps around the emitter spray tip at three different stages. The first one is designed to assist spray formation, the second and third one help to focus the spray plume into the MS inlet capillary. All three flows are created solely by the vacuum of the MS system, which requires that the entire source is vacuum sealed.

The spray emitter consists of a 2 cm long, 20 μm ID fused silica capillary. Its tip is etch-tapered, thus the inner diameter remains constant to the very end of the tip making it very robust against clogging. Furthermore, it also allows using the same emitter at flow rates ranging from 50 nl/min to 5 μl/min, thereby supporting a wide range of column types. Fused silica columns, which are often used for proteomics, are typically connected to the emitter via a low dead volume union (supplemental Fig. S1*A*), which also provides the electrical contact for keeping the electrospray at ground potential.

Minimizing Postcolumn Dead Volume—Using the described design, the CaptiveSpray source provides very stable ionization; however, when we initially coupled it to the LC set-up used in the Munich laboratory (17, 34), we observed broader LC peak elution distributions than we normally do (supplemental Fig. S1*B*). Furthermore, we wished to incorporate a column oven and pulled tip columns. We therefore constructed a modified source, which keeps the back end of the CaptiveSpray but replaces the front end by the standard set-up used in our department. The modified set-up incorporating the tip column is displayed in supplemental Fig. S1*C*. The modified design of the column holder allows for exact aligning and fixation of the column inside the CaptiveSpray source. Electrical grounding was applied using a connecting tee at the column head. This setup produced the desired, narrow LC peak distributions (supplemental Fig. S1*D*) and was used for the proteomic analyses described in this article.

LC-MS/MS Analysis—We used an Easy nLC-1000 (Thermo Fisher Scientific) on-line coupled to an impact II (Bruker Daltonics) with a CaptiveSpray ion source (Bruker Daltonics). The peptide mixtures (1 μg) were loaded onto an in-house packed column (50 cm, 75 μm inner diameter) filled with C_{18} material (ReproSil-Pur C_{18} AQ 1.9 μm reversed phase resin, Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). Chromatographic separation was carried out using a linear gradient of 5–30% buffer B (80% ACN and 0.1% FA) at a flow rate of

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

250 nl/min over 90 min. Because of loading and washing steps, the total time for an LC-MS/MS run was about 40 to 50 min longer.

Generally, LC-MS/MS data were acquired using a data-dependent auto-MS/MS method selecting the 17 most abundant precursor ions in cycle for fragmentation and an MS/MS summation time adjusted to the precursor intensity (Compass 1.8 acquisition and processing software, Bruker Daltonics). For the deep proteome measurements of a cell line in combination with peptide fractionation, we used a "dynamic method," with a fixed cycle time of 3 s. The mass range of the MS scan was set to extend from *m/z* 150 to 1750. Dynamic exclusion duration was 0.4 min. Isolation of precursor ions was performed using an *m/z* dependent isolation window of 1.5–5 Th. The collision energy was adjusted between 23–65 eV as a function of the *m/z* value.

For the quantitative analysis of the UPS standards in yeast we used a trapping column set-up (PepMap pre-column, 2 cm x 100 μm; Thermo Scientific) and a Dionex HPLC pump (Ultimate 3000, Thermo Scientific). For this experiment, peptides were separated on a PepMap UHPLC column (50 cm x 75 μm, 2 μm particles; Thermo Scientific) using a 90 min multistep ACN gradient (buffer A: 0.1% FA; buffer B 100% ACN in 0.1% FA). The unmodified CaptiveSpray ion source (see above) was used to interface the LC system to the impact II. For quantification full scan MS spectra were acquired at a spectra rate of 1Hz followed by acquisition of 1 MS/MS spectrum. Six replicates per sample were acquired. For data acquisition of the UPS-1 in yeast sample, the 17 most intense precursor ions were selected for fragmentation, resulting in a total cycle time of 1.2 s.

Intact Protein Analysis—Adalimumab was cleaved at the hinge region with IdeS (FabRICATOR, Genovis) and reduced to obtain the Fc/2, Fd and light chain sub units as recently described in (35). The subunits were separated by chromatography (35) prior to analysis on the impact II. Data was analyzed using the SNAP algorithm to fit the theoretical pattern (36, 37).

Development of MaxQuant for QTOF Data—In general all processing steps from the standard MaxQuant computational workflow, which was optimized for the analysis of Orbitrap data, are also applied to QTOF data. The nonlinear mass recalibration algorithm experienced major adaptations. Its original form for the Orbitrap applies a recalibration function with nonlinear dependence on the two variables *m/z* and retention time. It has been extended to include the peak intensity as a third dimension that the mass recalibration depends on. This is necessary because of appreciable systematic nonlinear intensity dependent peak mass shifts that are typically found in time of flight data. The intensity dependence is parameterized as a polynomial in the logarithm of peak intensities. The new mass recalibration algorithm allows for high mass accuracy without the use of internal or external calibrants.

We added a new instrument type called "Bruker QTOF" in which several relevant parameters of algorithms for the processing of spectra are set to default values that are suitable for the analysis of data generated by the impact family. These parameters include mass matching windows for the assembly of 3D peaks, mass tolerances for assembling isotope patterns and labeling pairs, initial peptide mass tolerance windows for the Andromeda search and minimum required number of scans per 3D peak. Raw data can be immediately read from the proprietary Bruker binary format and no conversion to intermediate file formats is needed. Peak centroids are utilized as determined by the centroiding algorithms of in the Bruker software. The viewer module of MaxQuant is enabled for QTOF data, among other features allowing to visualize MS data in *m/z*-retention time maps and to annotate and export MS/MS spectra to fulfil journal requirements for reporting of spectral evidence.

Analysis of Proteomic Data—All data were analyzed with the MaxQuant software (version 1.5.2.8 or version 1.5.0.1) (38, 39) with the Andromeda search engine (38) with the adaptions and developments described above. The false discovery rate (FDR) was set to 1% for both proteins and peptides and we specified a minimum length of seven amino acids. MaxQuant scored peptides for identification based on a search with an initial allowed mass deviation of the precursor ion of up to 0.07 Da after time-dependent recalibration of the precursor masses. The allowed fragment mass deviation was 40 ppm. The Andromeda search engine was used for the MS/MS spectra search against the Uniprot human database (downloaded on June 21, 2014, containing 88,976 entries and 247 contaminants), the Uniprot Saccharomyces cerevisiae database (downloaded on June 21, 2014, containing 6643 entries), the Uniprot mouse database (downloaded on June 21, 2014, containing 51,573 entries) and UPS fasta file provided by Sigma-Aldrich (http://www.sigmaaldrich.com/life-science/proteomics/mass-spectrometry/ups1-and-ups2-proteomic.html) for quantitative study. Enzyme specificity was set as C-terminal to Arg and Lys, also allowing cleavage at proline bonds and a maximum of two missed cleavages. Carbamidomethylation of cysteine was selected as fixed modification and N-terminal protein acetylation and methionine oxidation as variable modifications.

The "match between runs" feature of MaxQuant was used to transfer identifications to other LC-MS/MS runs based on their masses and retention time (maximum deviation 0.7 min) and this was also used in quantification experiments. Quantifications were performed with the label-free algorithms described recently (39). We required a minimum peptide ratio count of two and at least one "razor peptide" for quantification. For cerebellum, the quantification was based on normalized protein intensities. Further analysis of data was performed in the MaxQuant Viewer, in the Perseus post data acquisition package that is part of MaxQuant (all freely available at www.maxquant.org) and in the R statistical computing environment (40).

Potential contaminants as well as proteins identified only by site modification were strictly excluded from further analysis.

For the quantitative analysis of the UPS standards in yeast, entries were only accepted if they had valid values in all 12 replicates. Results were then filtered for Welch-significant regulation of UPS-2 proteins.

Analysis of the yeast samples were based on label-free intensities (LFQ values). After filtering (3 valid values in at least one group), remaining missing values were imputed from a normal distribution (width: 0.3; down shift: 1.8). Two-sample *t* test was performed with a FDR < 0.01.

For global cell line comparison, triplicates were analyzed twice for a total of six single shot measurements per cell line (except the NSC-34 cell line, which was only measured once). For the principal component analysis (PCA) of the different cell lines, we furthermore limited the data set (LFQ intensities) to entries with a minimum of four valid values in at least one group of six replicates. Remaining missing values were imputed from a normal distribution (see above).

Protein intensity (summed peptide intensity) for cerebellum samples were divided by the molecular weight for ranking by proteins abundance. Annotations (GO molecular function, biological process, cellular component; KEGG and Uniprot Keywords) were matched to protein groups with Perseus. We performed a 1D annotation enrichment (41) on the normalized protein intensities. To evaluate the proteome coverage, we counted the occurrence of categories in our sample and compared it to the category count for the complete murine proteome in Perseus.

MS raw data and data for protein and peptide identification and quantification were submitted as supplementary tables to the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD001592.

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

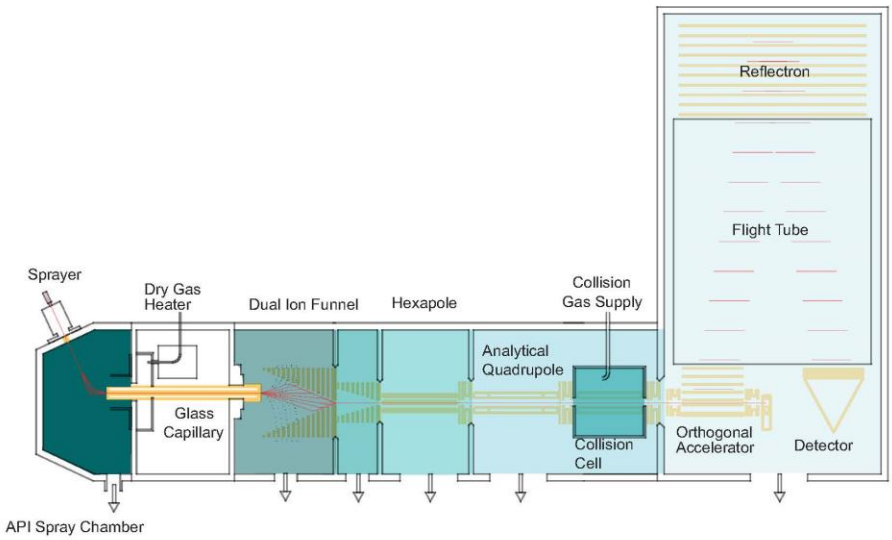


FIG. 1. **Schematic of the impact II mass spectrometer (not to scale).**

RESULTS AND DISCUSSION

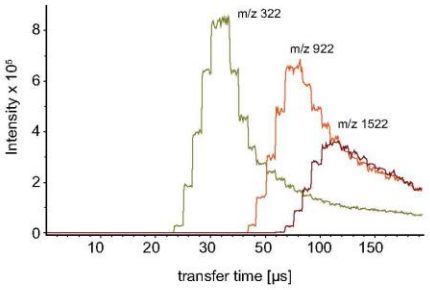
Overview of the Instrument—The Bruker impact II is a QTOF in a benchtop format, featuring several improvements in its design (Fig. 1). Briefly, ions are produced in the CaptiveSpray, which is in an encased nanoelectrospray source that features a well-defined gas stream to guide the ions into the vacuum via a capillary inlet. A double ion funnel, based on principles described by Smith and co-workers (42), is positioned off axis, which prevents neutrals from further transmission along the ion path. The pressure drops by several orders of magnitude from the capillary exit to the postfunnel stage (3 mbar to 3×10^{-4} mbar), while the ion current is virtually undiminished (see below). Additionally, the funnel allows for soft transfer based on low electrical field strength independent of the mass (typically 10 V/cm, much lower than in nozzle-skimmer designs). By introducing electrical acceleration in-between the two funnels, in-source fragmentation can still be achieved intentionally. There is a hexapole ion guide between funnel and the analytical quadrupole mass filter, which has a monolithic design based on high precision glass. Precursor ions can be isolated by this quadrupole for subsequent fragmentation in the collision cell. Intact ions or fragments can be stored and extracted from the collision cell and enter the orthogonal deflection region as a very narrowly focused ion beam ($< 500\ \mu m$). Here they are accelerated into a field-free drift region. A newly designed, two-stage reflectron further compensates the velocity distribution orthogonal to the beam direction. Finally, the ions impinge on an MCP detector coupled to a 10-bit, very high frequency (50 Gbit/s), zero noise digitizer. Data collection is coordinated by the Bruker Compass data system and in the experiments described here, post-acquisition data processing is performed in the MaxQuant environment.

Optimization of the Collision Cell—Efficient fragmentation of precursor ions on an LC-MS/MS time scale is a key for the identification of peptides in shotgun proteomics strategies. We optimized several aspects of the collision cell (supplemental Fig. S2*A*): Precise geometrical alignment allows focusing of the ions along the axis of the collision cell, directly translates into well-defined starting conditions for the orthogonal accelerator and is therefore mandatory for high mass resolution. This is implemented via a quadrupolar configuration of the collision cell device providing a narrow pseudo potential well (43). We also introduced a radial ejection step between any two MS or MS/MS experiments, in order to reduce the dead time. The ion path has to be emptied to avoid crosstalk between two consecutive spectra without introducing substantial ion losses. Most importantly, we optimized the time of ion fragmentation and extraction within the fragment spectra to ensure efficient high frequency MS/MS by implementing an electrical axial field gradient. This gradient directs the ions toward the exit of the collision cell, reducing the time it takes for the first ions to reach the extraction lens after quenching the collision cell. From here, they can be rapidly released toward the orthogonal accelerator, forming packages that match the orthogonal pulser frequency. The ion densities after 3 ms (upper red traces in supplemental Fig. S2*B*–S2*C*) reveals two important aspects: Using the axial field gradient results in comparable ion density at the collision cell exit in much shorter time than without, *i.e.* after about 1 ms instead of

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

A

	relative transmission			absolute transmission		
transfer time	m/z 322	m/z 922	m/z 1522	m/z 322	m/z 922	m/z 1522
optimum	100 %	100 %	100 %	8.34e5	6.51e5	3.57e5
100 µs	16.4 %	64.5 %	96.9 %	1.37e5	4.20e5	3.46e5



B

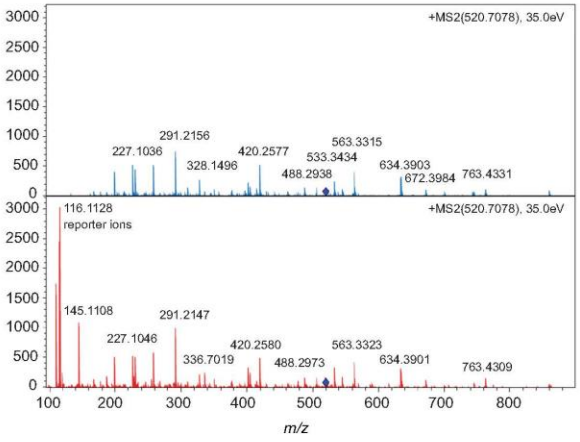


FIG. 2. ***A*, Absolute ion intensities of *m/z* 322, 922 and 1522 as function of the transfer time.** The maxima of the distributions for each *m/z* value are the optimal times for efficient transfer into the flight tube. *B*, Fragment spectrum of the iTRAQ labeled peptide LFTGHPETLEK without (blue) and with (red) transfer time stepping that adds the lower *m/z* range including the reporter ions without sacrificing intensity in the higher mass range.

3 ms, respectively. This suggests a three times faster ion transfer. Moreover, the overall number of ions within the collision cell is reduced to less than 50% at the moment of quenching, which should reduce ion losses related to quenching accordingly.

The pseudo potential calculations and simulation were confirmed in different experiments comparing the performance of the optimized collision cell on the fragmentation yield of Glu-Fibrino-Peptide B at different MS/MS acquisition rates (supplemental Fig. S2*D*). This revealed that the reduced quench losses in fact improve the number of ions detected at a spectra rate of 16 Hz by a factor of two. We further observed that the axial field gradient improves the stability of the system even in the presence of slight contaminations on the rods.

High Transfer Efficiency to the Orthogonal Acceleration Unit—The ions travel through the flight tube and require as much time as the largest *m/z* species needs to reach the detector, before the HV pulser can send the next ion package toward the detector (typically between 100 and 150 μs). To avoid excessive loss of ions, orthogonal TOF instruments are therefore often operated in a mode in which the ions are stored in the collision cell during the TOF scan and released in time for the next extraction pulse of the orthogonal accelerator. This would allow for 100% duty cycle if all ions were indeed transferred such that they arrive in the orthogonal accelerator at the same time and with the same kinetic energy. In practice, however, the extraction time from the collision cell toward the gate lens is a function of ion mobility. We have analyzed these combined effects by varying the time from opening the gate lens to the extraction pulse of the orthogonal accelerator (transfer time) for different *m/z* ratios (Fig. 2*A*). Simulations of ion trajectories and extraction times reveal that about 80% of a single ion species can be accelerated into the drift tube of the TOF analyzer under optimal transfer time conditions. In the impact II the high transfer efficiency is further optimized by reducing the distance between the trapping region and the orthogonal accelerator—it is about four times higher compared with conventional QTOF systems operated in continuous operation mode. Figure 2*A* summarizes the relative transmission efficiency of selected precursors (*m/z* = 322 Th, 922 Th and 1522 Th) at a transfer time of 100 μs, which is a standard setting to cover the mass range relevant in shotgun proteomics. The analysis highlights that ions with low *m/z* are compromised most, which we have previously counteracted in proteomics experiments by adding special extraction conditions (44). To tackle this problem in a more general way, Bruker introduced the "transfer time stepping" operation mode, where first high mobility species are extracted followed by species with a lower mobility. During the initially short opening times of the gate (typically 50% of the total scan) only the higher mobility, low *m/z* ions pass the gate lens, while the lower mobility ions are still accumulated in the collision cell. In the second transfer time steps, extraction times are increased to allow the low mobility, higher *m/z* ions to pass the gate lens. Spectra with and without transfer time stepping, reveal the beneficial effect on low mass ions without appreciable loss in the standard mass range (Fig. 2*B*). This is particularly beneficial for labeling experiments that rely on quantification of low mass reporter ions as shown in the figure. These ions are transferred with greater than 60% effi-

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

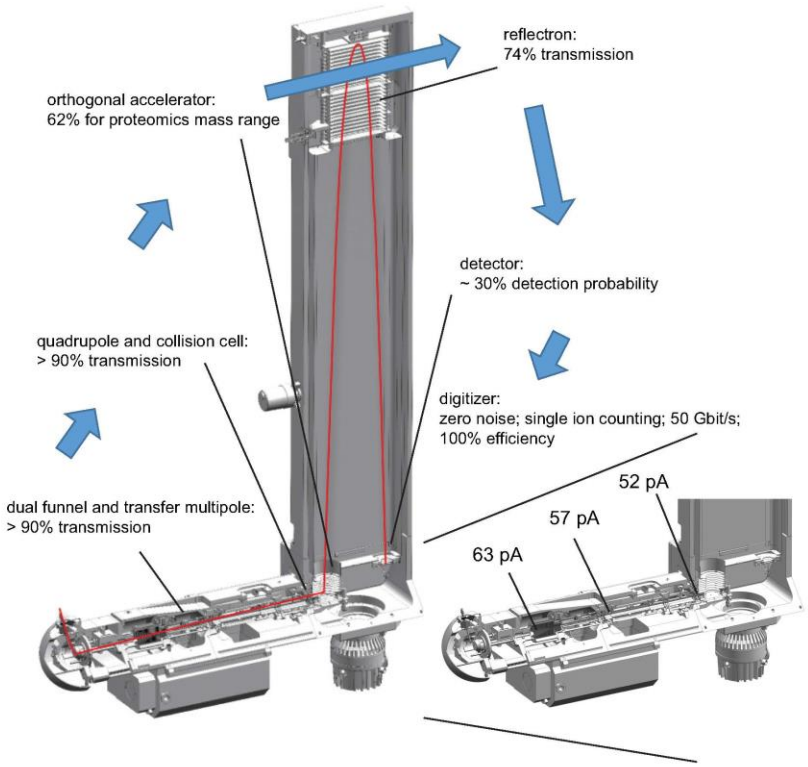


FIG. 3. **Ion transfer efficiency of mass range most relevant in shotgun-proteomics experiments (*m/z* 100–1500).** Inset shows net analyte ion currents at the indicated measurement points and transfer efficiencies at the orthogonal accelerator at various stages along the flight path of the impact II instrument.

ciency during the low *m/z* extraction phase and the total amount of ions in the analyzed fragment spectrum is increased by 58%.

Together, our developments led to an overall transmission efficiency of > 60% into the orthogonal acceleration unit. This compares favorably to a recent report, in which ion mobility was performed on fragment ions and their arrival times were synchronized with orthogonal extraction, which lead to an up to 10-fold improvement of detection sensitivities to standard operation (27).

Sensitivity and Ion Transfer—The number of ions that successfully pass through the instrument and are finally recorded determine a mass spectrometer's sensitivity. We were interested in the transfer efficiencies along the ion path, from entering the vacuum system to the detector. To experimentally determine this, we infused a 1 pmol/μl BSA solution or blank solution and measured the difference in the ion current between these conditions. When operating the outlet of the capillary and the funnel region as a Faraday cage, we measured a net ion current of 63 pA, which we defined as the starting value (100%) (Fig. 3). The large ion acceptance aperture of the first funnel efficiently captures in the ion flux leaving the capillary. It transfers the ions to the second funnel, which also passes the ions in an almost lossless manner through the next stage as evidenced by a net current reading of 57 pA (>90%) after the hexapole. Likewise, more than 90% of the ions are transferred through the quadrupole (in nonmass selecting operation) and the collision cell. More than 60% of these ions are transmitted into the flight tube by the orthogonal accelerator (see above). The reflectron contains two grids, which are each passed twice, leading to a geometrically defined overall transmission of 74%. Although all these ions hit the MCP, not all of them enter the channels and not all of them result in secondary electrons (detector quantum yield). However, when secondary ions are generated they are greatly amplified ($>10^6$ fold), and can be efficiently discriminated from electronic noise (zero noise detection). Together, this leads to an estimated detection probability of 30% for our MCP detector. Although not all of these measurements and estimates are very precise, they suggest an overall detection

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

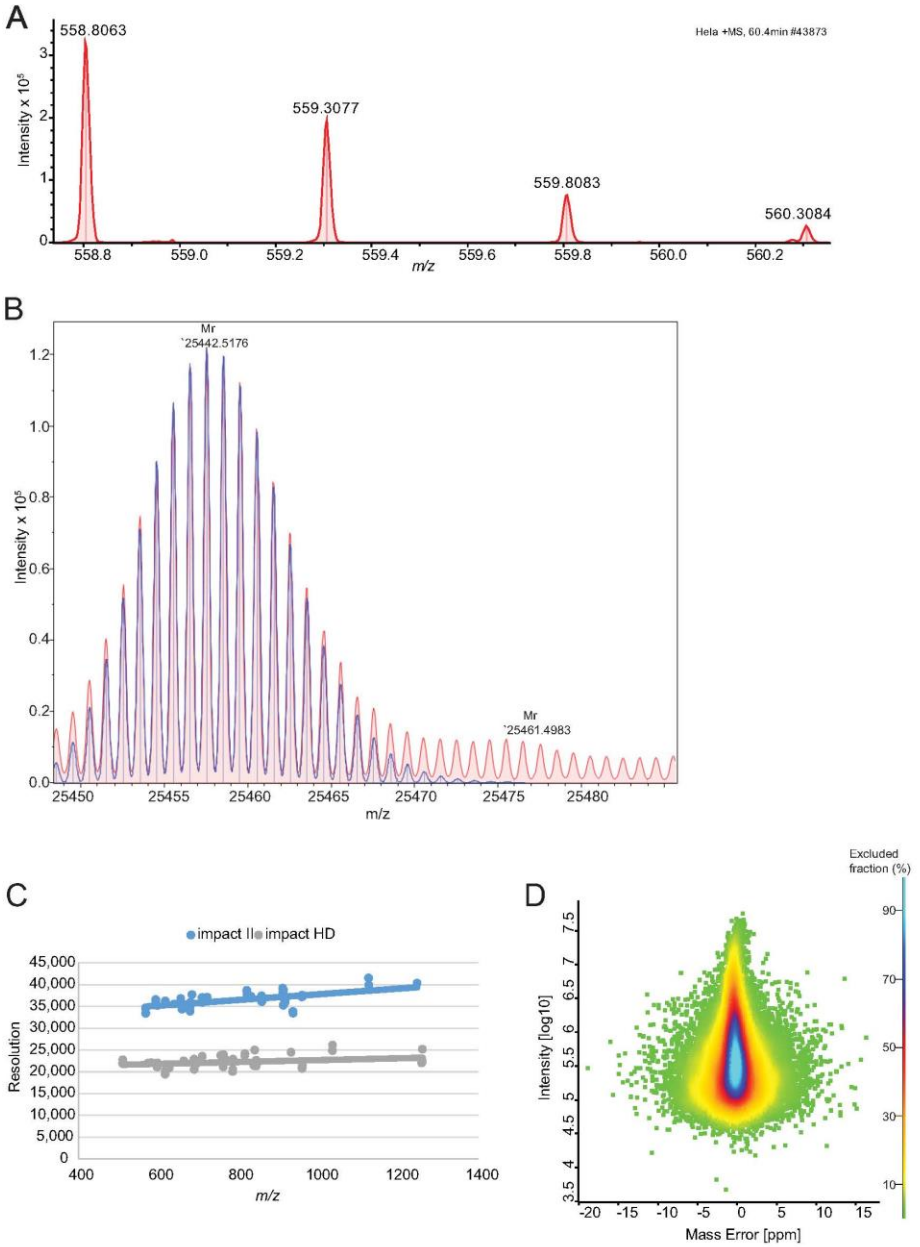


FIG. 4. **Resolution and mass accuracy of *A*, a peptide isotope cluster (*m/z* 558.8063, *r* = 33k) and *B*, Fd unit of Adalimumab (*m/z* 25442.5157, *r* = 63k, 0.26 ppm).** Overall improvement of the resolution with the improved detector, *C*, and the achieved mass accuracy dependent on the summed peptide intensity, *D*, in a shotgun proteomics experiment using the QTOF optimized version of MaxQuant.

TABLE I
Identification from HeLa and yeast lysate triplicate analysis using a standard 90 min gradient

	MS scans	Isotope pattern	MS/MS scans	Identification rate [%]	Peptide sequences identified	Proteins identified
HeLa_1	7002	759,774	79,704	47.09	35,547	4870
HeLa_2	7181	769,355	79,876	47.22	35,572	4864
HeLa_3	7272	796,086	81,389	46.57	35,621	4828
Total				46.96	48,172	5210
yeast_1	4873	528,682	63,682	31.08	17,066	3361
yeast_2	4732	541,194	66,441	30.88	16,921	3325
yeast_3	4691	556,675	66,978	31.98	17,494	3369
Total				31.32	24,131	3627

probability of ions transmitted into the vacuum system of about 10%. This excellent number is because of the fact that the continuous beam generated by the ES source can be utilized and that all the ion guiding elements have been successfully optimized for high transmission.

Resolution and Mass Accuracy—For the impact II several improvements were implemented: symmetrical shielding for better ion focusing; line grids to increase the transmission; low temperature coefficient ceramic spacers to decrease temperature related mass drift and improved axel bearings for precise alignment. Together this led to about 35% increased resolution over the full proteomics mass range. As this improvement is reached solely by better focusing of the ions, mass accuracy and signal to noise are expected to increase accordingly. Improvements to the MCP detector include an increased entrance aperture, higher electron accelerating fields and optimized shielding. Overall, these measures lead to 2-fold faster ion impact transients and 30% higher detection efficiency of the MCP.

In summary, the resolving power of the TOF analyzer is expected to increase by about 70 to 80% by the introduction of the new collision cell, reflectron and detector. To test this experimentally on a standard proteomic sample, we analyzed data from a HeLa digest. Resolution for typical peptides is in excess of 33,000 as illustrated by an example in Fig. 4*A*. Further increase in the resolution can be obtained by the "Focus mode," which involves real time processing and alignment of successive pulses and increases accuracy of flight time determination, when multiple ions of one species reach the detector at the same time (45). This can be helpful to resolve the isotope distributions of proteins, as shown in Fig. 4*B*, which features a resolution in excess of 60,000 for an antibody subunit (> 25 kDa). In TOF measurements, resolution tends to increase with *m/z* (Fig. 4*C*), and reached more than 40,000 for the TuneMix component at *m/z* 1222. This constitutes a 70 to 80% improvement over the previous impact model, the impact HD.

This increase in resolving power should also imply better mass accuracy in proteomics samples, which we tested using the software package MaxQuant, which we adapted to QTOF data as described under "Experimental Procedures." A special feature of MaxQuant is the extraction of individual mass accuracy values (46, 47), which allows to make efficient use of high mass accuracy in the identification of peptides. MaxQuant was originally developed on the basis of data from hybrid Orbitrap instruments. Here we developed MaxQuant further in order to analyze QTOF data and also in this context profit from the high mass accuracy provided by nonlinear mass recalibration algorithms. A special challenge in QTOF data is the drift in the mass scale because of thermal expansion caused by ambient temperature drift. MaxQuant employs a double search strategy, in effect supplying hundreds of reference masses internal to each proteomic sample. This feature efficiently removes any effect of the temperature related mass drift. This "software lock mass" feature makes it unnecessary to use dedicated molecular species for the calibration of spectra (48). With these developments in MaxQuant, we analyzed the peptide mass error distributions over a 90 min gradient run (Fig. 4*D*). This showed an average absolute mass deviation of around 1.45 ppm, which is excellent for a QTOF instrument.

Impact II Performance for Single Shot Analysis—To investigate the performance of the impact II for shotgun proteomics, we first analyzed a complex peptides mixture derived from a mammalian cell line in the single-run format (Experimental Procedures). We separated 1 μg of peptide digest by on-line HPLC with the standard 90 min gradient employed in our laboratory and performed triplicate analysis. The typical data dependent acquisition scheme in bottom-up proteomics consists of an MS scan followed by N MS/MS fragment scans of the most intense precursors (topN method). It is desirable to choose N such that the total cycle is less than a few seconds. For our measurements we aimed at a duty cycle of around 1.3 s and designed a top17 method, consisting of 200 ms for MS acquisition and a MS/MS integration time adapted to the precursor intensity. We found this to be a good balance between acquiring high S/N in the MS and achieving optimal ion intensities in the MS/MS spectra for the HeLa digest. This method reached more than 7000 MS and 79,700 MS/MS scans. In each run MaxQuant identified on average 35,580 unique peptide sequences, which results in total of 48,172 unique peptide sequences in the triplicate analysis

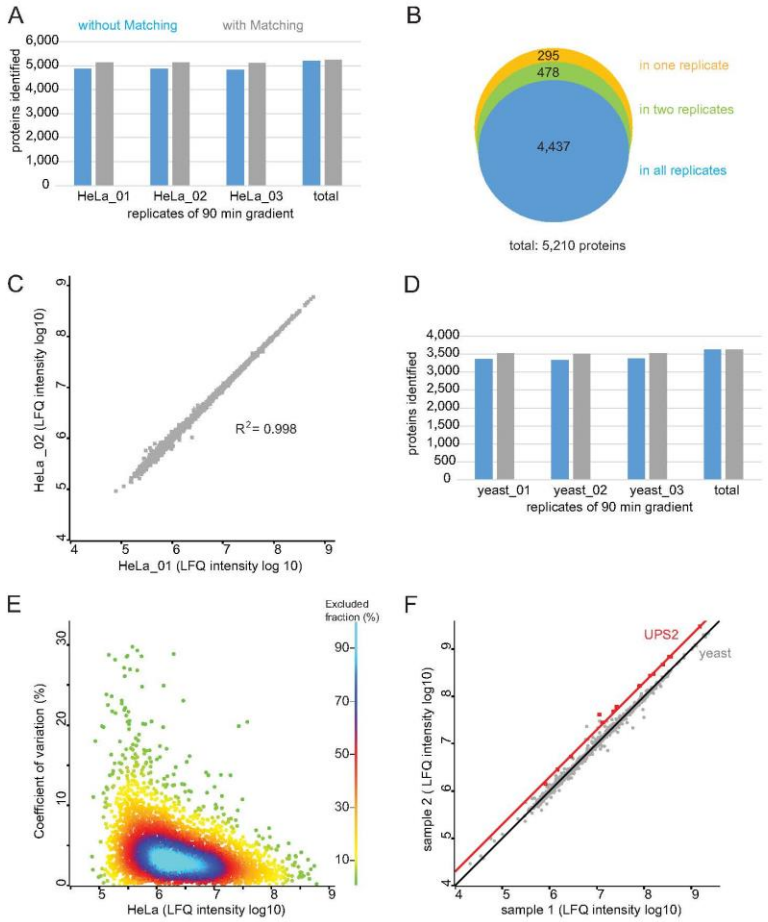


FIG. 5. **Triplicate analysis of HeLa and yeast digest using a 90 min gradient for single shot analysis.** *A*, Protein identification numbers of each replicate of 1 μg HeLa digest and *B*, the overlap of protein identities. *C*, Correlation between the rank ordered label-free quantification values of each identified protein in replicate1 (HeLa_01) and replicate2 (HeLa_02) (log10 LFQ intensities). *D*, Protein identification number of 1 μg yeast digest. *E*, Reproducibility across triplicate analysis of HeLa digest. CVs of all three replicates, representing 99% of the data points. *F*, Accuracy in quantification. Results of spike-in experiment showing the UPS-2 standard (orange) and in a yeast proteome background (gray). Yeast lysate was present in equal amounts in sample 1 and sample 2 and the UPS2 protein standard was present in twice the amount in sample 2.

(Table I). These peptides mapped to an average of 4854 proteins per run, and a total of 5210 proteins of the HeLa proteome with the three 90 min gradients (Fig. 5*A*), indicating that a deep coverage can be achieved using relatively short, single shot analysis. Transferring identification between the runs based on their mass precision and retention time ("match between runs" feature in MaxQuant) led to around 5100 proteins identifications per single run. Comparing protein identities between the triplicate analyses (without "matching between runs"), we observed that more than 90% of proteins were identified in each of them (Fig. 5*B*). This indicates high reproducibility and a minimal 'missing value' problem. This conclusion is further supported by excellent reproducibility (R^2 = 0.998) in the label-free intensities determined in pairwise comparison between runs (MaxQuant LFQ values (39, 49)) (Fig. 5*C*). We also determined the number of proteins identified in a single shot run of 1 μg yeast digest. On average, we identified 3352 proteins per single, 90 min gradient, and a total of 3627 proteins when combining the three single shot measurements (Fig. 5*D*).

Reproducibility and Accuracy of Quantification—To evaluate the reproducibility of the method for label-free quantifica-

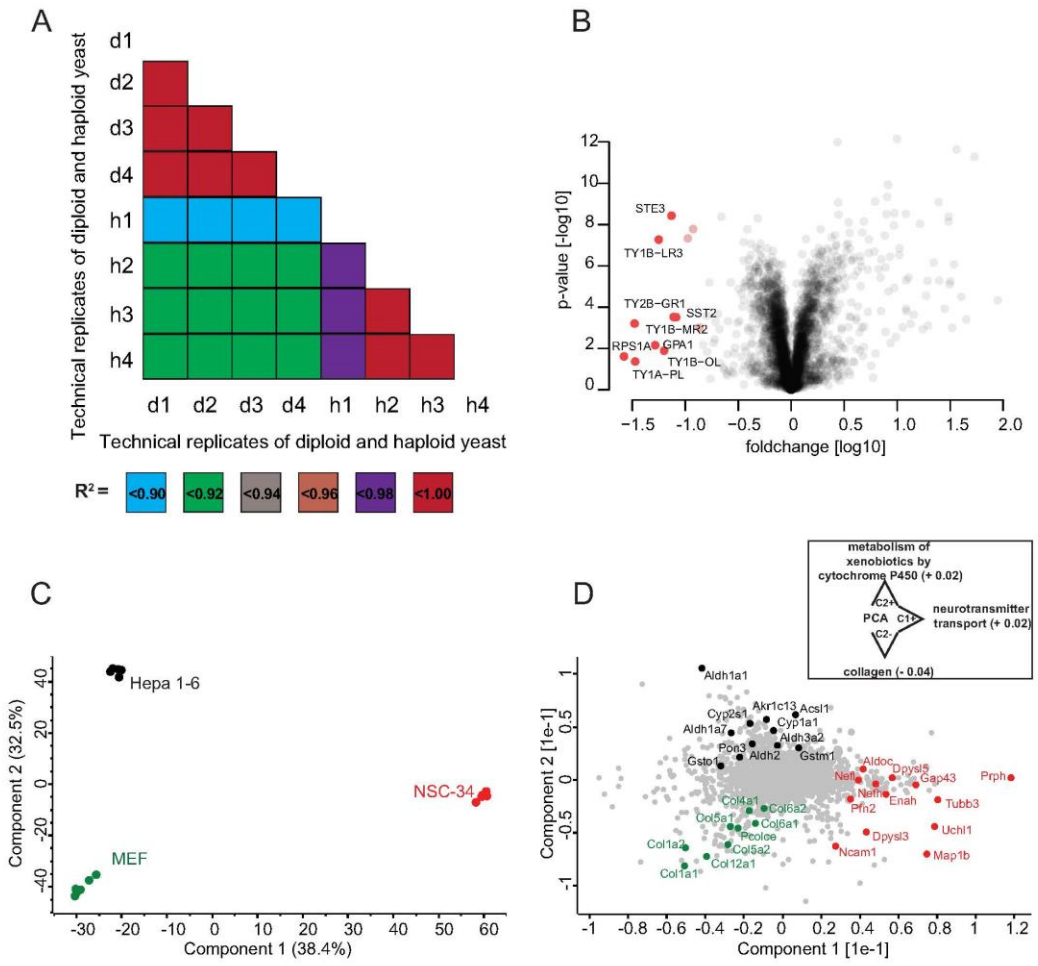


FIG. 6. ***A*, Correlation of single shot LC-MS/MS measurements of haploid (h) and diploid (d) yeast samples.** All technical replicate correlation values were > 0.98. *B*, Quantitative differences between the haploid and diploid yeast proteome. Proteins marked in red are significantly more abundant in haploid cells. *C*, Principal component analysis (PCA) of LFQ protein expression levels in mouse hepatoma (Hepa 1–6), mouse embryonic fibroblasts (MEF) and motoneuronal cell lines (NSC-34). *D*, Loading of the first two principal components as obtained from the PCA plotted in *C*. The inset indicates significantly enriched annotations along the PC axes (FDR < 0.05). The respective enrichment scores are specified in parentheses.

tion, we determined the coefficients of variation (CV) of the label-free intensities, determined in pair-wise comparison between three technical HeLa replicates (see above). For more than 90% of the quantified proteins the CV was smaller than 10% (Fig. 5*E*). For the lowest intensity quantile the median was 0.05 and for the highest one 0.03 (Fig. 5*E*, supplemental Fig. S3). This shows reproducible quantification over four orders of magnitude.

Accurate quantification of differently expressed proteins remains challenging over a wide concentration range and benefits from a very stable analytical platform. QTOF instrumentation has been used widely for label-free quantification, which in principle allows convenient analysis and comparison of an arbitrary number of samples. For evaluation of the label-free quantitative capabilities of the impact II platform in complex mixtures we wished to use a reference sample set with known ratios for a small subset of proteins. We spiked the Universal Protein Standard 2 (UPS-2), consisting of 48 proteins covering a dynamic range of five orders of magnitude, in two different concentrations into the yeast proteome.

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

This generated two samples, in which the yeast peptides should be 1:1 whereas ratios for UPS-2 peptides should be at 1:2. To increase the number of identified peptides we then used the equimolar UPS-1, which we also spiked in a yeast background. This allowed transfer of peptide identifications to unsequenced peptides in the UPS-2 in yeast runs via the 'match between runs' algorithm in MaxQuant. We identified all of the 48 human UPS proteins in the sample containing yeast with the equimolar UPS-1 standard. Of these proteins, 23 were identified in each one of the 12 single shot yeast measurements with the dynamic UPS-2 standard either directly via MS/MS or via match between runs and 18 of these proteins showed Welch-significance (Fig. 5*F*). We identified and quantified UPS-2 proteins over more than three orders of magnitude in these relatively fast measurements (90 min gradients). The UPS-2 proteins have an average fold change of 0.49 (± 0.06), which is close to the theoretical ratio of 0.5.

Quantification of Changes in the Yeast Proteome—To test the workflow in a systems biology context, we analyzed proteome changes of diploid and haploid (Matα cell) *S. cerevisiae*. We analyzed 2 μg of yeast digest from haploid and diploid cells in technical quadruplicates with our standard 90 min gradient. This identified 3769 proteins using "match between runs." For statistical analysis, we only considered LFQ intensities that were detected in at least three replicates of the haploid or diploid groups. After filtering, 3222 proteins remained for further analysis (Experimental Procedures). Remaining missing values were imputed from a normal distribution. The technical replicates correlated much more with each other than they correlated to the other genotype (R^2 greater 0.98 *versus* R^2 about 0.92; Fig. 6*A*). As in our previous large-scale analysis on SILAC labeled haploid and diploid yeast (50), we found transposons more abundant in haploid cells than in diploid cells (Fig. 6*B*). Ste3, the pheromone a factor receptor, was specific to haploid yeast, as expected from its mating status. Also absent in diploid but present in haploid cells were Sst2, a GTPase-activating protein for Gpa1 (51), which, consistently, showed higher expression in haploid cells. Conversely, Sps100, which is a sporulation-specific wall maturation protein turned out to be specific for diploid cells. Doing such systems-wide comparisons by traditional methods would have required thousands of individual Western blots. Even compared with our previous large-scale study performed by quantitative MS (50), we here used less than 1% of yeast input material and measurement time. This illustrates the rapidity by which MS-based proteomics is becoming a viable method for answering biological questions.

Global Proteomic Comparison of Different Cell Lines—As a second example of typical proteomics experiments, we applied the QTOF-based workflow to the characterization of common cellular disease model systems. For this purpose, we compared the proteomes of spinal cord neuron-neuroblastoma (NSC-34), mouse hepatoma (Hepa 1–6) and mouse embryonic fibroblast (MEF) cell lines in a quantitative manner. All cell lysates were analyzed in single shots using 90 min gradients and subsequently quantified using the MaxQuant label-free quantification algorithm. The observed LFQ intensities were highly reproducible between biological and technical replicates as indicated by Pearson correlations coefficients $>$ 0.97 (supplemental Fig. S4).

After stringent filtering (Experimental Procedures) we performed a principal component analysis (PCA) to evaluate the similarities and dissimilarities of the cell lines on a global scale. Replicates from a single cell line clustered very tightly in the PCA space and the first two principal components accounted for 71% of the cumulative variance within our data set (Fig. 6*C*). Interestingly, the motoneuronal cell line NSC-34 was clearly separated from Hepa 1–6 and MEF cells in the first principle component, whereas the variance between both non-neuronal cell lines was described by the second, orthogonal, principal component. In the latter component, NSC-34 is positioned half-way between Hepa 1–6 and MEF.

To assess individual proteins that are the main drivers for the separation between the three cell lines, we plotted the loadings of the first two principal components (Fig. 6*D*). Hepa 1–6 cells were characterized by gene products involved in regulatory and metabolic processes. As we had found before (52), proteins related to the glycolysis pathway, such as the aldehyde dehydrogenase (Aldh) family, were highly represented in Hepa 1–6. Not surprisingly, proteins driving the separation of MEF were predominantly linked to collagen synthesis, such as the COL gene family and the precollagen C-endopeptidase enhancer 1 (Pcolce). Proteins that differentiate along principal component 1 encompassed various components of the cytoskeleton axons, including Prph and the heavy and light chains of neurofilaments (Nefh and Nefl). In addition, gene products involved in axon guidance (Enah, Dypsl5 and Tubb3) and neuron projection (Uchl1, Gap43) were highly distinctive for NSC-34. Proteins separating NSC-34 from the other cell lines were significantly enriched for neurotransmitter transport while we observed enrichment for collagens and xenobiotic processes for the MEF and Hepa 1–6 cells on component 2, respectively, further showing how proteomics can highlight biological function. This is even more remarkable, given that we previously found that the proteomes of motoneuronal cell lines, including NSC-34, lack distinctive neuronal characteristics, as several key actors in axon growth and guidance were either depleted or low abundant (28). As a result, we had placed motoneuronal cell lines only halfway between *in vivo* motoneurons and non-neuronal controls. Nevertheless, our label-free QTOF-based workflow is very well suited to differentiate subtle alterations in biological systems in a short time of analysis.

Impact II Performance for Deep Proteome Analysis of a Cell Line—To evaluate the impact II for deep proteome coverage we performed high pH reverse-phase pre-fractionation with fraction concatenation as described in ref (31). We loaded 175 μg of a HeLa peptide mixture, collected 72 fractions and

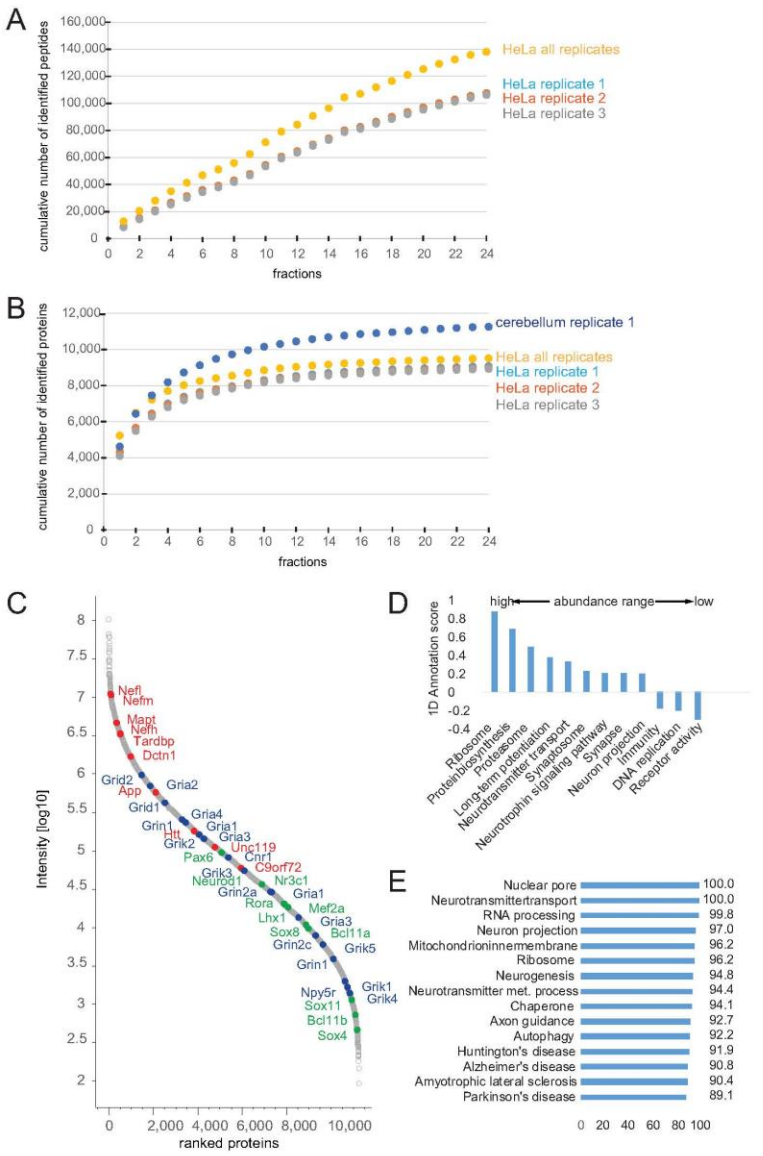


FIG. 7. **Triplicate cell line and singlet tissue analysis in 24 high-pH fractions.** *A*, Cumulative number of identified peptides of triplicate analysis of HeLa. *B*, Cumulative number of identified proteins of HeLa replicates and of the singlet analysis of mouse cerebellum (48 h total measurement time). *C*, Rank ordered intensities of each identified protein (log10 intensities) of cerebellum. Proteins involved in neurodegenerative disorders are marked in red; cerebellum specific expressed proteins are marked in blue and a selection of receptors in green. *D*, Protein abundance distribution and *E*, proteome coverage (expressed as a percentage) of different neuronal, metabolic and disease relevant processes in cerebellum.

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

combined them into 24 fractions (Experimental Procedures). These fractions were analyzed in technical triplicates using a standard 90 min LC-MS/MS gradient (total measurement time of 48 h per replicate). Instead of the simple Top17 method described above, we here used a so-called "dynamic method" with a fixed cycle time and a MS/MS integration time adapted to the precursor intensity. We found that this method helped to generate high quality MS/MS spectra also for low abundant peptides.

Figure 7*A* depicts the cumulative number of unique peptides identified as a function of the number of fractions analyzed. The increase is nearly linear, indicating a small overlap of peptide identification between fractions and the good orthogonal separation power as also observed by others using high-pH fractionation (49). The figure also indicates that the reproducibility was very high between the technical replicates. On average we identified 107,038 unique peptides, mapping to around 8995 proteins (protein FDR 1%; Fig. 7*B*). In total we identified 138,086 unique peptides, resulting in 9515 different protein groups. This is to our knowledge the deepest proteome coverage of a human cell line measured with a QTOF instrument.

Impact II Performance for Ultra-Deep Proteome Analysis of Tissue—Tissues are more challenging to analyze by proteomics than cell lines, because they are made up of different cell types, the extracellular matrix and other structural and connective elements. Among the different tissues, the brain is thought to be the most complex one. To evaluate the Impact II in this context, we homogenized the cerebellum of a single mouse, digested it using our standard workflow and separated part of the resulting tryptic peptides using high pH fractionation (Experimental Procedures). In total, we identified 11,257 proteins from a single analysis of 24 fractions (2 days of measurement time) (Fig. 7*B*, supplemental Table S1). To our knowledge, this is the deepest proteome measurement of any tissue reported by TOF instrumentation so far. Protein abundance, as indicated by the summed and normalized peptide signal varied by more than five orders of magnitudes (Fig. 7*C*). We identified many high to medium abundant proteins involved in neurodegenerative disorders (Fig. 7*C* marked in red). The transcription factors and DNA binders Pax6, Lhx1/5, Otx1/2, and Neurod1 (Fig. 7*C*, marked in blue) are examples of proteins that have been reported to be specifically expressed in cerebellum (53). Like various neuronal receptors, transcription factors populate the medium to low abundance range (Fig. 7*C*, marked in green). The distribution of molecular functions throughout the abundance range is similar to that observed in previous studies (Fig. 7*D*, supplemental Table S2) (54, 55). The very high depth of this cerebellum proteome is shown by almost complete coverage for neuronal, general metabolic and disease relevant processes (Fig. 7*E*, supplemental Table S2). Rapid estimation of protein abundances in the brain regions of a single mouse could be useful for studying tissue characteristics and disease specific alterations. For instance, knowledge about changes in the complete proteome would provide an additional layer of information on the pathological processes in neurodegeneration.

CONCLUSION

Here we have described the construction and performance of a state of the art QTOF instrument, the impact II. We documented significant improvements in the ion path, collision cell performance and, in particular, in the performance of the reflectron and detector. The latter allow a mass resolution and mass accuracy that is compatible with the high demands of shotgun proteomics of complex mixtures. For the first time, we measured and modeled the ion transmission from capillary to the detector, which revealed an excellent efficiency of about 10%. The new features of the impact II allow deep characterization in single shots, where we identified more than 5200 proteins in HeLa cells and 3600 proteins in yeast. Using off-line high pH reversed-phase fractionation we identified more than 9500 proteins in HeLa cells and 11,250 proteins in a single cerebellum tissue analysis. These are extremely high numbers for any platform and additional method developments should further improve these results. We also document excellent quantitative reproducibility and accuracy in a label-free format. In concordance with others (26, 27), we conclude that the improvements in QTOF technologies in recent years now clearly enable demanding, in-depth analysis of very complex proteomes.

Acknowledgments—We thank our colleagues from Bruker, in particular Christian Cyriacks, Christoph Gebhardt, Stefan Harsdorf, Wolfgang Jabs and Anja Wiechmann. At the Max-Planck Institute of Biochemistry, we thank Nagarjuna Nagaraj, Korbinian Mayr, Richard Scheltema and Gaby Sowa for technical support and Kirti Sharma, Garwin Pichler and Marco Hein for fruitful discussions.

* Partial support for this work was provided by the Max-Planck Society for the Advancement of Science and the Koerber foundation.

[S] This article contains supplemental Figs. S1 to S4 and Tables S1 and S2.

‖ To whom correspondence should be addressed: Dept. Proteomics and Signal Transduction, Max-Planck Institute of Biochemistry, Am Klopferspitz 18, Martinsried (near Munich) D-82152, Germany. Tel.: 49-89-8578 2557; Fax: 49-89-8578 2219; E-mail: mmann@biochem.mpg.de.

REFERENCES

1. Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60,** 2299–2301
2. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246,** 64–71
3. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422,** 198–207
4. Cravatt, B. F., Simon, G. M., and Yates, J. R., 3rd (2007) The biological impact of mass-spectrometry-based proteomics. *Nature* **450,** 991–1000
5. Altelaar, A. F., Munoz, J., and Heck, A. J. (2013) Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat. Rev. Genet.* **14,** 35–48
6. Richards, A. L., Merrill, A. E., and Coon, J. J. (2014) Proteome sequencing goes deep. *Curr. Opin. Chem. Biol.* **24C,** 11–17
7. Domon, B., and Aebersold, R. (2006) Challenges and opportunities in

proteomics data analysis. *Mol. Cell. Proteomics* **5,** 1921–1926
8. Mann, M., and Kelleher, N. L. (2008) Precision proteomics: the case for high-resolution and high mass-accuracy. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 18132–18138
9. Makarov, A. (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal. Chem.* **72,** 1156–1162
10. Zubarev, R. A., and Makarov, A. (2013) Orbitrap mass spectrometry. *Anal. Chem.* **85,** 5288–5296
11. Michalski, A., Damoc, E., Hauschild, J. P., Lange, O., Wieghaus, A., Makarov, A., Nagaraj, N., Cox, J., Mann, M., and Horning, S. (2011) Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol. Cell. Proteomics* **10,** M111 011015
12. Scheltema, R. A., Hauschild, J. P., Lange, O., Hornburg, D., Denisov, E., Damoc, E., Kuehn, A., Makarov, A., and Mann, M. (2014) The Q exactive HF, a benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. *Mol. Cell. Proteomics* **13,** 3698–3708
13. Kelstrup, C. D., Jersie-Christensen, R. R., Batth, T. S., Arrey, T. N., Kuehn, A., Kellmann, M., and Olsen, J. V. (2014) Rapid and deep proteomes by faster sequencing on a benchtop quadrupole ultra-high-field Orbitrap mass spectrometer. *J. Proteome Res.* **13(12),** 6187–6195
14. Zhou, H., Ning, Z., Wang, F., Seebun, D., and Figeys, D. (2011) Proteomic reactors and their applications in biology. *FEBS J.* **278,** 3796–3806
15. Kocher, T., Swart, R., and Mechtler, K. (2011) Ultra-high-pressure RPLC hyphenated to an LTQ-Orbitrap Velos reveals a linear relation between peak capacity and number of identified peptides. *Anal. Chem.* **83,** 2699–2704
16. Cox, J., and Mann, M. (2011) Quantitative, high-resolution proteomics for data-driven systems biology. *Annu. Rev. Biochem.* **80,** 273–299
17. Nagaraj, N., Kulak, N. A., Cox, J., Neuhauser, N., Mayr, K., Hoerning, O., Vorm, O., and Mann, M. (2012) System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Mol. Cell. Proteomics* **11,** M111 013722
18. Hebert, A. S., Richards, A. L., Bailey, D. J., Ulbrich, A., Coughlin, E. E., Westphall, M. S., and Coon, J. J. (2014) The one hour yeast proteome. *Mol. Cell. Proteomics* **13,** 339–347
19. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014) Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11,** 319–324
20. Michalski, A., Cox, J., and Mann, M. (2011) More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J. Proteome Res.* **10,** 1785–1793
21. Savitski, M. M., Nielsen, M. L., and Zubarev, R. A. (2006) ModifiComb, a new proteomic tool for mapping substoichiometric post-translational modifications, finding novel types of modifications, and fingerprinting complex protein mixtures. *Mol. Cell. Proteomics* **5,** 935–948
22. Morris, H. R., Paxton, T., Dell, A., Langhorne, J., Berg, M., Bordoli, R. S., Hoyes, J., and Bateman, R. H. (1996) High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Commun. Mass Spectr.* **10,** 889–896
23. Silva, J. C., Denny, R., Dorschel, C., Gorenstein, M. V., Li, G. Z., Richardson, K., Wall, D., and Geromanos, S. J. (2006) Simultaneous qualitative and quantitative analysis of the Escherichia coli proteome: a sweet tale. *Mol. Cell. Proteomics* **5,** 589–607
24. Silva, J. C., Gorenstein, M. V., Li, G. Z., Vissers, J. P., and Geromanos, S. J. (2006) Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol. Cell. Proteomics* **5,** 144–156
25. Gillet, L. C., Navarro, P., Tate, S., Rost, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* **11,** O111 016717
26. Distler, U., Kuharev, J., Navarro, P., Levin, Y., Schild, H., and Tenzer, S. (2014) Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **11,** 167–170
27. Helm, D., Vissers, J. P., Hughes, C. J., Hahne, H., Ruprecht, B., Pachl, F., Grzyb, A., Richardson, K., Wildgoose, J., Maier, S. K., Marx, H., Wilhelm, M., Becher, I., Lemeer, S., Bantscheff, M., Langridge, J. I., and Kuster, B. (2014) Ion mobility tandem mass spectrometry enhances performance of bottom-up proteomics. *Mol. Cell. Proteomics*, **13,** 3709–3715
28. Hornburg, D., Drepper, C., Butter, F., Meissner, F., Sendtner, M., and Mann, M. (2014) Deep proteomic evaluation of primary and cell line motoneuron disease models delineates major differences in neuronal characteristics. *Mol. Cell. Proteomics* **13,** 3410–3420
29. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, prefractionation, and storage of peptides for proteomics using StageTips. *Nature Protoc.* **2,** 1896–1906
30. Wang, H., Qian, W. J., Mottaz, H. M., Clauss, T. R., Anderson, D. J., Moore, R. J., Camp, D. G., 2nd, Khan, A. H., Sforza, D. M., Pallavicini, M., Smith, D. J., and Smith, R. D. (2005) Development and evaluation of a micro- and nanoscale proteomic sample preparation method. *J. Proteome Res.* **4,** 2397–2403
31. Dwivedi, R. C., Spicer, V., Harder, M., Antonovici, M., Ens, W., Standing, K. G., Wilkins, J. A., and Krokhin, O. V. (2008) Practical implementation of 2D HPLC scheme with accurate peptide retention prediction in both dimensions for high-throughput bottom-up proteomics. *Anal. Chem.* **80,** 7036–7042
32. Whitehouse, C. M., Dreyer, R. N., Yamashita, M., and Fenn, J. B. (1985) Electrospray interface for liquid chromatographs and mass spectrometers. *Anal. Chem.* **57,** 675–679
33. Franzen, J. (1998) Method and device for transport of ions in gas through a capillary. Google Patents
34. Ishihama, Y., Rappsilber, J., Andersen, J. S., and Mann, M. (2002) Micro-columns with self-assembled particle frits for proteomics. *J. Chromatogr. A* **979,** 233–239
35. Ayoub, D., Jabs, W., Resemann, A., Evers, W., Evans, C., Main, L., Baessmann, C., Wagner-Rousset, E., Suckau, D., and Beck, A. (2013) Correct primary structure assessment and extensive glyco-profiling of cetuximab by a combination of intact, middle-up, middle-down, and bottom-up ESI and MALDI mass spectrometry techniques. *mAbs* **5,** 699–710
36. Senko, M. W., Beu, S. C., and McLaffertycor, F. W. (1995) Determination of monoisotopic masses and ion populations for large biomolecules from resolved isotopic distributions. *J. Am. Soc. Mass Spectr.* **6,** 229–233
37. Tsybin, Y. O., Fornelli, L., Stoermer, C., Luebeck, M., Parra, J., Nallet, S., Wurm, F. M., and Hartmer, R. (2011) Structural analysis of intact monoclonal antibodies by electron transfer dissociation mass spectrometry. *Anal. Chem.* **83,** 8919–8927
38. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10,** 1794–1805
39. Cox, J., Hein, M. Y., Luber, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13,** 2513–2526
40. R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/
41. Cox, J., and Mann, M. (2012) 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC Bioinformatics* **13,** S12
42. Shaffer, S. A., Prior, D. C., Anderson, G. A., Udseth, H. R., and Smith, R. D. (1998) An ion funnel interface for improved ion focusing and sensitivity using electrospray ionization mass spectrometry. *Anal. Chem.* **70,** 4111–4119
43. Gerlich, D. (1992) *Inhomogeneous RF fields a versatile tool for the study of processes with slow ions.*, John Wiley & Sons, Inc, Hoboken, NJ, U.S.A.
44. Steen, H., Kuster, B., Fernandez, M., Pandey, A., and Mann, M. (2001) Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. *Anal. Chem.* **73,** 1440–1448
45. Räther, O. (2005) High resolution detection for time-of-flight mass spectrometers. US6870156 B2 patent
46. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnol.* **26,** 1367–1372
47. Cox, J., Hubner, N. C., and Mann, M. (2008) How much peptide sequence information is contained in ion trap tandem mass spectra? *J. Am. Soc.*

Mass Spectr. **19,** 1813–1820
48. Cox, J., Michalski, A., and Mann, M. (2011) Software lock mass by two-dimensional minimization of peptide mass errors. *J. Am. Soc. Mass Spectr.* **22,** 1373–1380
49. Wang, Y., Yang, F., Gritsenko, M. A., Wang, Y., Clauss, T., Liu, T., Shen, Y., Monroe, M. E., Lopez-Ferrer, D., Reno, T., Moore, R. J., Klemke, R. L., Camp, D. G., 2nd, and Smith, R. D. (2011) Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics* **11,** 2019–2026
50. de Godoy, L. M., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Frohlich, F., Walther, T. C., and Mann, M. (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455,** 1251–1254
51. Apanovitch, D. M., Slep, K. C., Sigler, P. B., and Dohlman, H. G. (1998) Sst2 is a GTPase-activating protein for Gpa1: purification and characterization of a cognate RGS-Galpha protein pair in yeast. *Biochemistry* **37,** 4815–4822
52. Pan, C., Kumar, C., Bohl, S., Klingmueller, U., and Mann, M. (2009) Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. *Mol. Cell. Proteomics* **8,** 443–450
53. Suzuki, H., Okunishi, R., Hashizume, W., Katayama, S., Ninomiya, N., Osato, N., Sato, K., Nakamura, M., Iida, J., Kanamori, M., and Hayashizaki, Y. (2004) Identification of region-specific transcription factor genes in the adult mouse brain by medium-scale real-time RT-PCR. *FEBS Letters* **573,** 214–218
54. Geiger, T., Wehner, A., Schaab, C., Cox, J., and Mann, M. (2012) Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol. Cell. Proteomics* **11,** M111 014050
55. Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature* **473,** 337–342

MCP MOLECULAR & CELLULAR PROTEOMICS ASBMB

2.2. Increase in sequencing speed and sensitivity by using Parallel Accumulation – Serial Fragmentation on a TIMS-QTOF instrument

Florian Meier*, **Scarlet Beck***, Niklas Grassl, Markus Lubeck, Melvin A. Park, Oliver Raether, and Matthias Mann

*contributed equally

Parallel Accumulation – Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device

Journal of Proteome Research, 2015 December; 14(12)

The major bottleneck of the data-dependent acquisition mode is that a large number of peptides elute from the column at once, but peptides are selected and fragmented one by one. Over the last decades improvements in mass spectrometry have tried to address these challenges by an increase in sequencing speed and higher resolution. Nevertheless, only a limited number of peptides can be targeted by the mass spectrometer and even less are identified. Our group recently showed that under standard conditions only 16% of the eluted peptide features were targeted for fragmentation [114].

In 2011, trapped ion mobility spectrometry (TIMS) was introduced by Park and co-workers at Bruker, which can easily be combined with mass spectrometry. As described in the introduction, the TIMS device consists of pairs of stacked electrodes to which an RF and DC field are applied. A gas flow drags the ions along as they enter the TIMS device through the tunnel until they experience an equal counteracting force due to an electric field. Depending on their size and shape, the ions rest at a specific position. After a pre-defined accumulation time, stored ions are released from the TIMS tunnel by decreasing the electric field strength. Ions with larger cross sections elute before the more compact ones.

In this second project, Florian Meier and I teamed up with scientists from Bruker to develop a method on a TIMS-MS instrument that can drastically increase the sequencing speed and/or the sensitivity. Inspired by the adjustable elution profile of ions from the TIMS device, we were interested to see if it was possible to use multiple precursors in one TIMS scan (or TIMS-MS cycle) to multiply sequencing speed without losing sensitivity. For this purpose we first evaluated the necessary hardware requirements. Rise and lag times of the power supplies could be adjusted

so that the quadrupole switched on the sub-millisecond time scale. This means that the quadrupole could in principle target 12-20 precursors in the proteomic mass when using a TIMS scan time of 50 ms. The major bottleneck on our prototype was the time needed to calculate the quadrupole set values and the application of those values by the instrument controller. In the future this limitation could be circumvented by using a real time field-programmable gate array.

After demonstrating that sub-millisecond switching times of the quadrupole are possible, we investigated an ESI-TIMS-MS analysis by direct infusion of a peptide mixture composed of alcohol dehydrogenase (ADH), BSA, enolase and phosphorylase b. Instead of one precursor, four precursors could be targeted in a TIMS scan time of 50 ms. We termed the method parallel accumulation – serial fragmentation (PASEF), since precursor ions are accumulated in parallel and released sequentially dependent on their ion mobility.

Finally, we investigated the effect of such an increase in MS/MS speed by modeling the outcome on typical shotgun proteomics data. We observed that around 250,000 features were detected within a 90 min gradient and of those, 45,000 features were targeted without using PASEF. By applying PASEF, with a four-fold increase in sequencing speed, almost all peptide features could in principle be targeted. However due to the limit of detection of the mass spectrometer most of the low intensity peptides would not be identified. By using PASEF for increasing sequencing speed and simultaneously increasing sensitivity by targeting low intensity features repeatedly, we expect an increase of 300% of targeted and potentially identified peptide features with current hardware specifications.

Journal of proteome research

Article

pubs.acs.org/jpr

Parallel Accumulation–Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device

Florian Meier,[†] Scarlet Beck,[†] Niklas Grassl,[†] Markus Lubeck,[‡] Melvin A. Park,[§] Oliver Raether,[‡] and Matthias Mann*[,†]

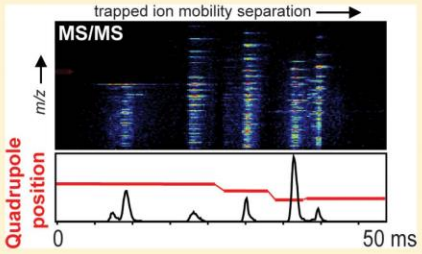
[†]Proteomics and Signal Transduction, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

[‡]Bruker Daltonik GmbH, Fahrenheitstrasse 4, 28359 Bremen, Germany

[§]Bruker Daltonics Inc., 40 Manning Road, Billerica, Massachusetts 01821, United States

Supporting Information

ABSTRACT: In liquid chromatography-mass spectrometry (LC-MS)-based proteomics, many precursors elute from the column simultaneously. In data-dependent analyses, these precursors are fragmented one at a time, whereas the others are discarded entirely. Here we employ trapped ion mobility spectrometry (TIMS) on an orthogonal quadrupole time-of-flight (QTOF) mass spectrometer to remove this limitation. In TIMS, all precursor ions are accumulated in parallel and released sequentially as a function of their ion mobility. Instead of selecting a single precursor mass with the quadrupole mass filter, we here implement synchronized scans in which the quadrupole is mass positioned with sub-millisecond switching times at the m/z values of appropriate precursors, such as those derived from a topN precursor list. We demonstrate serial selection and fragmentation of multiple precursors in single 50 ms TIMS scans. Parallel accumulation–serial fragmentation (PASEF) enables hundreds of MS/MS events per second at full sensitivity. Modeling the effect of such synchronized scans for shotgun proteomics, we estimate that about a 10-fold gain in sequencing speed should be achievable by PASEF without a decrease in sensitivity.



KEYWORDS: *proteomics, MS/MS, ion mobility, TIMS, peptide sequencing, multiplexing, time-of-flight, high resolution*

INTRODUCTION

High-resolution mass spectrometry (MS)-based proteomics has emerged as a powerful technique for large-scale profiling of thousands of proteins with many applications in molecular and cellular biology.[1−3] A typical bottom-up shotgun proteomics workflow starts with the extraction and solubilization of the protein material prior to enzymatic digestion. The peptide mixture is subsequently separated via liquid chromatography (LC) and electrosprayed into a mass spectrometer. To derive sequence information, suitable precursor ions are isolated by their mass-to-charge ratio (m/z) and subjected to collision-induced fragmentation, followed by database identification. Precursor scan and fragmentation are commonly performed with a data-dependent topN method, in which a MS survey spectrum is followed by fragmentation spectra of the N most abundant precursors. In complex mixtures, the depth of the analysis is thus foremost limited by the sequencing speed and sensitivity of the mass spectrometer. Previously we demonstrated that only 16% out of 100,000 peptide features eluting during a 90 min LC gradient were targeted for MS/MS.[4] State-of-the-art proteomics MS instruments partly address these challenges by increasing sequencing speed and resolving power.[5,6] However, as more and more MS/MS spectra are acquired per second, less and less acquisition time is available for each precursor—an inherent consequence of the serial nature of the MS/MS process. Parallel fragmentation of unselected precursors in methods such as MS^{E}[7] or SWATH[8] addresses this problem at the expense of multiplexing the MS/MS spectra, making peptide identification more challenging and precluding iTRAQ[9] and TMT[10] based multiplexing.[10,11]

Time-of-flight (TOF) instruments acquire spectra at very high frequency. This makes them capable of using the majority of the precursor and fragment ions, thus promising optimal sensitivity and sequencing rates. However, although MS/MS rates of up to 100 Hz[12] can be readily achieved, in proteomic practice, sensitivity is generally not sufficient to detect an adequate number of fragment ions for peptide identification at such high scan speeds. That is, because the quadrupole filter transmits a given type of precursor ion for only a small fraction of the time in which that species is eluting from the LC column, only a small fraction of available precursor ions are utilized

Received: October 6, 2015
Published: November 5, 2015

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

when operating at high MS/MS spectral rates. Nevertheless, quadrupole TOF instruments have improved their performance over the last few years, and appear ready to become a viable alternative to the prevalent Orbitrap analyzer-based technology.[13−16] We recently described a state of the art QTOF instrument, with high resolution and ion transmission, which enabled rapid and in-depth analysis of complex proteomes, including the identification of more than 11,000 different proteins in brain tissue.[17]

Due to their high scanning speeds (about 0.1 ms/spectrum), TOF instruments are compatible with ion mobility spectrometry (IMS), which happens on the time scale of tens of milliseconds.[18−21] While IMS-MS can increase speed, selectivity, and sensitivity,[22−24] available platforms have entailed a considerable increase in instrumental complexity and reduced ion transmission.[25] A particular form of ion mobility, termed trapped ion mobility spectrometry (TIMS), features a particularly compact construction, without compromising resolution or transmission.[26−29] In the TIMS device, ions are accumulated in an RF-only tunnel at a position where the force of a gas flow equals the opposing force of an electric field. Ion mobility separated species are released from the device as a function of their collisional cross section. Here we asked if this sequential release after parallel accumulation can be exploited to drastically increase the speed and sensitivity of MS/MS experiments.

■ MATERIALS AND METHODS

Sample Preparation

Purified and predigested standards of enolase, phosphorylase b, alcohol dehydrogenase (ADH), and bovine serum albumin (BSA) were purchased from Waters GmbH (Eschborn, Germany) and resuspended in 0.1% formic acid to prepare stock solutions at a concentration of 10 pmol/μL each. These stock solutions were combined in an equimolar ratio and diluted in 50% water/50% acetonitrile/0.1% formic acid (v/v/v) to a final concentration of 100 fmol/μL.

Trapped ion mobility spectrometry−mass spectrometry

To mimic conditions during the analysis of complex proteomics samples, we directly infused the digested four protein mixture into a prototype, high-resolution QTOF mass spectrometer equipped with a TIMS device (flow rate 3 μL/min). For an overview of the instrument, see the Results and Discussion below. A detailed description of the construction and operation of the TIMS analyzer employed here has been published elsewhere.[26,30] Briefly, the TIMS device is composed of stacked ring electrodes, which form three distinct sections: an entrance funnel, the TIMS tunnel, and an exit funnel. The experiments were performed using nitrogen as a bath gas at room temperature, and the gas flow velocity was kept constant by regulating the pressure at the inlet and the outlet of the TIMS cartridge. Ions were accumulated for 50 ms, and mobility separation was achieved by ramping the entrance potential from −180 V to −40 V within 50 ms (435 TOF scans of 115 μs each). TIMS and MS operation were controlled and synchronized using the instrument control software otofControl (Bruker Daltonik, Bremen, Germany).

Synchronized TIMS and quadrupole operation

To enable the PASEF method, precursor m/z and mobility information was first derived from full scan TIMS-MS experiments (with a mass range of m/z 150−1850). Resulting quadrupole mass, collision energy, and switching times were manually transferred to the instrument controller as a function of the total cycle time via direct firmware commands. The quadrupole isolation width was set to 3 Th and, for fragmentation, the collision energies were varied between 30 and 60 eV depending on precursor mass and charge.

Data analysis and modeling of increased sequencing rates

Ion mobility resolved mass spectra, nested ion mobility vs m/z distributions, as well as summed fragment ion intensities were extracted from the raw data files with a prototype version of DataAnalysis (Bruker Daltonik). S/N ratios were increased by summations of individual TIMS scans. Mobility peak positions and peak half-widths were determined based on extracted ion mobilograms (±0.05 Da) with an in-house script written in Python, using the peak detection algorithm implemented in the DataAnalysis software.

To model the anticipated benefit from PASEF scans for shotgun proteomics, we reanalyzed a single shot HeLa experiment from our recently published data set acquired on an impact II QTOF instrument.[17] Feature detection and peptide identification were performed with MaxQuant version 1.5.2.8 applying the previously described search parameters. Further analysis of the MaxQuant output was performed in the R statistical computational environment.[31] For each isotope pattern ("MS feature"), MaxQuant reports the minimum and maximum scan number where it was detected.[32] To simulate a topN method, each detected feature was assigned to a single scan (at the minimum scan number) and the N most abundant features of each bin were selected according to the original MS peak picking criteria (m/z > 300 and charge >1). This procedure reproduced the actual peak picking procedure performed by the Bruker acquisition software.

To restrict the simulations to likely peptide precursors, the MS features present in the MaxQuant output, as well as the target species assigned as described above, were further filtered for retention times between 10−100 min, charge state 2−5 and m/z values >450, which yielded a log-normal intensity distribution.

■ RESULTS AND DISCUSSION

Trapped ion mobility spectrometry (TIMS) − Mass spectrometry

To investigate the PASEF method that is the subject of this study, we made use of a prototype TIMS-QTOF mass spectrometer (Figure 1). In this instrument, a TIMS tunnel was incorporated into the first pumping stage, upstream from the transfer and selection quadrupoles and the high resolution mass analyzer.

Briefly, ions are generated in an electrospray source, transferred into the vacuum system through a glass capillary, deflected by 90°, and focused by the entrance funnel into the TIMS tunnel. This tunnel consists of pairs of stacked electrodes to which an RF field is applied. A DC field in the longitudinal direction is superimposed on this RF field (Figure 1B). Gas originating from the capillary flows through the sealed tunnel at a pressure of about 2−3 mbar. Ions entering the tunnel experience a drag due to the gas flow and a counteracting electrical field. They come to rest at a position where these two forces are equal and are radially confined by the pseudopotential induced by the RF field. Since the drag is proportional to the collisional cross section, ions of different mobility are trapped at different positions along the longitudinal axis, with

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

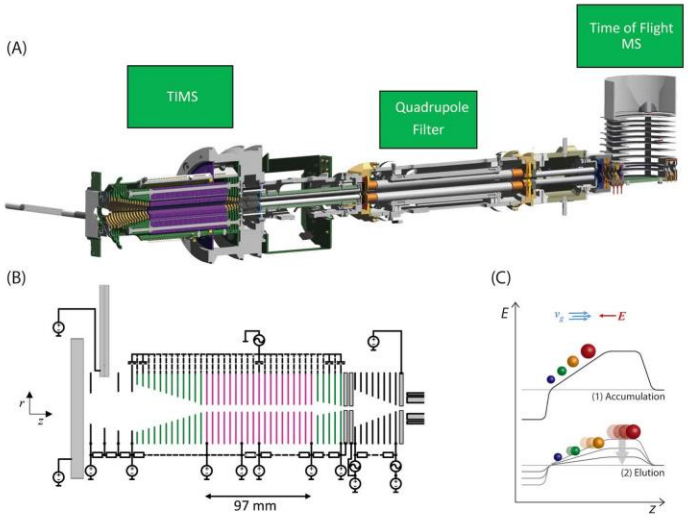


Figure 1. Trapped ion mobility spectrometry coupled to a QTOF mass spectrometer. (A) Instrument schematic of the prototype TIMS-QTOF instrument used in this study. (B) Detailed schematic of the TIMS tunnel (purple), enclosed by the entrance and exit funnels (green). (C) General mode of TIMS operation, including ion accumulation (1) and serial elution (2) of ion mobility separated ions from the TIMS device by decreasing the electrical field. The directed forces of the gas flow and electrical field are indicated by v_g and E.

high mobility ions at the entrance and low mobility ions near the exit of the tunnel (Figure 1C). After a desired ion accumulation time, further ions are prevented from entering the tunnel by a change of the potential on the deflection plate. Stored ions are then "eluted" by decreasing the electric field strength. Eluted ions are focused by the exit funnel and pass through a second funnel and transfer quadrupole. Precursor ions can be isolated by the analytical quadrupole mass filter for optional subsequent fragmentation in the collision cell. Afterward narrow ion packages are accelerated into a field-free drift region by the orthogonal deflection unit. A two-stage reflectron compensates differences in the kinetic energy of the ions and they are detected on an MCP detector coupled to a 10-bit digitizer, similar to the Bruker impact II instrument.[17]

The total length of the TIMS device is only about 10 cm (Figure 1B), and it is operated with modest potentials of less than 300 V. Since the resolution is mainly determined by the rate at which the electrical field is decreased, the user is free to adjust it based on experimental needs. For slow ramp times, ion mobility resolutions (expressed as $R = \Omega/\Delta\Omega$, where Ω is the collisional cross-section) of more than 200 have been demonstrated[28,29] and faster scan out times of about 50 ms still allow $R > 40$.[30]

Parallel accumulation–serial fragmentation (PASEF)

Ion mobility–mass spectrometry adds an additional dimension of separation to the standard MS scans. The elution ramp described above defines a complete cycle of TIMS-MS, with the ions that have the lowest mobility (largest collisional cross sections) in relation to their charge state passing through the instrument first. During the elution ramp, TOF spectra are recorded at high frequency (~8.7 kHz). We term the data structure generated during a TIMS-MS cycle a "TIMS scan". Several TIMS scans can be added to obtain a TIMS-MS spectrum of a desired signal-to-noise.

The TOF spectra from one scan can be summed (i.e., projected onto the m/z axis) to determine the m/z and intensity of all the ions present. This information can then be used to create a topN target list for fragmentation. MS/MS is performed in the usual way by setting the quadrupole transmission window to the values in the topN list and applying the collision energy that is optimal for the precursor mass and charge state. The MS/MS spectra produced when incorporating ion mobility are similar in most respects to those obtained by MS/MS only, with two principal differences: First, compared to operation without IMS, in which the signal is recorded continuously as long as the precursor is selected, the signal is compressed into a short time—i.e. the duration of the mobility peak. This leads to better signal-to-noise as signal is concentrated whereas noise is distributed. Second, the different precursors present in the quadrupole selection window are separated from each other by their different ion mobilities even at the same m/z. This alleviates the 'Precursor Ion Fraction (PIF) problem', which comes about because PIF values are much smaller than one for a majority of precursors in proteomics experiments.[4,33]

Despite these two advantages, MS/MS experiments performed in the above-described way still only make use of one selection window per scan, in common with all methods that use a mass selecting quadrupole to isolate precursors (Figure 2, upper panel). Inspired by the controllable nature of the release of ions from the TIMS device and the fact that cross sections roughly correlate with precursor mass,[34] we wondered if it would be possible to utilize several precursors in one scan, without giving up the mass selection of each of them by the quadrupole. In the PASEF method described and demonstrated here, the quadrupole is positioned at the m/z of a precursor eluting from the TIMS analyzer and is then rapidly moved to the next one as soon as it has eluted (Figure 2, lower panel). In

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

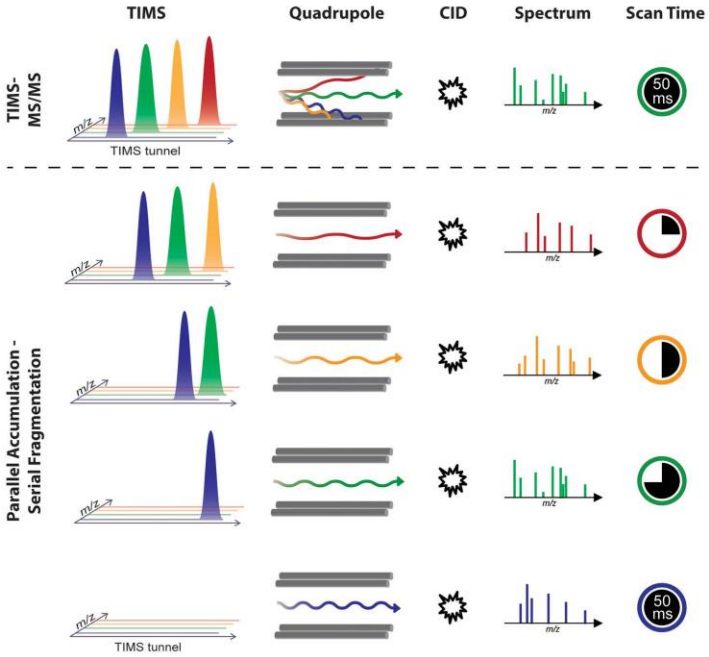


Figure 2. Illustration of the PASEF method in comparison with the standard TIMS-MS/MS operation mode. The top panel shows the selection of one precursor from a single TIMS scan, while all others are discarded. Conversely, the PASEF method (bottom panel) involves rapid switching of the quadrupole mass position to select multiple precursors at different m/z on the very same time scale. In this case, all targeted ions are fully used for fragmentation.

this way, the full intensity of the precursors that have been accumulated together can be utilized in one TIMS scan. This increases the speed of MS/MS by the number of precursors that are targeted. Alternatively, the same precursors can be selected in different TIMS scans to gain sensitivity; again by a factor up to the number of selected precursors per TIMS scan, or speed and sensitivity advantages can be combined as desired.

Time scales and required switching times

Key to realizing the concept of PASEF are the efficient storage of ions of the intended precursor range, high ion mobility resolution, and the extremely rapid switching of the quadrupole between precursors on the time scale of a single TIMS scan. To establish the hardware requirements for this, it is worthwhile to first consider the time scales of the individual processes involved in the LC-MS/MS experiment. With the QTOF setup in our laboratory,[17] we typically achieve average chromatographic peak widths of about 7 s (FWHM) during a 90 min LC gradient. The MS acquisition cycle for a data-dependent top17 method is completed within 1.2 to 1.4 s. Given a 200 ms interval for summing MS survey scans of 110 μs each, up to 17 precursors per second are selected by the quadrupole mass filter and fragmented in the collision cell. Therefore, the quadrupole mass position, as well as the collision energy, are switched every 60 ms. During this time, the TOF mass analyzer performs about 550 scans.

When incorporating ion mobility separation into this workflow, it becomes evident that a typical TIMS scan of 20 to 60 ms is orders of magnitude shorter than either the chromatographic peak widths or the topN cycle times (Figure 3). Conversely, the TIMS scan is about 100-fold longer than the time required for a single TOF scan and thus readily fits in between the chromatographic time scale and the TOF scan time. Individual ion mobility peaks have half widths of one millisecond or less (see below). Since the PASEF method aims to select multiple precursors from a single TIMS scan, precursor isolation should also happen within one millisecond. Note that this requires a much faster precursor isolation than the 60 ms of the normal top17 method, ideally by about 100-fold. To investigate the feasibility of quadrupole switching at these time scales, we first evaluated the rise times of the power supplies and lag times that potentially affect ion transmission. This revealed that the quadrupole is in principle capable of stepping its mass position by 3000 Th within 1.5 ms or less. Considering the much narrower m/z range relevant for proteomics, this should make even sub-millisecond switching times possible. Adding the width of a typical ion mobility peak results in a quadrupole isolation time per precursor of currently 2.5 to 4 ms. This translates into a maximum of 12−20 precursors per 50 ms TIMS scan, or 240−400 precursors per second. Importantly, the signal per precursor still corresponds to the full accumulation time. Thus, PASEF should be able to measure MS/MS spectra up to 12−20 times faster without losing sensitivity. Clearly, this number exceeds the number of precursors that can usefully be selected for fragmentation in contemporary shotgun proteomics experiments. Therefore,

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

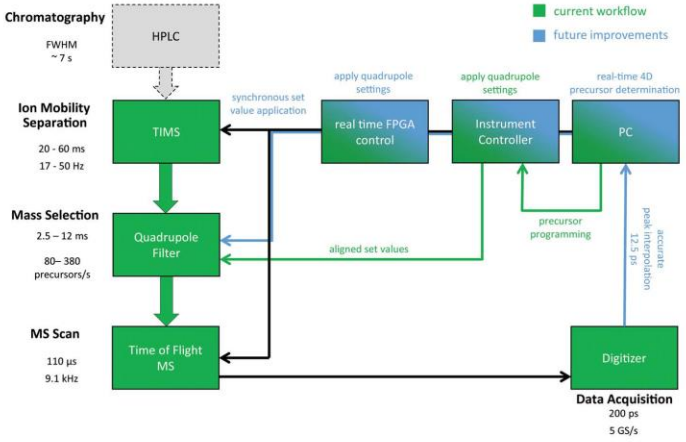


Figure 3. Investigation of hardware requirements for the PASEF method, including time scales and switching times of a typical proteomics LC-TIMS-MS/MS experiment. Arrows and lines visualize the hardware connections in the currently used setup (green) and planned future improvements (blue). GS/s, Giga-samples per second.

excess sequencing speed can be expended on improving identification scores or MS/MS-based quantification as explained in more detail below.

Precursor selection at this frequency can also be limited by the rate at which the quadrupole set values are calculated and applied by the instrument controller. In fact, we found this to be one of the major bottlenecks of our current setup since the set values are transferred via a comparably slow serial interface, precluding a fully synchronized operation of the quadrupole and TIMS. For our proof of principle experiments, we here circumvent this limitation by determining the elution times from TIMS precursor scans and applying the mass position and switching times for MS/MS scans to the quadrupole via the instrument controller in an asynchronous manner. This enabled us to select four precursors for a single TIMS scan. In the future, we will overcome this limitation by moving the calculations and set value application to the real time field-programmable gate array (FPGA), which already synchronizes TIMS and TOF analyzer. An additional direct interface to the quadrupole driver will allow set value applications in less than 50 μs, thus no longer compromising the quadrupole switching times (depicted in blue in Figure 3).

Note that the PASEF operation performed here will not change due to these improvements, except that we will be able to make use of the maximum mass selection rate of the quadrupole. Moreover, advanced peak detection and compression algorithms executed by the digitizer, capable of handling several million peaks per second, will enable real-time precursor determination in four dimensions (LC retention time, elution time from the TIMS device, m/z value and intensity) as required for data-dependent topN methods. To not compromise the proteome coverage in a shotgun experiment, this will also involve dynamic exclusion of already sequenced peptides. However, as the precursor determination itself can take up to a few milliseconds, we plan to parallelize precursor search and data acquisition. Notably, this strategy would also support more sophisticated precursor search algorithms without reducing the duty cycle for data acquisition. From the discussion above, we conclude that current limitations are not of a fundamental nature, but tasks for engineering and data-handling which are likely to be accomplished in the near future.

Application of PASEF to complex protein digests

Having established sub-millisecond switching of the quadrupole isolation window, we next aimed to examine the novel method under circumstances that mimic the simultaneous elution of many peptides from the chromatographic column in shotgun proteomics. We thus directly infused an unseparated mixture of digested ADH, BSA, enolase and phosphorylase b to generate high peptide complexity (Materials and Methods).

Figure 4A shows the result of an ESI-TIMS-MS analysis of the four protein mixture. Without ion mobility separation, this experiment would have yielded a very complex mass spectrum with multiple overlapping signals as indicated by the projection on the m/z axis on the right. TIMS separates these overlapping ion species by their mobility, resulting in much less complex mass spectra per TOF scan. In accordance with others,[19,35] we observe a significant correlation between m/z and mobility, whereas higher m/z species are less mobile and vice versa. There are two main populations that can be assigned to different charge states, with singly charged species being less mobile than their multiply charged counterparts at similar m/z. Charge 2−5 species—potential tryptic peptides—make up a more dense population. Nevertheless, close inspection of the mobility peak widths in this regime revealed that the elution time chosen (ramping down the TIMS tunnel gradient in 50 ms) was sufficient to achieve a mobility resolution above 40 and to separate these species.

From the heat map created by the multiply charged tryptic peptides of the protein mixture, we first selected four mobility separated precursors. Corresponding mass positions at m/z 810.3 (1), 714.3 (2), 559.3 (3) and 560.6 (4) and the appropriate switching times were uploaded to the instrument controller. As apparent from Figure 4B, the quadrupole correctly isolated these precursors on the TIMS time scale. In each case, the entire peak was quantitatively captured. Projection onto the m/z scale also shows successful isolation,

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

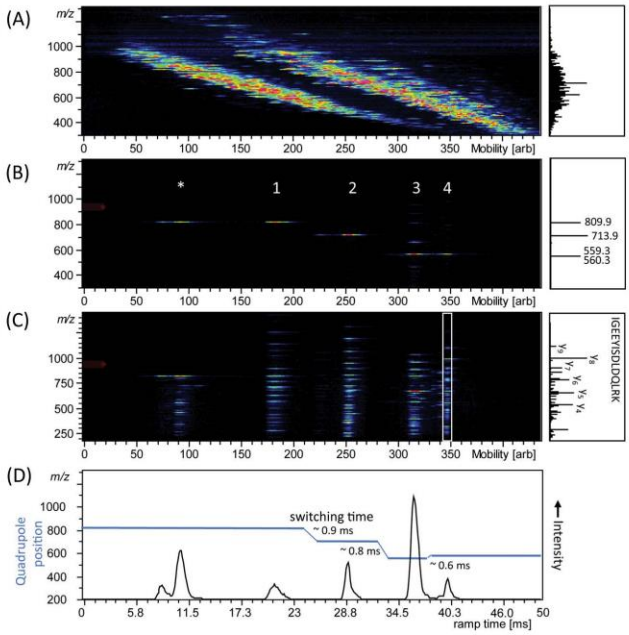


Figure 4. TIMS-QTOF analysis of electrosprayed peptides from a combined tryptic digest of ADH, BSA, phosphorylase b, and enolase. (A) Nested m/z and ion mobility distribution as detected in full scan MS. (B) Sequential isolation of four ions at different m/z that are separated by mobility after parallel accumulation. (C) Parallel accumulation–serial fragmentation (PASEF) of the precursors as isolated in (B). (D) Arrival time distribution of the summed fragment ions as observed by the PASEF method together with the quadrupole isolation mass as a function of the TIMS ramp time. Projected mass spectra are shown on the right of the panels, in (C) the fragments of the white box are projected.

and comparison to the spectrum without quadruple isolation demonstrates a drastic simplification of the precursor population. Precursors (3) and (4) only differ in one m/z unit, representing the classical PIF problem, but are clearly separated by the combination of ion mobility and quadrupole isolation. Another notable feature is the peak isolated before peptide (1) and marked with an asterisk. This is caused by the fact that in the current configuration the quadruple immediately started selection from the first m/z value. Consequently, it isolated a precursor from the singly charged population, which is widely separated from the actual target and upon inspection turns out to consists of two species that are distinct in mass and ion mobility.

In the next step, we applied the PASEF method by performing MS/MS on the isolated precursors. This led to a characteristic ladder of fragment ions at each precursor (Figure 4C, Supplementary Figure S1). Projection of the ladder generated by peptide (4) shows a standard MS/MS spectrum, corresponding to the sequence IGEEISDLDQLRK from phosphorylase B. The other peptides corresponded to VLGI-DGGEGKEELFR (1) from enolase, HLQIIYEINQR (2) from phosphorylase B and VAAAFPGDVDR (3), also from phosphorylase B. Interestingly, we were also able to assign one of the singly charged species to ADH based on its peptide fragment spectrum (YVVDTSK). Projection of all fragment ions onto the ion mobility axis shows coherence in arrival times, with fragment ion distributions very similar to their precursors (Figure 4D). These peaks were less than 3 ms wide, with typical half-widths of around 1 ms. In contrast, the chosen quadrupole isolation times were at least three times longer and could therefore in principle have been shortened considerably. Combined with the fast switching times apparent in the figure, this shows that at least ten precursors could have been targeted in this 50 ms ramp.

To demonstrate PASEF on a larger scale, we extended the previously described experiment to a total of ten sets of four precursors each. The results on the 40 precursors were very similar to the example discussed above and are summarized in Table 1. The selected precursors span a mass range from m/z 418 to 956 and exhibited ion mobilities corresponding to TIMS elution times between 17 and 44 ms with an average half width of 0.8 ± 0.2 ms. In this experiment we also quantified the gain due to PASEF by comparing it to standard fragmentation of the 40 precursors with TIMS. With full realization of the PASEF concept, we would expect that we would preserve full signal intensity despite being 4-fold faster (or N-fold in general, where N is the number of selected precursors). This was fully validated by the results, which showed a ratio indistinguishable from one between the two cases (0.94 ± 0.12, Table 1).

Modeling of coverage improvements in shotgun proteomics

Having demonstrated that PASEF is capable of a four to more than 10-fold increase in MS/MS speed at full sensitivity, we next modeled the effect of such an improvement on typical shotgun proteomics data. We built on our previous analysis, in

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

Table 1. PASEF Analysis of 40 Precursors from a Complex Mixture of Four Digested Proteins

	Precursor				Mobility	FWHM	Ratio
Scan	Protein	Peptide Sequence Targeted	*m/z*	Charge	[ms]	[ms]	PASEF/TIMS[a]
1	Enolase	VLGIDGGEGKEELFR	809.956	2	22.6	1.3	0.87
	Phos b	HLQIIYEINQR	713.897	2	30.4	0.7	0.85
	Phos b	VAAAFPGDVDR	559.253	2	37.4	0.9	0.97
	Phos b	IGEEYISDLDQLRK	560.261	3	40.9	0.5	0.96
2	Enolase	SIVPSGASTGVHEALEMR	921.033	2	17.4	1.3	0.95
	Phos b	IGEEYISDLDQLR	775.914	2	28.2	0.7	0.97
	BSA	LVNELTEFAK	582.292	2	36.2	0.7	0.77
	BSA	LFTFHADICTLPDTEK	636.633	3	39.6	1.1	0.90
3	Phos b	TCAYTNHTVLPEALER	938.031	2	22.0	1.5	0.94
	Phos b	VLYPNDNFFEGK	721.862	2	28.7	0.7	1.00
		no match	658.303	4	35.1	0.6	1.12
	Phos b	LITAIGDVVNHDPVVGDR	630.659	3	39.7	0.9	1.00
4	Enolase	TAGIQIVADDLTVTNPK	878.527	2	23.2	0.7	1.07
		no match	767.925	2	29.8	0.9	0.93
	BSA	DAIPENLPPLTADFAEDKDVCK	820.099	3	34.2	0.7	1.10
	Phos b	TNFDAFPDK	527.705	2	40.3	0.6	0.82
5	Phos b	WLVLCNPGLAEIIAER	927.568	2	22.4	0.8	0.95
	Enolase	AVDDFLISLDGTANK	789.928	2	29.3	0.8	0.92
		no match	594.311	2	35.5	0.7	1.01
	Phos b	TCAYTNHTVLPEALER	625.624	3	39.9	0.9	0.82
6		*no match*	955.970	2	17.3	0.7	1.09
	Phos b	DFNVGGYIQAVLDR	783.927	2	27.6	0.8	0.91
	ADH	SISIVGSYVGNR	626.323	2	34.5	0.6	0.92
	Enolase	VLGIDGGEGKEELFR	540.249	3	39.0	0.8	0.92
7	BSA	YNGVFQECCQAEDK	874.403	2	27.1	0.9	0.93
	Phos b	VFADYEEYVK	631.784	2	33.6	0.6	(0.37)[b]
	Phos b	QRLPAPDEK	527.248	2	37.5	0.8	0.90
	BSA	DDPHACYSTVFDK	518.847	3	43.4	0.6	0.72
8	Phos b	IGEEYISDLDQLRK	839.973	2	24.2	0.8	1.05
	ADH	GLAGVENVTELKK	679.383	2	31.4	0.8	1.04
	Phos b	ARPEFTLPVHFYGR	563.934	3	37.7	0.6	0.78
	ADH	IGDYAGIK	418.668	2	43.6	0.6	1.03
9	BSA	KVPQVSTPTLVEVSR	820.502	2	26.4	0.8	0.98
	Phos b	LLSYVDDEAFIR	720.877	2	30.2	1.0	0.88
	Phos b	VAIQLNDTHPSLAIPELMR	706.720	3	34.9	0.8	1.05
	Phos b	APNDFNLK	459.682	2	41.2	0.6	(0.28)[b]
10		*no match*	782.429	2	23.8	0.7	1.16
	BSA	HLVDEPQNLIK	653.351	2	32.6	0.8	1.13
	Enolase	IGSEVYHNLK	580.281	2	35.3	0.7	0.67
	Phos b	NLAENISR	458.691	2	39.7	0.6	0.70

[a]Median summed fragment ion intensities were extracted for each precursor from PASEF ($N = 331$) and TIMS-MS/MS ($N = 9$) scans with identical quadrupole isolation settings. [b]As an artifact resulting from the asynchronous operation of TIMS and quadrupole, these two precursors were not isolated in each PASEF scan.

which we used MaxQuant to determine the number of isotope patterns ("peptide features") as a function of retention time in a 90 min analysis of a HeLa digest and used this to evaluate speed, sensitivity and separation power achievable in shotgun proteomics.[4] We repeated this analysis using our data from the impact II QTOF instrument,[17] on which our prototype is based (Materials and Methods). From the first to last eluting peptides, about 50 to 70 unique features/s are detected by MaxQuant, of which about 12 had been fragmented per second (Figure 5A). The 4-fold improvement in sequencing speed already demonstrated above would allow targeting 80% of all detected features. We noted above that a 10-fold improvement is entirely consistent with the actually achieved switching times and ion mobility resolution. This would result in a targeting rate of about 125 features/s, far above the number of features detected in the previous data set. That said, adding ion mobility could be expected to increase the number of resolvable peptide features in shotgun proteomics runs, in which case even a 10-fold increased sequencing speed could in principle be used entirely on unique peptide features.

However, not all detected features are suitable for MS/MS fragmentation because their peptide intensity may be too low to result in useful MS/MS spectra. Redoing the peptide histogram analysis that we had done on the original Orbitrap data[4] revealed 3-fold larger numbers for the impact II data set but the same proportions: About 250,000 peptide features can be detected in the 90 min gradient, of which about 45,000 were fragmented, and about 30,000 with attendant peptide identification (Figure 5B).

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

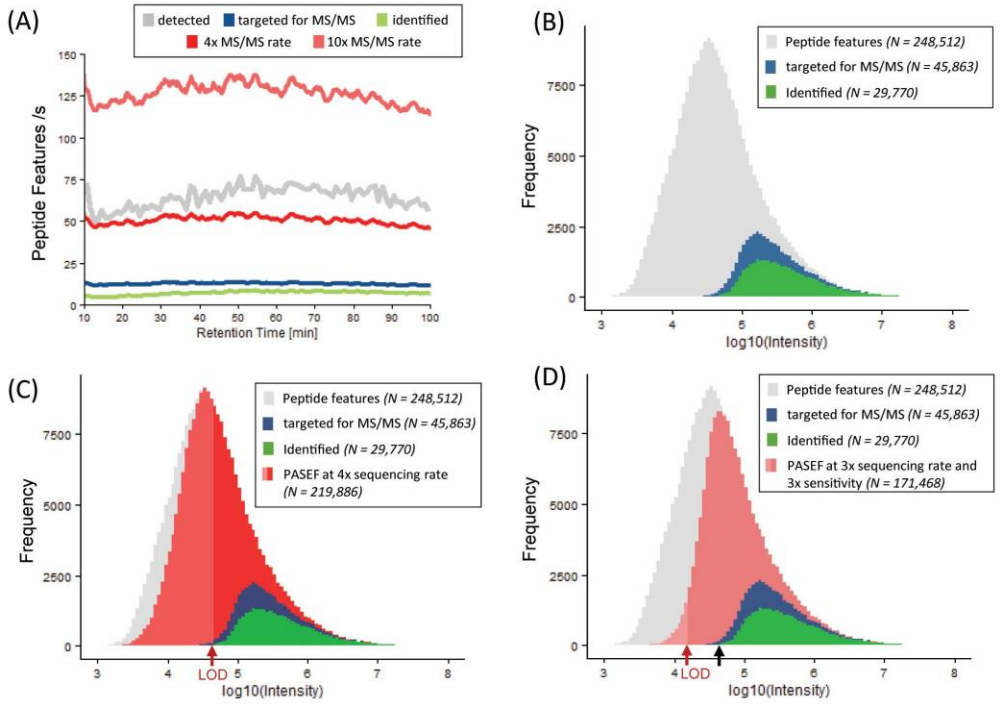


Figure 5. Modeling the benefits of the PASEF method in shotgun proteomics. (A) Peptide features per second eluting during a 90 min gradient. The gray trace indicates detected peptide features during conventional LC-MS/MS analysis of a tryptic HeLa digest on a QTOF instrument; blue and green traces indicate features that were targeted for MS/MS and successfully identified in the impact II data set.[17] Red traces represent the simulated sequencing speed of PASEF with four and ten precursor ions per TIMS scan, respectively. (B) Intensity distribution of detected (gray), targeted (blue), and identified (green) peptide features extracted by MaxQuant from the impact II data set. (C) Same as (B) with a simulation of targeted peptides at 4-fold sequencing speed with the PASEF method (red). (D) Same as (B) with a simulation of a PASEF experiment with three precursors per scan and 3-fold increased sensitivity due to retargeting (red). MS features below the anticipated limit of detection (LOD) are indicated by reduced opacity.

With 4-fold higher sequencing speed, essentially the entire population of detected peptide features can now be targeted, but lower intensity precursors would diminish identification success (Figure 5C). The sensitivity of the impact II instrument for peptide identification is given by the lower limit of the green population in the peptide histogram. If we assume that the current instrument reaches at least the same sensitivity, then a 4-fold improvement in sequencing at full sensitivity, allows filling in about 40% of the entire peptide precursor population, a 250% improvement from before.

As mentioned above, the PASEF advantage could be used for increasing sequencing speed or sensitivity—through targeting the same feature repeatedly. In practice, we imagine that one would use a combination of both, for instance by fragmenting sufficiently abundant precursors once but summing MS/MS spectra for lower abundant ones. To model this in a simple way, we equally assigned the expected 10-fold improvement due to PASEF to sequencing speed and sensitivity gain (Figure 5D). This leads to a similarly shaped distribution of targeted precursor as in the experimental distribution, with the difference that the target and likely identified populations are increased by 300%. As a result, 70% of the overall precursor population can now be targeted and potentially identified by MS/MS. Future experiments, using a full LC-TIMS-MS/MS set up, will reveal optimal combinations of refragmentation and precursors per TIMS scan. We also note that more sophisticated selection strategies for MS/MS could further improve the proportion of successfully identified peptides and that improvements in ion beam intensity would increase both the total number of detected features as well as the number of targetable and identifiable peptides.

■ CONCLUSIONS

Coupling ion mobility spectrometry to MS comprises several advantages, such as separation of ions from protein mixtures according to their size-to-charge ratio. Using the compact TIMS analyzer implemented into a state of the art QTOF mass spectrometer, we have here introduced the concept of PASEF. Sub-millisecond quadrupole switching times allowed us to select multiple precursors for fragmentation during a 50 ms ion mobility scan instead of only one. Importantly, we demonstrated that synchronized quadrupole and TIMS operation is fully quantitative in that the signal is not diminished compared to single precursor selection. An important advantage of PASEF

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

is that the resulting spectra—in addition to the ion mobility dimension—are fully precursor mass resolved, unlike recently proposed data independent strategies. This also makes PASEF compatible with reporter ion based chemical multiplexing strategies, such as iTRAQ or TMT. The about 10-fold gain that should be achievable by PASEF in shotgun proteomics experiments can be employed as increased sequencing speed without a decrease in sensitivity. However, modeling suggests that a combination of targeting more precursors and targeting weak precursors repeatedly, will be most effective. While demonstrated here for the TIMS-QTOF combination, the PASEF principle could be applied to any ion mobility−mass spectrometer configuration with the required sub-millisecond scan speed in the MS read out.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.5b00932.

Supplementary Figure S-1: Annotated PASEF-MS/MS spectra of the peptides shown in Figure 4 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +49-89-8578-2557. E-mail: mmann@biochem.mpg.de.

Author Contributions

All authors contributed to the research design; F.M., S.B., and M.L. performed research; F.M., S.B., M.L., O.R., and M.M. analyzed data; F.M., S.B., O.R., and M.M. wrote the manuscript. All authors have given approval to the final version of the manuscript. F.M. and S.B. contributed equally.

Notes

The authors declare the following competing financial interest(s): Several authors are employees of Bruker.

■ ACKNOWLEDGMENTS

We thank our colleagues in the Department of Proteomics and Signal Transduction and at Bruker for fruitful discussions. We particularly acknowledge our co-workers in Bremen, Niels Goedecke, Joerg Hoffmann, and Stefan Harsdorf, for support in setting up the prototype, as well as Peter Sander and Thomas Betz for providing data analysis tools. We acknowledge Max-Planck Society for the Advancement of Sciences for funding.

■ ABBREVIATIONS

ADH, alcohol dehydrogenase; BSA, bovine serum albumin; FPGA, field-programmable gate array; FWHM, full width at half-maximum; GS/s, Giga-samples per second; IM, ion mobility; LC, liquid chromatography; MS, mass spectrometry; MS/MS, tandem mass spectrometry; PASEF, parallel accumulation−serial fragmentation; PIF, precursor ion fraction; QTOF, quadrupole time-of-flight; RF, radiofrequency; TIMS, trapped ion mobility spectrometry; TOF, time-of-flight

■ REFERENCES

(1) Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422* (6928), 198−207.

(2) Cravatt, B. F.; Simon, G. M.; Yates, J. R., 3rd. The biological impact of mass-spectrometry-based proteomics. *Nature* **2007**, *450* (7172), 991−1000.

(3) Altelaar, A. F.; Munoz, J.; Heck, A. J. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat. Rev. Genet.* **2013**, *14* (1), 35−48.

(4) Michalski, A.; Cox, J.; Mann, M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J. Proteome Res.* **2011**, *10* (4), 1785−93.

(5) Kelstrup, C. D.; Jersie-Christensen, R. R.; Batth, T. S.; Arrey, T. N.; Kuehn, A.; Kellmann, M.; Olsen, J. V. Rapid and Deep Proteomes by Faster Sequencing on a Benchtop Quadrupole Ultra-High-Field Orbitrap Mass Spectrometer. *J. Proteome Res.* **2014**, *13*, 6187.

(6) Scheltema, R. A.; Hauschild, J. P.; Lange, O.; Hornburg, D.; Denisov, E.; Damoc, E.; Kuehn, A.; Makarov, A.; Mann, M. The Q Exactive HF, a Benchtop Mass Spectrometer with a Pre-filter, High-performance Quadrupole and an Ultra-high-field Orbitrap Analyzer. *Mol. Cell. Proteomics* **2014**, *13* (12), 3698−708.

(7) Li, G. Z.; Vissers, J. P.; Silva, J. C.; Golick, D.; Gorenstein, M. V.; Geromanos, S. J. Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. *Proteomics* **2009**, *9* (6), 1696−719.

(8) Gillet, L. C.; Navarro, P.; Tate, S.; Rost, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* **2012**.

(9) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3* (12), 1154−69.

(10) Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **2003**, *75* (8), 1895−904.

(11) Wiese, S.; Reidegeld, K. A.; Meyer, H. E.; Warscheid, B. Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. *Proteomics* **2007**, *7* (3), 340−50.

(12) Andrews, G. L.; Simons, B. L.; Young, J. B.; Hawkridge, A. M.; Muddiman, D. C. Performance characteristics of a new hybrid quadrupole time-of-flight tandem mass spectrometer (TripleTOF 5600). *Anal. Chem.* **2011**, *83* (13), 5442−6.

(13) Silva, J. C.; Denny, R.; Dorschel, C. A.; Gorenstein, M.; Kass, I. J.; Li, G. Z.; McKenna, T.; Nold, M. J.; Richardson, K.; Young, P.; Geromanos, S. Quantitative proteomic analysis by accurate mass retention time pairs. *Anal. Chem.* **2005**, *77* (7), 2187−200.

(14) Silva, J. C.; Denny, R.; Dorschel, C.; Gorenstein, M. V.; Li, G. Z.; Richardson, K.; Wall, D.; Geromanos, S. J. Simultaneous qualitative and quantitative analysis of the Escherichia coli proteome: a sweet tale. *Mol. Cell. Proteomics* **2006**, *5* (4), 589−607.

(15) Distler, U.; Kuharev, J.; Navarro, P.; Levin, Y.; Schild, H.; Tenzer, S. Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **2014**, *11* (2), 167−70.

(16) Helm, D.; Vissers, J. P.; Hughes, C. J.; Hahne, H.; Ruprecht, B.; Pachl, F.; Grzyb, A.; Richardson, K.; Wildgoose, J.; Maier, S. K.; Marx, H.; Wilhelm, M.; Becher, I.; Lemeer, S.; Bantscheff, M.; Langridge, J. I.; Kuster, B. Ion mobility tandem mass spectrometry enhances performance of bottom-up proteomics. *Mol. Cell. Proteomics* **2014**, *13*, 3709.

(17) Beck, S.; Michalski, A.; Raether, O.; Lubeck, M.; Kaspar, S.; Goedecke, N.; Baessmann, C.; Hornburg, D.; Meier, F.; Paron, I.; Kulak, N. A.; Cox, J.; Mann, M. The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics. *Mol. Cell. Proteomics* **2015**, *14* (7), 2014−29.

(18) Hoaglund, C. S.; Valentine, S. J.; Sporleder, C. R.; Reilly, J. P.; Clemmer, D. E. Three-dimensional ion mobility/TOFMS analysis of electrosprayed biomolecules. *Anal. Chem.* **1998**, *70* (11), 2236−42.

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

(19) Valentine, S. J.; Counterman, A. E.; Hoaglund, C. S.; Reilly, J. P.; Clemmer, D. E. Gas-phase separations of protease digests. *J. Am. Soc. Mass Spectrom.* **1998**, *9* (11), 1213−6.
(20) Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H., Jr. Ion mobility-mass spectrometry. *J. Mass Spectrom.* **2008**, *43* (1), 1−22.
(21) May, J. C.; McLean, J. A. Ion mobility-mass spectrometry: time-dispersive instrumentation. *Anal. Chem.* **2015**, *87* (3), 1422−36.
(22) Hoaglund-Hyzer, C. S.; Clemmer, D. E. Ion trap/ion mobility/quadrupole/time-of-flight mass spectrometry for peptide mixture analysis. *Anal. Chem.* **2001**, *73* (2), 177−84.
(23) Shliaha, P. V.; Bond, N. J.; Gatto, L.; Lilley, K. S. Effects of traveling wave ion mobility separation on data independent acquisition in proteomics studies. *J. Proteome Res.* **2013**, *12* (6), 2323−39.
(24) Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nat. Chem.* **2014**, *6* (4), 281−94.
(25) Cumeras, R.; Figueras, E.; Davis, C. E.; Baumbach, J. I.; Gracia, I. Review on ion mobility spectrometry. Part 1: current instrumentation. *Analyst* **2015**, *140* (5), 1376−90.
(26) Fernandez-Lima, F. A.; Kaplan, D. A.; Park, M. A. Note: Integration of trapped ion mobility spectrometry with mass spectrometry. *Rev. Sci. Instrum.* **2011**, *82* (12), 126106.
(27) Fernandez-Lima, F.; Kaplan, D. A.; Suetering, J.; Park, M. A. Gas-phase separation using a trapped ion mobility spectrometer. *Int. J. Ion Mobility Spectrom.* **2011**, *14* (2−3), 93.
(28) Silveira, J. A.; Ridgeway, M. E.; Park, M. A. High resolution trapped ion mobility spectrometery of peptides. *Anal. Chem.* **2014**, *86* (12), 5624−7.
(29) Ridgeway, M. E.; Silveira, J. A.; Meier, J. E.; Park, M. A. Microheterogeneity within conformational states of ubiquitin revealed by high resolution trapped ion mobility spectrometry. *Analyst* **2015**, *140* (20), 6964−72.
(30) Michelmann, K.; Silveira, J. A.; Ridgeway, M. E.; Park, M. A. Fundamentals of trapped ion mobility spectrometry. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (1), 14−24.
(31) R Development Core Team. *R: A Language and Environment for Statistical Computing*; 2014.
(32) Cox, J.; Matic, I.; Hilger, M.; Nagaraj, N.; Selbach, M.; Olsen, J. V.; Mann, M. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* **2009**, *4* (5), 698−705.
(33) Houel, S.; Abernathy, R.; Renganathan, K.; Meyer-Arendt, K.; Ahn, N. G.; Old, W. M. Quantifying the impact of chimera MS/MS spectra on peptide identification in large-scale proteomics studies. *J. Proteome Res.* **2010**, *9* (8), 4152−60.
(34) Revercomb, H. E.; Mason, E. A. Theory of Plasma Chromatography Gaseous Electrophoresis - Review. *Anal. Chem.* **1975**, *47* (7), 970−983.
(35) Taraszka, J. A.; Counterman, A. E.; Clemmer, D. E. Gas-phase separations of complex tryptic peptide mixtures. *Fresenius' J. Anal. Chem.* **2001**, *369* (3−4), 234−45.

DOI: 10.1021/acs.jproteome.5b00932
J. Proteome Res. 2015, 14, 5378−5387

2.3. Improving isobaric mass tagging quantification by trapped ion mobility mass spectrometry

Scarlet Beck, Florian Meier, Heiner Koch, Oliver Raether, Markus Lubeck, Niels Goedecke, Juergen Cox, and Matthias Mann

Reducing the ratio compression problem in isobaric mass tagging experiments by trapped ion mobility spectrometry (TIMS) – mass spectrometry

In preparation

A well-known problem associated with isobaric mass tagging as well as with data-dependent acquisition experiments is the interference of multiple precursors that is caused because the resolution of liquid chromatography and the mass isolation window is not enough to isolate each eluting peptide from each other.

Different publication have investigated this problem because on the one hand it decreases the number of peptide identifications and on the other hand it leads to ratio distortion in isobaric mass tagging experiments. Several methods have been developed to eliminate the ratio compression problem, but they still have disadvantages such as longer duty cycles, resulting in less peptide identifications.

Ion mobility mass spectrometry separates ions based on their ion mobility and *m/z* and therefore has the potential to reduce or even eliminate the generation of chimeric MS/MS spectra. In my third project, I have investigated if trapped ion mobility spectrometry (TIMS) - mass spectrometry (MS) can be used to reduce the ratio compression problem. For this purpose, I have first investigated which resolution can be achieved when applying TIMS by varying the release times. Then I have calculated the number of interfering peptides with and without applying TIMS. With a median ion mobility resolution of 78 the effect of interference could be reduced by two-fold. Several examples show that peptides that would not be distinguishable without applying TIMS, are now separated based on their ion mobility. Additional investigations determined that narrowing the isolation mass window or increasing the ion mobility resolution would help to further reduce interference.

Reducing the ratio compression problem in isobaric mass tagging experiments by trapped ion mobility spectrometry (TIMS) – mass spectrometry

Scarlet Beck[1], Florian Meier[1], Heiner Koch[1], Markus Lubeck[2], Oliver Raether[2], Jürgen Cox[3] and Matthias Mann[1]*

[1]Proteomics and Signal Transduction, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

[2]Bruker Daltonik GmbH, Fahrenheitstr. 4, 28359 Bremen, Germany

[3]Computational Systems Biochemistry, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

*To whom correspondence should be addressed: mmann@biochem.mpg.de

Running title: Reducing ratio compression by TIMS

Abbreviations: iTRAQ – isobaric tags for relative and absolute quantitation, MS – mass spectrometry, LC – liquid chromatography, MS1 – full mass spectrum, MS2 – tandem mass spectrum, PIF – precursor ion fraction, QTOF – quadrupole time-of-flight, SPS – synchronous precursor selection, TIMS – trapped ion mobility spectrometry, TMT – tandem mass tag, TOF – time-of-flight, TWIMS – travelling wave ion mobility spectrometry

SUMMARY

Proteomics experiments can be multiplexed by isobaric tagging approaches such as the tandem mass tag (TMT) reagent. However, more than one precursor can contribute to the reporter ions, leading to a compression of the measured TMT ratios compared to the theoretically expected ones. Here we combine trapped ion mobility spectrometry – mass spectrometry (TIMS-MS) with isobaric labeling to reduce or eliminate ratio compression without compromising analysis speed or sensitivity. Release (ramp out time) from the TIMS device determine the ion mobility resolution. By using a ramp out time of 25 ms, 50 ms, 75 ms and 100 ms, we achieved median mobility resolutions of 22, 42, 60 and 78 respectively. To investigate the effect of ion mobility resolution on ratio compression, we assessed interfering ions within the volume elements defined by chromatographic peak width, mass selection window and ion mobility width. Choosing a mass selection window of 2 Th and an ion mobility resolution corresponding to 100 ms release time caused more than 70% of the volume elements to be free of co-isolated species. This is a vast improvement from not using TIMS, where this proportion was less than 30%. Correspondingly, TMT ratios in selected volume elements agreed with theoretically expected ones. We conclude that TIMS combined with TMT is a promising avenue for multiplexing in proteomics.

Mass spectrometry (MS)-based proteomics has emerged as a powerful tool for the identification and quantification of thousands of proteins in biological samples such as cell or tissue lysates [1, 2]. Over the last years different quantification strategies have been developed [3] that can be categorized into label-free quantification, in which protein abundance is calculated based on peptide intensities across runs [4, 5] and label-based quantification, in which different natural isotopes are used to quantify combined samples within the same spectra, either at the full mass spectra (MS1) or at tandem mass spectra (MS/MS) level, based on reporter ions [6-9]. Depending on the labeling strategy and the available number of different labels, multiplexing of samples can reduce MS measurement time several-fold. In isobaric tagging methods such as isobaric tags for relative and absolute quantification (iTRAQ) [10] or tandem mass tag (TMT) [11], samples are combined only after digestion, potentially giving rise to artifactual quantitative differences caused by up front sample handling [3]. However, ratios can often be determined very consistent within the reporter ion patterns and therefore isobaric tagging strategies often have high precision of measurement.

The main drawback of reporter ion based quantitation strategies is that co-isolation and co-fragmentation of multiple peptides within a defined quadrupole isolation window leads to chimeric MS/MS spectra. While co-isolated peptides could be distinguished by backbone fragmentation, all of them contribute to the very same reporter ion channels. This leads to distortion of the measured reporter ion ratios in comparison to the theoretically expected ones [12-14]. As the peptide of interest is usually one with a non-equal reporter ion pattern, whereas the background is not regulated, this tends to produce an underestimation of the actual ratios, and this is termed the ‘ratio compression problem’.

To overcome these challenges, different strategies have been developed in the last few years. Several groups have restricted the selection window of the precursor in the mass or elution dimension, have extensively fractionated peptide mixtures or have attempted to correct the reporter ion ratios for the observed precursor ion fraction (PIF) in the selection window [12, 15-17]. However, none of these approaches have been entirely successful and as an example, a recent large-scale studies still reports a mean ratio underestimation by a factor two [18]. Two methods manipulate ions in the mass spectrometer by either changing their charge states [16] or further fragmenting one of the MS/MS fragments [13]. The latter method has been further refined by co-isolating several fragments, which increases sensitivity, however, potentially re-introduces ratio compression if they originate from different precursors [19]. This synchronous precursor selection (SPS) method, while not solving the ratio compression problems completely [20],

requires a very high end mass spectrometer and still has a relatively slow cycle time, but nevertheless has been shown to enable quite large scale, multiplexed measurements [21-23].

It has recently been reported that ion mobility spectrometry in combination with mass spectrometry potentially decreases the number of co-isolated peptides. For example Shliaha et al. reported that travelling wave ion mobility spectrometry (TWIMS) - MS reduces reporter ions contamination [24]. However, TWIMS did not generally solve the ratio compression problem, perhaps due the limited ion mobility resolution currently achievable with this technology.

Several different ion mobility principles have been developed over the last few years that can be combined with MS. One of the latest is trapped ion mobility spectrometry (TIMS), introduced by Park and co-workers in 2011 [25]. In TIMS, ions are separated in an ion tunnel based on size and shape (collision cross section). In contrast to other approaches, this is achieved by a gas flow that drags the different ion species through the TIMS tunnel, while they experience a counteracting force due to an electrical field, bringing them to rest at the specific position where the two forces balance. Attractive features of the TIMS device include its compact nature, low voltage requirements and the fact that its resolving power is tunable through the release time of the ions from the device [26-28].

In this paper, we investigated the application of trapped ion mobility spectrometry (TIMS) in combination with quadrupole time-of-flight (QTOF) mass spectrometry for accurate quantification in isobaric tagging experiments. We observed that TIMS-MS removes the ratio compression problem in isobaric tagging experiments for a large majority of the peptides.

EXPERIMENTAL PROCEDURES

Sample preparation of HeLa and yeast lysates: HeLa cells (ATCC, S3 subclones) were cultured in Dulbecco`s modified Eagle`s medium containing 10% fetal bovine serum, 20 mM glutamine and 1% penicillin-streptomycin (all from PAA Laboratories). 5 x 10^7 cells were centrifuged at 200 x g for 10 min, washed once with cold phosphate buffered saline (PBS) and then centrifuged again. Supernatant was discarded and the cell pellet shock frozen in liquid nitrogen and stored at -80°C until further use.

Yeast strain (BY4741) clones were seeded in 2 x 30 ml of YPD medium (including 2% glucose w/v) and were grown over night at 30 °C at 200 rpm. On the next day, 500 ml YPD (including 2% glucose w/v) were pre-warmed in a 3 l Erlenmeyer flask (with baffles) and were grown for 5-6 h at 30 °C at 200 rpm to OD_{600} 0.8. Then 40 ml of the culture were pelleted at 3,500 g for 5 min at 4 °C. Pellets were resuspended in 25 ml cold PBS and centrifuged at 3,500 g for 5 min at 4 °C. They were resuspended in 1 ml cold PBS and centrifuged again at 3,500 g for 5 min at 4 °C. Finally, they were snap frozen in liquid nitrogen and stored at -80 °C until further use.

The cell pellets were resuspended in 1 % (w/v) sodium deoxycholate, 10 mM TCEP, 40 mM 2-chloroacetamide (CAA), 100 mM Tris-HCl, pH 8.5 for cell lysis. Lysates were heated to 100 °C for 10 min at 1400 rpm to enhance protein denaturation and to stop protease and phosphatase activity. Lysates were homogenized 15 min in a bioruptor (Diagenode) and LysC (Wako Chemicals GmbH) was added 1:100 and incubated at 30 °C overnight. The samples were transferred to StageTips (Empore) [29] containing SDB-RPS material, and were washed two times with 200 µl isopropanol 1% TFA. Afterwards the StageTips were washed with 200 µl of 0.2% TFA in ddH_20 and samples were eluted in 60 µl of 1% (w/v) TAE buffer, 80% ACN. Samples were dried in a speedvac and stored at -20 °C until further use. They were then subjected to single shot LC-MS/MS or LC-TIMS-MS/MS measurements with or without previous labeling with TMT 6-plex labeling reagent (Thermo Scientific).

Labeling of HeLa and yeast lysates: The HeLa peptides or yeast peptides (see above) were dissolved in 35 µl 50 mM HEPES pH 8.5 and labeled with 5 µl of the six TMT labels, leading to a final concentration of 7.5 mM TMT reagent. Afterwards the samples were diluted with 0.1% FA to acidify them and to lower the concentration of ACN to 1%. Samples were combined in a ratio of 10:10:10:0:0:0 (HeLa) and 10:4:1:1:4:10 (yeast). Labeled HeLa and yeast sample were dissolved in 0.1% FA and combined 1:1 (v:v). The combined samples were desalted on StageTips. Briefly,

they were equilibrated with 40 µl 0.1% FA, then the sample was loaded and washed two-times with 0.1% FA. Samples were eluted with 60% ACN, 0.1% FA and dried followed by storage at -80 °C before further measurement.

Trapped ion mobility spectrometry (TIMS) – mass spectrometry (MS): All analyses were performed on a prototype, high-resolution QTOF mass spectrometer equipped with a TIMS device (Fig. 1A). Detailed description of the TIMS device and QTOF mass spectrometer have been published before [26, 30, 31].

LC-TIMS-MS/MS analysis: We used a trapping column set-up (PepMap pre-column, 2 cm x 100 µm; Thermo Scientific) and a Dionex HPLC pump (Ultimate 3000 nRSLC, Thermo Scientific). They were on-line coupled to a prototype TIMS-QTOF instrument with a CaptiveSpray ion source (both Bruker Daltonics). Peptides were separated on a Pep-Map UHPLC column (50 cm x 75 µm, 2 µm particles; Thermo Scientific) using a 90 min multistep ACN gradient (buffer A: 0.1% FA, buffer B 100% ACN in 0.1% FA). In the TIMS tunnel, ions were accumulated for 25 ms, and mobility separated. They were released from the TIMS device by ramping the entrance potential from 207 V to 77 V within 25, 50, 75 or 100 ms (219, 438, 657, 877 TOF scans of 114 µs each). We used nitrogen as a bath gas at 31 °C at a constant flow velocity. The ion mobility scale was calibrated using selected masses from ESI-L Tuning Mix (Agilent). After acquiring a full scan MS spectrum, the N most intense precursors (topN) were selected for fragmentation. To keep the duty cycle equal for the different ramp times, we used a top40 method for the 25 ms ramp time, top20 for 50 ms, top13 for 75 ms and top10 for 100 ms. The total cycle time was around 1.2 s. If no ion mobility separation was needed for the experiment, the TIMS tunnel was switched to the transmission only mode where it functions as a funnel.

Data analysis: For the data analysis with the MaxQuant software (version 1.5.5.6) [32] with the Andromeda search engine [33], we cut all raw files into 5 min retention time slices by using a python script and performed the data processing only from retention 60 to 65. The peptide spectrum match and protein false discovery rates (FDR) were set to 1%. The minimum length of amino acids was specified to 7. We used an initial allowed mass deviation of the precursor ions of up to 70 ppm, which was then reduced by MaxQuant after time-dependent recalibration of the

precursor masses and allowed a fragment mass deviation of 35 ppm. We searched against the Uniprot human database (downloaded on June 21, 2014, containing 88,976 entries and 247 contaminants). Enzyme specificity was set as C-terminal to lysine and arginine, also allowing cleavage at proline bonds and a maximum of two missed cleavages. Carbamidomethylation of cysteine was selected as fixed modification and N-terminal protein acetylation and methionine oxidation as variable modifications.

Data were further analyzed with the MaxQuant Viewer and processed using the R statistical programming environment [34].

RESULTS AND DISCUSSION

Reducing interference by trapped ion mobility spectrometry (TIMS)

For the investigation of the interference problem in trapped ion mobility spectrometry (TIMS) – mass spectrometry (MS) we used a prototype QTOF mass spectrometer equipped with a TIMS device (Fig. 1A). Briefly, ions are generated in an electrospray source, transferred into the vacuum system through a glass capillary and deflected by 90° by the deflection plate before entering the entrance funnel of the TIMS device (total length around 14 cm). There ions are focused and further guided into the TIMS tunnel (10 cm long), where they are separated based on ion mobility. The TIMS tunnel is composed of stacked ring electrodes to which a radio-frequency field and an electrostatic field current field are applied. A gas flow at a pressure of about 2-3 mbar originating from the capillary drags the ions through the TIMS tunnel until they experience an equal countervailing force due to an opposing electric field. Thus, ions remain at a specific position dependent on their mobility. Larger ions are positioned close to the exit of the tunnel, while very small or compact ion species are located at the entrance of the tunnel. After ions are accumulated for a specific time period, they are released by decreasing the electric field strength. Ion packages that elute from the TIMS tunnel are focused into the exit funnel of the TIMS device and guided through the transfer quadrupole. This is followed by the analytical quadrupole, where precursor ions can be selected for subsequent fragmentation in the collision cell. Then, ions are accelerated in the field-free flight tube by the orthogonal accelerator and are deflected by a reflectron before they are detected on an MCP detector coupled to a 10-bit digitizer.

In liquid chromatography (LC)-tandem mass spectrometry (MS/MS) analysis a large number of peptides elute from the column simultaneously after separation by reversed-phase chromatography. If these ion species differ only slightly in mass, they are co-isolated by the quadrupole and co-fragmented in the collision cell, resulting in chimeric MS/MS spectra. Figure 1B (upper panel) schematically illustrates the LC-MS/MS analysis of a pair of co-eluting peptides. These peptides are also co-isolated and co-fragmented, producing chimeric MS/MS spectra as well as chimeric reporter ion patterns. In LC-TIMS-MS/MS analyses, peptides that elute from the column are first separated dependent on their mobility before specific ions are selected by the quadrupole for further analysis. Figure 1B (lower panel) shows how separation based on their ion mobility followed by separate isolation and fragmentation, results in uncontaminated reporter ion spectra for each peptide. This additional dimension of separation based on collisional cross sections therefore has the potential to suppress the generation of chimeric MS/MS spectra partially or completely, without a reduction in sensitivity or analysis speed.

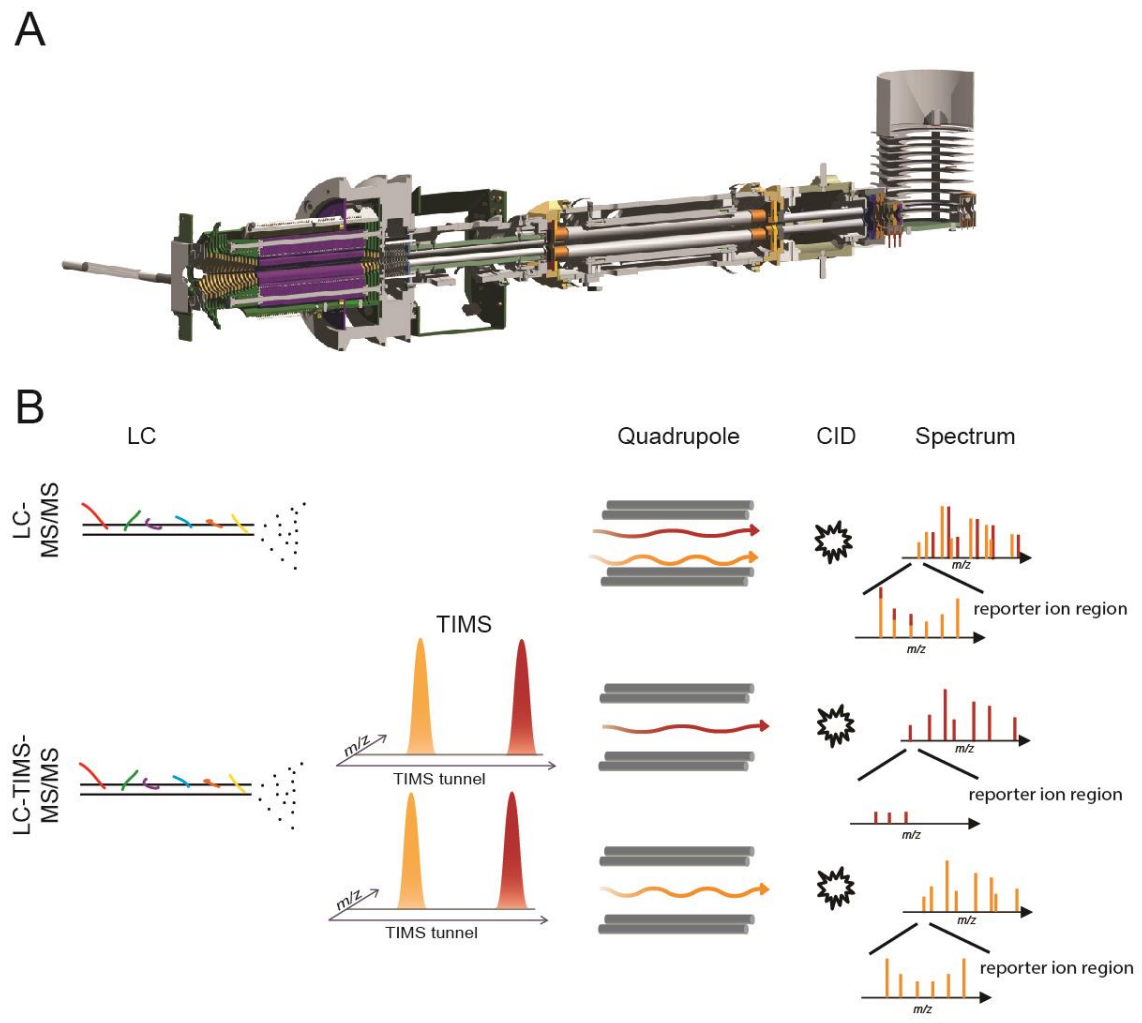


Figure 1: Trapped ion mobility spectrometry (TIMS) – mass spectrometry (MS). (A) Scheme of the prototype TIMS-QTOF mass spectrometer. (B) Illustration of isobaric interference in LC-MS/MS and LC-TIMS-MS/MS. Two co-eluting peptides of similar mass are co-isolated and co-fragmented, resulting in distorted reporter ion pattern. TIMS separates the peptides based on their ion mobility and fragments them separately, potentially resulting in interference free reporter ion measurement.

Investigating interference at the MS level

Shotgun proteomics data feature two dimensions at the MS level, the mass range and the retention time of the peptides eluting from the chromatographic column. Figure 2A illustrates the extremely high complexity of shotgun proteomics data acquired from a complete mammalian cell lysate separated over a two hour gradient. Zooming into a very dense range of the heat map

shows that a large number of peptide that elute from the column simultaneously only differ slightly in mass. A 2 Th precursor isolation window in the quadrupole co-fragments them in the collision cell. Those peptides all contribute to the reporter ion intensities in isobaric mass tagging experiments, resulting in in-accurate quantitation. In addition, chimeric MS/MS spectra can hamper peptide identifications, although some database search algorithm can identify from more than one peptide from one MS/MS spectrum. With the additional dimension of separation due to trapped ion mobility spectrometry, peptides are further separated based on their mobility. As an example, Figure 2B shows that within an isolation window of 2 Th peptides are distinguishable based on their difference in ion mobility but not on their *m/z*. Note, however, that the mass and the mobility are correlated because larger ions tend to have larger cross sections. Therefore, we wished to investigate how high the ion mobility resolution would need to be to meaningfully reduce the co-fragmentation problem.

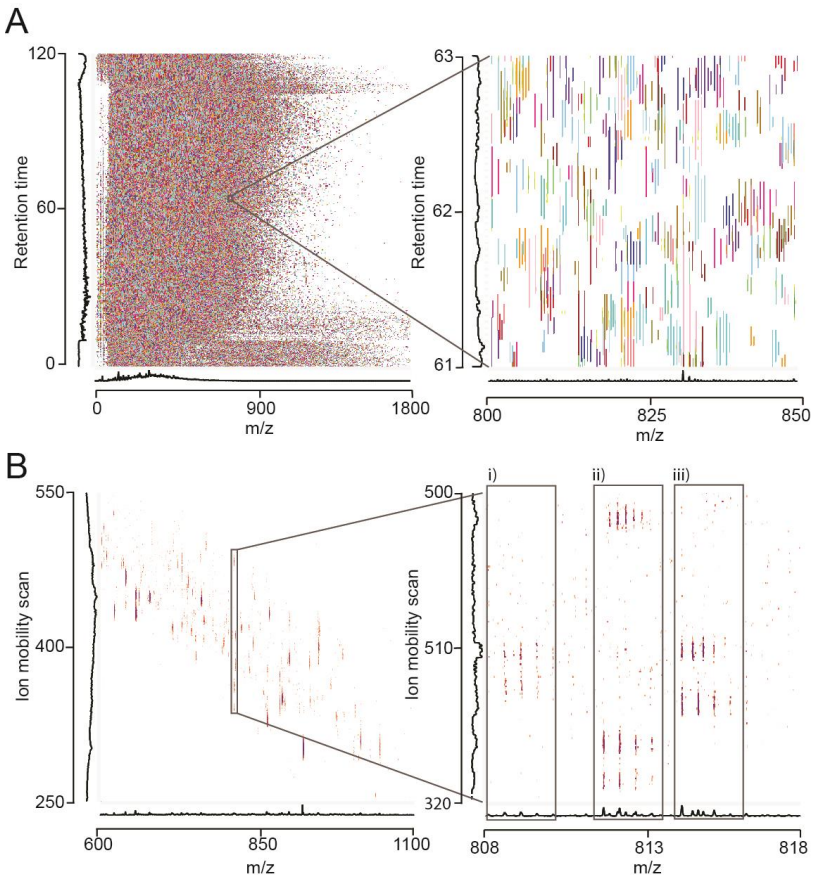


Figure 2: Peptide heat maps of a mammalian cell lysate. (A) LC-MS heat map and zoom showing a large number of co-eluting peptide features. (B) LC-TIMS-MS heat map showing nearly isobaric, co-eluting peptides (i), ii), iii)) within an isolation window of 2 Th that are separated based on their mobility.

Investigating the interference problem in LC-TIMS-MS

To investigate if TIMS-MS can be used to reduce the ratio compression problem to achieve accurate quantification ratios of peptides in isobaric mass tagging experiments, we first explored which ion mobility resolution can be achieved with the TIMS device in a typical proteomics experiment. We varied the release time for ions from the TIMS tunnel from 25 ms up to 100 ms, while keeping fill time constant at 25 ms (Fig. 3A). We calculated the resolution for all peptide features that were determined by the MaxQuant environment using standard settings within a 5 min window (from retention time 60 min to 65 min). Within a 5 min window more than 450 MS spectra were acquired containing more than 10,000 peptide features. A release time of 25 ms yielded a median resolution of 22, which increased to 78 with a release time of 100 ms. Moreover, increasing ion mobility resolution leads to the detection of more peptide features. Next, we investigated how often co-isolation and co-fragmentation appears in LC-MS and LC-TIMS-MS analysis. For this purpose we used all the peptide features determined by MaxQuant from the LC-TIMS-MS with a release time of 100 ms to assess co-eluting peptides in an isolation window of 2 Th with and without applying TIMS. Figure 3B demonstrates that more than 70% of all detectable features show interference in an isolation window of 2 Th without the ion mobility separation dimension. When applying TIMS with a release time of 100 ms, peptide interference can be reduced by around two-fold.

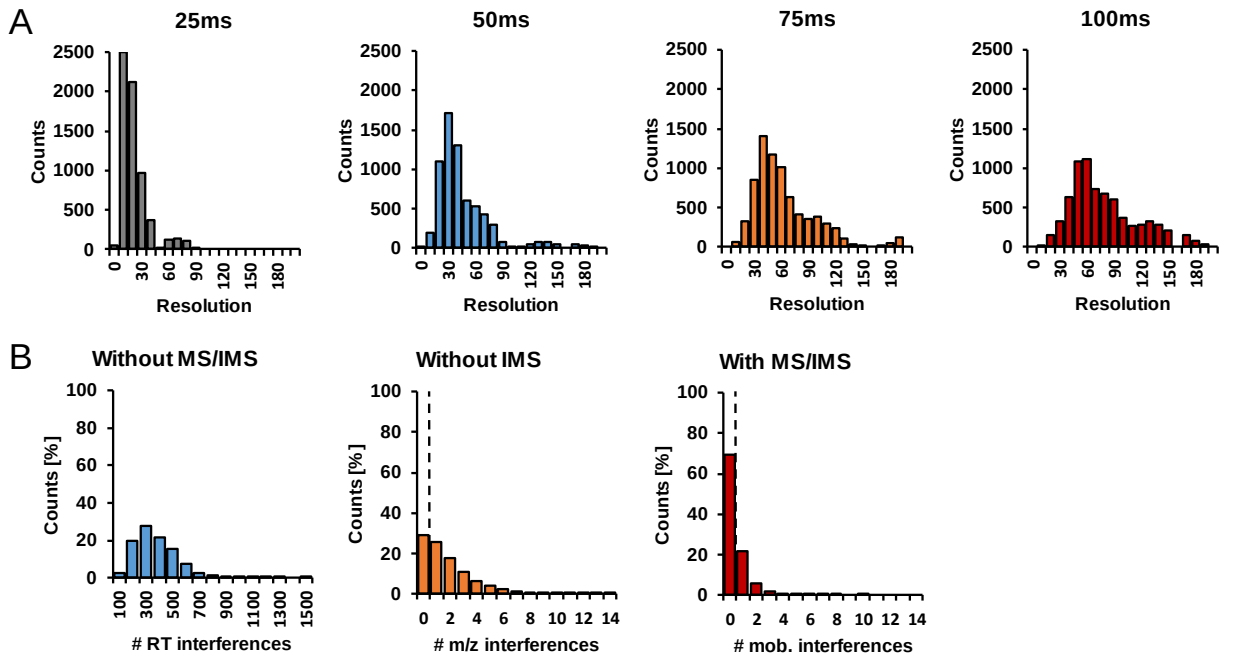


Figure 3: (A) Increase in IMS resolution by increasing the ramp time of the TIMS device. (B) Alleviating the interference problem by applying TIMS-MS. Retention time interferences were counted within a +/- FWHM of the retention time peak, MS interferences within a 2 Th window and ion mobility interferences within a 2 Th window and +/-FWHM of the ion mobility peak.

Using a narrow isolation window should help to further reduce co-isolation and co-fragmention [35]. Therefore, we next modeled the effects of varying the isolation window from 2 Th to 0.1 Th on our data. Figure 4A shows that an isolation window of 0.1 Th would almost eliminate co-isolation and co-fragmentation, however such an isolation window would be impractical due to a drastically reduced the ion transmission efficiency. Currently, an isolation window of between 1 to 2 Th is common in MS-based proteomics, and with this isolation width up to 80% of peptide precursors are pure. With a release time of 100 ms, 30% of all peptide features still show distortion. Therefore we next investigated if a higher resolution could help to eliminate the interference problem completely. For this purpose, we calculated the number of interfering peptides within +/- 1 to +/- 0.1 times the mobility peak half width at full maximum (FWHM). Figure 4B shows that very sharp peaks on the mobility axis would be necessary to completely eliminate the problem of co-isolation.

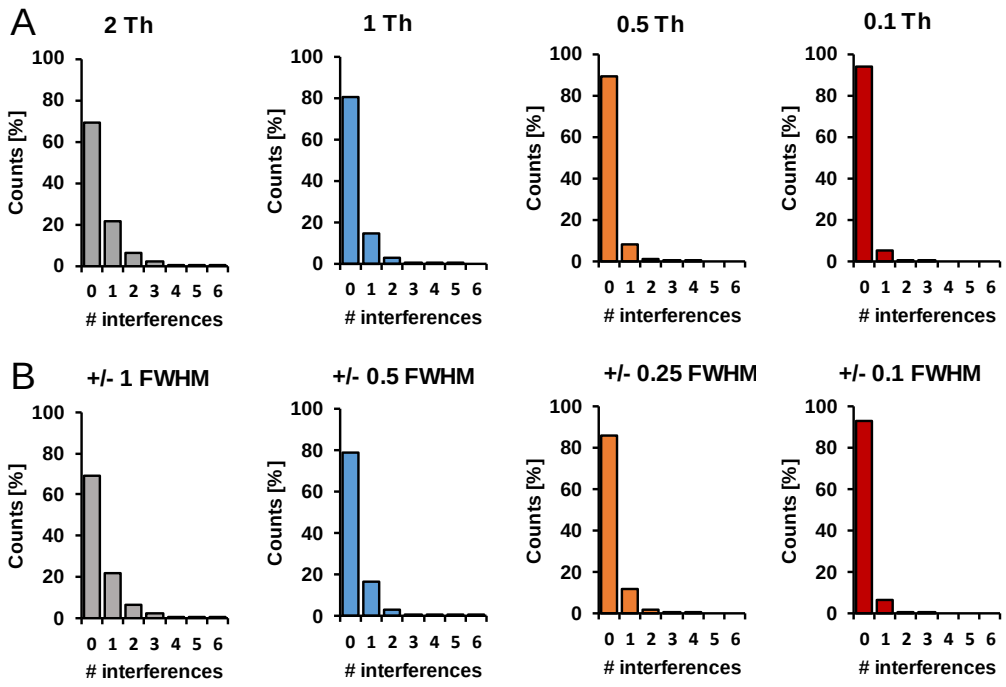


Figure 4: Reduction of interference by reducing the isolation window or mobility half width (FWHM). (A) Interferences dependent on the isolation window of the quadrupole. (B) Interferences dependent on the mobility half width. Selection of peptides in narrow ion mobility windows lowers the number of interfering peptides.

Reporter ion decompression by TIMS

To benchmark the TIMS-MS method for isobaric mass tagging experiments, we used a two-proteome model yeast/HeLa to measure ratio distortion as described previously [13]. We digested a yeast sample with LysC and labeled separate aliquots using TMT 6-plex, and then mixed those at ratios of 10:4:1:1:4:10. We did the same with a HeLa lysate, but only labeled three aliquots and mixed them 10:10:10. Finally, we mixed the two labelled peptide mixtures in a 1:1 ratio. In this kind of experiment, co-isolation and co-fragmentation of TMT labeled yeast and HeLa peptides is clearly detectable (Figure 5, marked in green). Figure 5 shows an example in which ratio distortion is eliminated by applying TIMS.

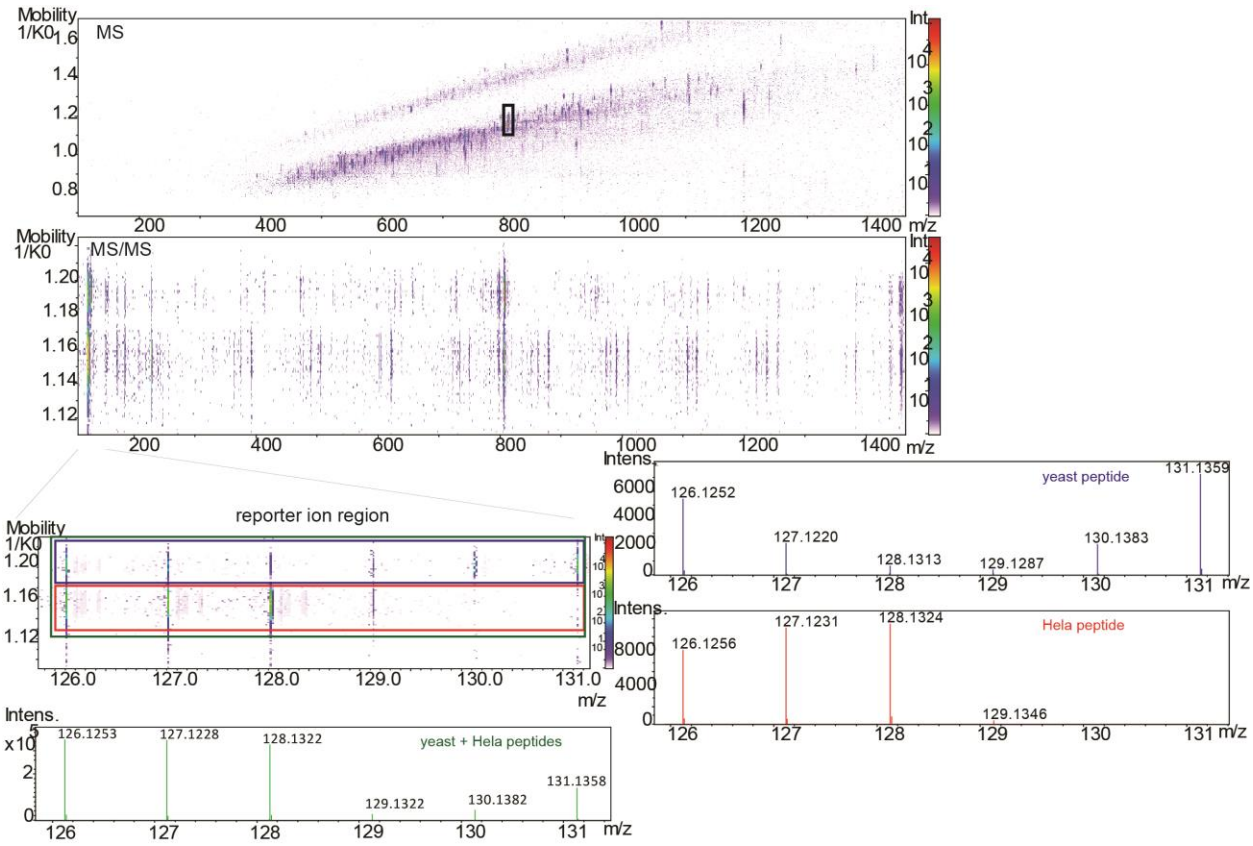


Figure 5: Example of successful decompression of reporter ion patterns. The two co-eluting peptides marked in black in the MS spectrum are separated based on their mobility before isolation by a quadrupole mass isolation window of 2 Th. One peptide is derived from yeast (blue) with an expected ratio of 10:4:1:1:4:10 and the other peptide from HeLa with an expected ratio of 10:10:10 (red).

CONCLUSION AND OUTLOOK

Here, we have shown that coupling trapped ion mobility spectrometry (TIMS) to MS provides an additional dimension of separation based on mobility. By varying the release time (25 ms to 100 ms) from the TIMS tunnel a resolution of 22 to 78 can be achieved. Further, we have demonstrated that even an ion mobility resolution of around 78 is still not sufficient to completely solve the ratio compression problem, however at least the interference of peptides can be reduced two-fold. A further increase in ion mobility resolution and using narrower isolation windows would help to further reduce interfering peptides. Furthermore, computational strategies could be applied to correct the ratios based on the observed patterns of changing ratios in ion mobility and chromatographic retention times.

REFERENCES

1. Ong, S.-E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative.* Nat Chem Biol, 2005. **1**(5): p. 252-262.
2. Kline, K.G. and M.R. Sussman, *Protein quantitation using isotope-assisted mass spectrometry.* Annu Rev Biophys, 2010. **39**: p. 291-308.
3. Bantscheff, M., et al., *Quantitative mass spectrometry in proteomics: a critical review.* Anal Bioanal Chem, 2007. **389**(4): p. 1017-31.
4. Lundgren, D.H., et al., *Role of spectral counting in quantitative proteomics.* Expert Rev Proteomics, 2010. **7**(1): p. 39-53.
5. Old, W.M., et al., *Comparison of label-free methods for quantifying human proteins by shotgun proteomics.* Mol Cell Proteomics, 2005. **4**(10): p. 1487-502.
6. Boersema, P.J., et al., *Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics.* Nat Protoc, 2009. **4**(4): p. 484-94.
7. Hsu, J.L., et al., *Stable-isotope dimethyl labeling for quantitative proteomics.* Anal Chem, 2003. **75**(24): p. 6843-52.
8. Ong, S.E., et al., *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.* Mol Cell Proteomics, 2002. **1**(5): p. 376-86.
9. Gygi, S.P., et al., *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.* Nat Biotechnol, 1999. **17**(10): p. 994-9.
10. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.* Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
11. Thompson, A., et al., *Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS.* Anal Chem, 2003. **75**(8): p. 1895-904.
12. Savitski, M.M., et al., *Measuring and managing ratio compression for accurate iTRAQ/TMT quantification.* J Proteome Res, 2013. **12**(8): p. 3586-98.
13. Ting, L., et al., *MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics.* Nat Methods, 2011. **8**(11): p. 937-40.
14. Ow, S.Y., et al., *iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly".* J Proteome Res, 2009. **8**(11): p. 5347-55.
15. Ow, S.Y., et al., *Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation.* Proteomics, 2011. **11**(11): p. 2341-6.
16. Wenger, C.D., et al., *Gas-phase purification enables accurate, multiplexed proteome quantification with isobaric tagging.* Nat Methods, 2011. **8**(11): p. 933-5.
17. Savitski, M.M., et al., *Delayed Fragmentation and Optimized Isolation Width Settings for Improvement of Protein Identification and Accuracy of Isobaric Mass Tag Quantification on Orbitrap-Type Mass Spectrometers.* Analytical Chemistry, 2011. **83**(23): p. 8959-8967.
18. Keshishian, H., et al., *Multiplexed, Quantitative Workflow for Sensitive Biomarker Discovery in Plasma Yields Novel Candidates for Early Myocardial Injury.* Mol Cell Proteomics, 2015. **14**(9): p. 2375-93.
19. McAlister, G.C., et al., *MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes.* Anal Chem, 2014. **86**(14): p. 7150-8.
20. Wuhr, M., et al., *The Nuclear Proteome of a Vertebrate.* Curr Biol, 2015. **25**(20): p. 2663-71.
21. Isasa, M., et al., *Multiplexed, Proteome-Wide Protein Expression Profiling: Yeast Deubiquitylating Enzyme Knockout Strains.* J Proteome Res, 2015. **14**(12): p. 5306-17.
22. Paulo, J.A. and S.P. Gygi, *A comprehensive proteomic and phosphoproteomic analysis of yeast deletion mutants of 14-3-3 orthologs and associated effects of rapamycin.* Proteomics, 2015. **15**(2-3): p. 474-86.

23. Chick, J.M., et al., *Defining the consequences of genetic variation on a proteome-wide scale.* Nature, 2016. **534**(7608): p. 500-505.
24. Shliaha, P.V., et al., *Additional precursor purification in isobaric mass tagging experiments by traveling wave ion mobility separation (TWIMS).* J Proteome Res, 2014. **13**(7): p. 3360-9.
25. Fernandez-Lima, F., et al., *Gas-phase separation using a trapped ion mobility spectrometer.* Int J Ion Mobil Spectrom, 2011. **14**(2-3).
26. Michelmann, K., et al., *Fundamentals of trapped ion mobility spectrometry.* J Am Soc Mass Spectrom, 2015. **26**(1): p. 14-24.
27. Silveira, J.A., M.E. Ridgeway, and M.A. Park, *High resolution trapped ion mobility spectrometery of peptides.* Anal Chem, 2014. **86**(12): p. 5624-7.
28. Meier, F., et al., *Parallel Accumulation-Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device.* J Proteome Res, 2015. **14**(12): p. 5378-87.
29. Rappsilber, J., M. Mann, and Y. Ishihama, *Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips.* Nat Protoc, 2007. **2**(8): p. 1896-906.
30. Fernandez-Lima, F.A., D.A. Kaplan, and M.A. Park, *Note: Integration of trapped ion mobility spectrometry with mass spectrometry.* Rev Sci Instrum, 2011. **82**(12): p. 126106.
31. Beck, S., et al., *The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics.* Mol Cell Proteomics, 2015. **14**(7): p. 2014-29.
32. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* Nat Biotechnol, 2008. **26**(12): p. 1367-72.
33. Cox, J., et al., *Andromeda: a peptide search engine integrated into the MaxQuant environment.* J Proteome Res, 2011. **10**(4): p. 1794-805.
34. Team, R.C., *R: A Language and Environment for Statistical Computing.* 2014.
35. Michalski, A., J. Cox, and M. Mann, *More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS.* J Proteome Res, 2011. **10**(4): p. 1785-93.

ACKNOWLEDGMENT

We thank Alexander Reim, Sophia Doll, Philipp Geyer and other members of our groups in Munich and at Bruker for discussion and help. This work was partly supported by the Max Planck Society for the advancement of Science.

3. CONCLUSION AND OUTLOOK

In this thesis construction and performance of novel quadrupole time-of-flight mass spectrometers were described and evaluated for applications in shotgun proteomics. In the first project we have measured and modelled the ion transmission through the whole instrument starting from the capillary outlet to the detector for the first time. 10% of the ions are transferred, indicating excellent efficiency. Developments and improvements on the collision cell, reflectron and detector lead to an increase in ion transmission (up to two-fold) and in resolution (80% higher in comparison to the previous instrument). In the next step, we optimized the instrument settings and showed that performance is now comparable to other state-of-the-art mass spectrometers by analyzing yeast and HeLa peptide mixtures. For example, more than 4800 proteins of the HeLa proteome could be identified in a single run within a 90 min gradient with high reproducibility (90% of all identified proteins were detected in all replicates). In addition, I achieved the deepest proteome coverage that was reporter with a QTOF instrument so far. In proteomics experiments, a deep coverage of the proteome is of interest, but even more so the absolute or relative amount of proteins within each biological sample. Therefore, I also evaluated the instrument for label-free quantification. With the novel QTOF instrument accurate (measured fold change of 0.49 ± 0.06, theoretical ratio 0.5) and reproducible quantification can be achieved. For 90% of the identified proteins the CV was below 10%. Moreover an $R^2 > 0.99$ for triplicate analysis of Hela digests was achieved. I also have tested our workflow in biological contexts by comparing the proteome between diploid and haploid S. cerevisiae and by performing global proteomic comparison of different cell lines (spinal cord neuron-neuroblastoma, mouse hepatoma and mouse embryonic fibroblast cell lines). With a 90 min gradient in total 3769 proteins could be identified in quadruplicates. Our group has described the same system a few years ago using Orbitrap technology, but there much more start material was used and much more measurement time was needed. Therefore, we conclude that the Bruker QTOF instrument can be used for shotgun proteomics analyses of complex proteomes and with further developments and improvements coverage of the complete proteome could be achieved.

Motivated by the positive outcome of the first project as well as the potential of ion mobility spectrometry, we continued developing and improving the described QTOF instrument described above. Due to their high sequencing speed, QTOFs can be easily combined with ion mobility spectrometry and this has several advantages. It improves the selectivity since co-eluting peptides with the same *m/z* can be separated based on differences in ion mobility. Moreover, by the additional dimension of separation the limit of detection can be improved by removing

background interference. An increase in speed can be achieved because peptides are separated much faster with IMS in comparison to LC. In the next project, we have equipped a QTOF mass spectrometer with a trapped ion mobility spectrometry (TIMS) device up-front the analytical quadrupole to investigate such an instrument for shotgun proteomics applications. More precisely we were interested if multiple precursors could be selected in a single TIMS scan, by Parallel Accumulation – Serial Fragmentation (PASEF). We showed that the hardware and firmware allowed sub-millisecond switching time of the quadrupole, which was important for PASEF to be applied. We next investigated the PASEF method in a complex protein digest. Instead of one precursor, we could target four, resulting in a four-fold sequencing speed increase. Modeling the effect of such an improvement for shotgun proteomics data suggests that twice as many peptide features can be successfully targeted. This increase in speed can be used to target more peptide features and to target low abundant peptides several time to increase sensitivity. This method therefore has the potential to solve the challenges of data-dependent MS-based proteomics nearly completely. All the measurements in the second project were performed by direct infuse the complex mixture. Therefore, the next step would be to apply it to LC-MS/MS measurements to realize the potential of the PASEF method in practice.

In the last project I have applied TIMS-MS for the reduction of interfering peptides. I showed that a median resolution of 78 is not sufficient to eliminate the interference problem completely. However, using TIMS to MS reduces the number of interfering peptides up to two-fold. Moreover, I showed in a typical isobaric tagging experiment that peptides that slightly differ in mass can be distinguished based on their ion mobility, resulting in non-chimeric MS/MS spectra and avoiding ratio distortion. Further improvements on instrument settings such as reducing the mass isolation window or increasing the ion mobility resolution would help to further reduce interfering peptides.

In summary, results described in this thesis clearly show the potential of quadrupole time-of-flight instruments for shotgun proteomics, especially when combining it with ion mobility spectrometry. With the QTOF mass spectrometry alone, already a very deep coverage of a proteome with high reproducibility was achieved. Combining such an instrument with TIMS will clearly increase the number of peptide and protein identifications. In addition by applying the PASEF method in TIMS-MS the increase in sequencing speed will help to target more peptides for subsequent MS/MS spectra as well as increasing the sensitivity for shotgun proteomics experiments. Moreover, TIMS-MS can be used for multiplexing, resulting in a decrease in MS measurement time and particularly has the potential for accurate quantification in isobaric labeling experiments. High specificity, sensitivity and throughput are all potentially provided by the TIMS-QTOF mass spectrometer and

this is important for in-depth profiling of complex protein mixtures such as human tissues or body fluids in clinical proteomics. For example, the chance to identify low abundant proteins in blood plasma is now increased by applying the PASEF method. Moreover, the increase in sensitivity could also be beneficial biomarker discovery since protein biomarker candidates may be present at the lower end of the plasma protein concentration range. Simultaneously, approaches for decreasing the sample complexity such as fractionation will in many cases not be necessary anymore, resulting in a reduction of MS measurement time. Likewise depletion of the high abundant plasma proteins, which may introduce a bias in the sample may not be necessary when applying TIMS-MS. The increase in sequencing speed by PASEF can also be used to develop a mixed targeted and data-dependent acquisition. A great advantage here is that there is no need to develop peptide assays before the actual measurement, as is the case in SRM targeted proteomics. Using multiplexing strategies samples from large patient cohorts can be analyzed in a reasonable time. Combined with further developments and improvements in instrumental robustness and robust high-throughput sample preparation workflow will make MS-based proteomics ready for the clinics.

4. REFERENCES

1. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors.* Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
2. Sanger, F., et al., *Nucleotide sequence of bacteriophage [phi]X174 DNA.* Nature, 1977. **265**(5596): p. 687-695.
3. Maxam, A.M. and W. Gilbert, *A new method for sequencing DNA.* Proceedings of the National Academy of Sciences of the United States of America, 1977. **74**(2): p. 560-564.
4. Sanger, F., et al., *The nucleotide sequence of bacteriophage phiX174.* J Mol Biol, 1978. **125**(2): p. 225-46.
5. Venter, J.C., et al., *The sequence of the human genome.* Science, 2001. **291**(5507): p. 1304-51.
6. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
7. Schena, M., et al., *Quantitative monitoring of gene expression patterns with a complementary DNA microarray.* Science, 1995. **270**(5235): p. 467-70.
8. Vahey, M., et al., *Performance of the Affymetrix GeneChip HIV PRT 440 platform for antiretroviral drug resistance genotyping of human immunodeficiency virus type 1 clades and viral isolates with length polymorphisms.* J Clin Microbiol, 1999. **37**(8): p. 2533-7.
9. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics.* Nat Rev Genet, 2009. **10**(1): p. 57-63.
10. O'Farrell, P.H., *High resolution two-dimensional electrophoresis of proteins.* J Biol Chem, 1975. **250**(10): p. 4007-21.
11. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.
12. Hillenkamp, F., et al., *Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Biopolymers.* Analytical Chemistry, 1991. **63**(24): p. 1193A-1203A.
13. Link, A.J., et al., *Direct analysis of protein complexes using mass spectrometry.* Nat Biotechnol, 1999. **17**(7): p. 676-82.
14. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.
15. Clamp, M., et al., *Distinguishing protein-coding and noncoding genes in the human genome.* Proceedings of the National Academy of Sciences, 2007. **104**(49): p. 19428-19433.
16. Hebert, A.S., et al., *The one hour yeast proteome.* Mol Cell Proteomics, 2014. **13**(1): p. 339-47.
17. Anderson, N.L. and N.G. Anderson, *The human plasma proteome: history, character, and diagnostic prospects.* Mol Cell Proteomics, 2002. **1**(11): p. 845-67.
18. McLafferty, F.W., et al., *Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics.* FEBS J, 2007. **274**(24): p. 6256-68.
19. Kellie, J.F., et al., *The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput.* Mol Biosyst, 2010. **6**(9): p. 1532-9.
20. Toby, T.K., L. Fornelli, and N.L. Kelleher, *Progress in Top-Down Proteomics and the Analysis of Proteoforms.* Annu Rev Anal Chem (Palo Alto Calif), 2016. **9**(1): p. 499-519.
21. Catherman, A.D., O.S. Skinner, and N.L. Kelleher, *Top Down proteomics: facts and perspectives.* Biochem Biophys Res Commun, 2014. **445**(4): p. 683-93.
22. Heck, A.J., *Native mass spectrometry: a bridge between interactomics and structural biology.* Nat Methods, 2008. **5**(11): p. 927-33.

23. Compton, P.D. and N.L. Kelleher, *Spinning up mass spectrometry for whole protein complexes.* Nat Methods, 2012. **9**(11): p. 1065-6.
24. Domon, B. and R. Aebersold, *Mass spectrometry and protein analysis.* Science, 2006. **312**(5771): p. 212-7.
25. Cravatt, B.F., G.M. Simon, and J.R. Yates, 3rd, *The biological impact of mass-spectrometry-based proteomics.* Nature, 2007. **450**(7172): p. 991-1000.
26. Choudhary, C. and M. Mann, *Decoding signalling networks by mass spectrometry-based proteomics.* Nat Rev Mol Cell Biol, 2010. **11**(6): p. 427-39.
27. Nilsson, T., et al., *Mass spectrometry in high-throughput proteomics: ready for the big time.* Nat Methods, 2010. **7**(9): p. 681-5.
28. Mallick, P. and B. Kuster, *Proteomics: a pragmatic perspective.* Nat Biotechnol, 2010. **28**(7): p. 695-709.
29. Yates, J.R., C.I. Ruse, and A. Nakorchevsky, *Proteomics by mass spectrometry: approaches, advances, and applications.* Annu Rev Biomed Eng, 2009. **11**: p. 49-79.
30. Olsen, J.V., S.E. Ong, and M. Mann, *Trypsin cleaves exclusively C-terminal to arginine and lysine residues.* Mol Cell Proteomics, 2004. **3**(6): p. 608-14.
31. Tsiatsiani, L. and A.J. Heck, *Proteomics beyond trypsin.* FEBS J, 2015. **282**(14): p. 2612-26.
32. Nagaraj, N., et al., *Deep proteome and transcriptome mapping of a human cancer cell line.* Mol Syst Biol, 2011. **7**: p. 548.
33. Choudhary, G., et al., *Multiple enzymatic digestion for enhanced sequence coverage of proteins in complex proteomic mixtures using capillary LC with ion trap MS/MS.* J Proteome Res, 2003. **2**(1): p. 59-67.
34. Swaney, D.L., C.D. Wenger, and J.J. Coon, *Value of using multiple proteases for large-scale mass spectrometry-based proteomics.* J Proteome Res, 2010. **9**(3): p. 1323-9.
35. Guo, X., et al., *Confetti: a multiprotease map of the HeLa proteome for comprehensive proteomics.* Mol Cell Proteomics, 2014. **13**(6): p. 1573-84.
36. Glatter, T., et al., *Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion.* J Proteome Res, 2012. **11**(11): p. 5145-56.
37. Kulak, N.A., et al., *Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells.* Nat Methods, 2014. **11**(3): p. 319-24.
38. Neverova, I. and J.E. Van Eyk, *Role of chromatographic techniques in proteomic analysis.* J Chromatogr B Analyt Technol Biomed Life Sci, 2005. **815**(1-2): p. 51-63.
39. Ritorto, M.S., et al., *Hydrophilic strong anion exchange (hSAX) chromatography for highly orthogonal peptide separation of complex proteomes.* J Proteome Res, 2013. **12**(6): p. 2449-57.
40. Lau, E., et al., *Combinatorial use of offline SCX and online RP-RP liquid chromatography for iTRAQ-based quantitative proteomics applications.* Mol Biosyst, 2011. **7**(5): p. 1399-408.
41. Boersema, P.J., S. Mohammed, and A.J. Heck, *Hydrophilic interaction liquid chromatography (HILIC) in proteomics.* Anal Bioanal Chem, 2008. **391**(1): p. 151-9.
42. Wang, Y., et al., *Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells.* Proteomics, 2011. **11**(10): p. 2019-26.
43. Batth, T.S., C. Francavilla, and J.V. Olsen, *Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics.* J Proteome Res, 2014. **13**(12): p. 6176-86.
44. Villen, J. and S.P. Gygi, *The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry.* Nat. Protocols, 2008. **3**(10): p. 1630-1638.
45. Macek, B., M. Mann, and J.V. Olsen, *Global and site-specific quantitative phosphoproteomics: principles and applications.* Annu Rev Pharmacol Toxicol, 2009. **49**: p. 199-221.

46. Olsen, J.V., et al., *Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap.* Mol Cell Proteomics, 2005. **4**(12): p. 2010-21.
47. Elias, J.E., et al., *Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations.* Nat Methods, 2005. **2**(9): p. 667-75.
48. Scigelova, M. and A. Makarov, *Orbitrap mass analyzer--overview and applications in proteomics.* Proteomics, 2006. **6 Suppl 2**: p. 16-21.
49. Glish, G.L. and R.W. Vachet, *The basics of mass spectrometry in the twenty-first century.* Nat Rev Drug Discov, 2003. **2**(2): p. 140-150.
50. Makarov, A., et al., *Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer.* Anal Chem, 2006. **78**(7): p. 2113-20.
51. Makarov, A. and E. Denisov, *Dynamics of Ions of Intact Proteins in the Orbitrap Mass Analyzer.* Journal of the American Society for Mass Spectrometry, 2009. **20**(8): p. 1486-1495.
52. Zubarev, R.A. and A. Makarov, *Orbitrap Mass Spectrometry.* Analytical Chemistry, 2013. **85**(11): p. 5288-5296.
53. Makarov, A., *Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis.* Analytical Chemistry, 2000. **72**(6): p. 1156-1162.
54. Kingdon, K.H., *A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures.* Physical Review, 1923. **21**(4): p. 408-418.
55. Zubarev, R.A. and A. Makarov, *Orbitrap mass spectrometry.* Anal Chem, 2013. **85**(11): p. 5288-96.
56. Makarov, A., E. Denisov, and O. Lange, *Performance evaluation of a high-field Orbitrap mass analyzer.* J Am Soc Mass Spectrom, 2009. **20**(8): p. 1391-6.
57. Michalski, A., et al., *Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes.* Mol Cell Proteomics, 2012. **11**(3): p. O111 013698.
58. Kelstrup, C.D., et al., *Rapid and Deep Proteomes by Faster Sequencing on a Benchtop Quadrupole Ultra-High-Field Orbitrap Mass Spectrometer.* J Proteome Res, 2014.
59. Michalski, A., et al., *Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer.* Mol Cell Proteomics, 2011. **10**(9): p. M111 011015.
60. Scheltema, R.A., et al., *The Q Exactive HF, a Benchtop Mass Spectrometer with a Pre-filter, High-performance Quadrupole and an Ultra-high-field Orbitrap Analyzer.* Mol Cell Proteomics, 2014. **13**(12): p. 3698-708.
61. W.E., S., *A Pulsed Mass Spectrometer with Time Dispersion.* Physical Review, 1946. **69**(11-12): p. 674-674.
62. Wiley, W.C. and I.H. McLaren, *Time-of-Flight Mass Spectrometer with Improved Resolution.* Review of Scientific Instruments, 1955. **26**(12): p. 1150-1157.
63. Moniatte, M., et al., *Characterisation of the heptameric pore-forming complex of the Aeromonas toxin aerolysin using MALDI-TOF mass spectrometry.* FEBS Letters, 1996. **384**(3): p. 269-272.
64. Mamyrin B.A., K.V.I., Shmikk D.V., Zagulin V.A. , *The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution.* Journal of Experimental and Theoretical Physics. **37**(1): p. 45.
65. O`Halloran G.J., F.R.A., Betts J.F. and Everett W.L., *Technical Documentarty Report.* The Bendix Corporation, Research Laboratory Division, 1964. **ASD-TDR-62-664**.
66. Dovonof A.F., C.I.V.a.L.V.V., Proceedings of 12th International Mass Sepctrometry Conference, 1991: p. 153.
67. Dawson, J.H.J. and M. Guilhaus, *Orthogonal-acceleration time-of-flight mass spectrometer.* Rapid Communications in Mass Spectrometry, 1989. **3**(5): p. 155-159.

68. Johnson, J.V., et al., *Tandem-in-space and tandem-in-time mass spectrometry: triple quadrupoles and quadrupole ion traps.* Analytical Chemistry, 1990. **62**(20): p. 2162-2172.
69. Gillet, L.C., A. Leitner, and R. Aebersold, *Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing.* Annu Rev Anal Chem (Palo Alto Calif), 2016. **9**(1): p. 449-72.
70. Wells, J.M. and S.A. McLuckey, *Collision-induced dissociation (CID) of peptides and proteins.* Methods Enzymol, 2005. **402**: p. 148-85.
71. Hunt, D.F., et al., *Protein sequencing by tandem mass spectrometry.* Proc Natl Acad Sci U S A, 1986. **83**(17): p. 6233-7.
72. Olsen, J.V., et al., *Higher-energy C-trap dissociation for peptide modification analysis.* Nat Methods, 2007. **4**(9): p. 709-12.
73. Syka, J.E., et al., *Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry.* Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9528-33.
74. Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides.* Biomed Mass Spectrom, 1984. **11**(11): p. 601.
75. Biemann, K., *Mass spectrometry of peptides and proteins.* Annu Rev Biochem, 1992. **61**: p. 977-1010.
76. Yang, Y.H., et al., *Low mass cutoff evasion with q(z) value optimization in ion trap.* Anal Biochem, 2009. **387**(1): p. 133-5.
77. Kocher, T., et al., *High precision quantitative proteomics using iTRAQ on an LTQ Orbitrap: a new mass spectrometric method combining the benefits of all.* J Proteome Res, 2009. **8**(10): p. 4743-52.
78. Mikesh, L.M., et al., *The utility of ETD mass spectrometry in proteomic analysis.* Biochim Biophys Acta, 2006. **1764**(12): p. 1811-22.
79. Steen, H. and M. Mann, *The ABC's (and XYZ's) of peptide sequencing.* Nat Rev Mol Cell Biol, 2004. **5**(9): p. 699-711.
80. Peterson, A.C., et al., *Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics.* Mol Cell Proteomics, 2012. **11**(11): p. 1475-88.
81. Gillet, L.C., et al., *Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis.* Mol Cell Proteomics, 2012. **11**(6): p. O111 016717.
82. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* Nat Biotechnol, 2008. **26**(12): p. 1367-72.
83. Cox, J., et al., *A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics.* Nat Protoc, 2009. **4**(5): p. 698-705.
84. Hein M.Y., S.K., Cox J., Mann M. , *Proteomic Analysis of Cellular Systems.* Handbook of Systems Biology, Academic Press, 2013. **Chapter 1**: p. 3-25.
85. Elias, J.E. and S.P. Gygi, *Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry.* Nat Methods, 2007. **4**(3): p. 207-14.
86. Bondarenko, P.V., D. Chelius, and T.A. Shaler, *Identification and Relative Quantitation of Protein Mixtures by Enzymatic Digestion Followed by Capillary Reversed-Phase Liquid Chromatography−Tandem Mass Spectrometry.* Analytical Chemistry, 2002. **74**(18): p. 4741-4749.
87. Cox, J., et al., *Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ.* Mol Cell Proteomics, 2014. **13**(9): p. 2513-26.
88. Geiger, T., et al., *Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins.* Mol Cell Proteomics, 2012. **11**(3): p. M111 014050.

89. Ong, S.-E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative.* Nat Chem Biol, 2005. **1**(5): p. 252-262.
90. Bantscheff, M., et al., *Quantitative mass spectrometry in proteomics: a critical review.* Anal Bioanal Chem, 2007. **389**(4): p. 1017-31.
91. Everley, R.A., et al., *Increasing Throughput in Targeted Proteomics Assays: 54-Plex Quantitation in a Single Mass Spectrometry Run.* Analytical Chemistry, 2013. **85**(11): p. 5340-5346.
92. Ong, S.E., et al., *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.* Mol Cell Proteomics, 2002. **1**(5): p. 376-86.
93. Geiger, T., et al., *Super-SILAC mix for quantitative proteomics of human tumor tissue.* Nat Methods, 2010. **7**(5): p. 383-5.
94. Hebert, A.S., et al., *Neutron-encoded mass signatures for multiplexed proteome quantification.* Nat Methods, 2013. **10**(4): p. 332-4.
95. Gygi, S.P., et al., *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.* Nat Biotechnol, 1999. **17**(10): p. 994-9.
96. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.* Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
97. Werner, T., et al., *Ion coalescence of neutron encoded TMT 10-plex reporter ions.* Anal Chem, 2014. **86**(7): p. 3594-601.
98. Mertins, P., et al., *iTRAQ labeling is superior to mTRAQ for quantitative global proteomics and phosphoproteomics.* Mol Cell Proteomics, 2012. **11**(6): p. M111 014423.
99. Hardt, M., et al., *Assessing the effects of diurnal variation on the composition of human parotid saliva: quantitative analysis of native peptides using iTRAQ reagents.* Anal Chem, 2005. **77**(15): p. 4947-54.
100. Choi, S., et al., *Comparative proteome analysis using amine-reactive isobaric tagging reagents coupled with 2D LC/MS/MS in 3T3-L1 adipocytes following hypoxia or normoxia.* Biochem Biophys Res Commun, 2009. **383**(1): p. 135-40.
101. Unwin, R.D., *Quantification of proteins by iTRAQ.* Methods Mol Biol, 2010. **658**: p. 205-15.
102. Ow, S.Y., et al., *iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly".* J Proteome Res, 2009. **8**(11): p. 5347-55.
103. Savitski, M.M., et al., *Delayed Fragmentation and Optimized Isolation Width Settings for Improvement of Protein Identification and Accuracy of Isobaric Mass Tag Quantification on Orbitrap-Type Mass Spectrometers.* Analytical Chemistry, 2011. **83**(23): p. 8959-8967.
104. Karp, N.A., et al., *Addressing accuracy and precision issues in iTRAQ quantitation.* Mol Cell Proteomics, 2010. **9**(9): p. 1885-97.
105. Mathur, R. and P.B. O'Connor, *Artifacts in Fourier transform mass spectrometry.* Rapid Commun Mass Spectrom, 2009. **23**(4): p. 523-9.
106. Rauniyar, N. and J.R. Yates, 3rd, *Isobaric labeling-based relative quantification in shotgun proteomics.* J Proteome Res, 2014. **13**(12): p. 5293-309.
107. Ow, S.Y., et al., *Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation.* Proteomics, 2011. **11**(11): p. 2341-6.
108. Ting, L., et al., *MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics.* Nat Meth, 2011. **8**(11): p. 937-940.
109. McAlister, G.C., et al., *MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes.* Anal Chem, 2014. **86**(14): p. 7150-8.
110. Sturm, R.M., C.B. Lietz, and L. Li, *Improved isobaric tandem mass tag quantification by ion mobility mass spectrometry.* Rapid Commun Mass Spectrom, 2014. **28**(9): p. 1051-60.
111. Thakur, S.S., et al., *Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation.* Mol Cell Proteomics, 2011. **10**(8): p. M110 003699.

112. Nagaraj, N., et al., *System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap.* Mol Cell Proteomics, 2012. **11**(3): p. M111 013722.
113. Han, J., et al., *Towards high peak capacity separations in normal pressure nanoflow liquid chromatography using meter long packed capillary columns.* Anal Chim Acta, 2014. **852**: p. 267-73.
114. Michalski, A., J. Cox, and M. Mann, *More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS.* J Proteome Res, 2011. **10**(4): p. 1785-93.
115. MP., L., Ann de Chim et de Phys, 1905(8): p. 245-288.
116. MP., L., Ann de Chim et de Phys, 1903(7): p. 289-384.
117. Makinen, M.A., O.A. Anttalainen, and M.E. Sillanpaa, *Ion mobility spectrometry and its applications in detection of chemical warfare agents.* Anal Chem, 2010. **82**(23): p. 9594-600.
118. Makinen, M., M. Nousiainen, and M. Sillanpaa, *Ion spectrometric detection technologies for ultra-traces of explosives: a review.* Mass Spectrom Rev, 2011. **30**(5): p. 940-73.
119. Vautz, W., et al., *Ion mobility spectrometry for food quality and safety.* Food Addit Contam, 2006. **23**(11): p. 1064-73.
120. Verkouteren, J.R. and J.L. Staymates, *Reliability of ion mobility spectrometry for qualitative analysis of complex, multicomponent illicit drug samples.* Forensic Sci Int, 2011. **206**(1-3): p. 190-6.
121. Lokhnauth, J.K. and N.H. Snow, *Solid phase micro-extraction coupled with ion mobility spectrometry for the analysis of ephedrine in urine.* J Sep Sci, 2005. **28**(7): p. 612-8.
122. Srebalus, C.A., et al., *Gas-Phase Separations of Electrosprayed Peptide Libraries.* Analytical Chemistry, 1999. **71**(18): p. 3918-3927.
123. Fernandez-Lima, F., et al., *Gas-phase separation using a trapped ion mobility spectrometer.* Int J Ion Mobil Spectrom, 2011. **14**(2-3).
124. Barnes, W.S., D.W. Martin, and E.W. McDaniel, *Mass Spectrographic Identification of the Ion Observed in Hydrogen Mobility Experiments.* Physical Review Letters, 1961. **6**(3): p. 110-111.
125. McDaniel, E.W., D.W. Martin, and W.S. Barnes, *Drift Tube-Mass Spectrometer for Studies of Low-Energy Ion-Molecule Reactions.* Review of Scientific Instruments, 1962. **33**(1): p. 2-7.
126. K. B. McAfee, J. and D. Edelson, *Identification and Mobility of Ions in a Townsend Discharge by Time-resolved Mass Spectrometry.* Proceedings of the Physical Society, 1963. **81**(2): p. 382.
127. Liu, X., et al., *Mapping the human plasma proteome by SCX-LC-IMS-MS.* J Am Soc Mass Spectrom, 2007. **18**(7): p. 1249-64.
128. McLean, J.A., et al., *Ion mobility–mass spectrometry: a new paradigm for proteomics.* International Journal of Mass Spectrometry, 2005. **240**(3): p. 301-315.
129. Myung, S., et al., *Development of high-sensitivity ion trap ion mobility spectrometry time-of-flight techniques: a high-throughput nano-LC-IMS-TOF separation of peptides arising from a Drosophila protein extract.* Anal Chem, 2003. **75**(19): p. 5137-45.
130. Distler, U., et al., *Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics.* Nat Methods, 2014. **11**(2): p. 167-70.
131. Zinnel, N.F., P.J. Pai, and D.H. Russell, *Ion mobility-mass spectrometry (IM-MS) for top-down proteomics: increased dynamic range affords increased sequence coverage.* Anal Chem, 2012. **84**(7): p. 3390-7.
132. Isailovic, D., et al., *Profiling of human serum glycans associated with liver cancer and cirrhosis by IMS-MS.* J Proteome Res, 2008. **7**(3): p. 1109-17.
133. Isailovic, D., et al., *Delineating diseases by IMS-MS profiling of serum N-linked glycans.* J Proteome Res, 2012. **11**(2): p. 576-85.

134. Gaye, M.M., et al., *Ion mobility-mass spectrometry analysis of serum N-linked glycans from esophageal adenocarcinoma phenotypes.* J Proteome Res, 2012. **11**(12): p. 6102-10.
135. Dwivedi, P., et al., *Metabolic profiling by ion mobility mass spectrometry (IMMS).* Metabolomics, 2008. **4**(1): p. 63-80.
136. Guo, X. and E. Lankmayr, *Multidimensional approaches in LC and MS for phospholipid bioanalysis.* Bioanalysis, 2010. **2**(6): p. 1109-23.
137. Kaplan, K., et al., *Monitoring dynamic changes in lymph metabolome of fasting and fed rats by electrospray ionization-ion mobility mass spectrometry (ESI-IMMS).* Anal Chem, 2009. **81**(19): p. 7944-53.
138. Fernandez-Lima, F.A., et al., *Petroleum crude oil characterization by IMS-MS and FTICR MS.* Anal Chem, 2009. **81**(24): p. 9941-7.
139. Li, Z., S.J. Valentine, and D.E. Clemmer, *Complexation of amino compounds by 18C6 improves selectivity by IMS-IMS-MS: application to petroleum characterization.* J Am Soc Mass Spectrom, 2011. **22**(5): p. 817-27.
140. E.A. Mason, E.W.M., *Transport Properties of Ions in Gases.* WILEY-VCH Verlag GmbH & Co. KGCaA Weinheim, 1988.
141. Mason E.A., c.H.W., *Mobility of gaseous ion in weak electric fields.* Ann. Phys. , 1958. **4**: p. 233-270.
142. Bush, M.F., et al., *Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology.* Analytical Chemistry, 2010. **82**(22): p. 9557-9565.
143. May, J.C., et al., *Conformational Ordering of Biomolecules in the Gas Phase: Nitrogen Collision Cross Sections Measured on a Prototype High Resolution Drift Tube Ion Mobility-Mass Spectrometer.* Analytical Chemistry, 2014. **86**(4): p. 2107-2116.
144. Jia, C., et al., *Site-Specific Characterization of d-Amino Acid Containing Peptide Epimers by Ion Mobility Spectrometry.* Analytical Chemistry, 2014. **86**(6): p. 2972-2981.
145. Mason, E.A. and H.W. Schamp, *Mobility of gaseous lons in weak electric fields.* Annals of Physics, 1958. **4**(3): p. 233-270.
146. Valentine, S.J., et al., *Toward plasma proteome profiling with ion mobility-mass spectrometry.* J Proteome Res, 2006. **5**(11): p. 2977-84.
147. Baker, E.S., et al., *An LC-IMS-MS Platform Providing Increased Dynamic Range for High-Throughput Proteomic Studies.* Journal of Proteome Research, 2010. **9**(2): p. 997-1006.
148. Baker, E.S., et al., *Ion mobility spectrometry-mass spectrometry performance using electrodynamic ion funnels and elevated drift gas pressures.* J Am Soc Mass Spectrom, 2007. **18**(7): p. 1176-87.
149. Clowers, B.H., et al., *Enhanced ion utilization efficiency using an electrodynamic ion funnel trap as an injection mechanism for ion mobility spectrometry.* Anal Chem, 2008. **80**(3): p. 612-23.
150. Ibrahim, Y.M., et al., *Improving Ion Mobility Measurement Sensitivity by Utilizing Helium in an Ion Funnel Trap.* Analytical Chemistry, 2014. **86**(11): p. 5295-5299.
151. Blase, R.C., et al., *Increased ion transmission in IMS: A high resolution, periodic-focusing DC ion guide ion mobility spectrometer.* International Journal of Mass Spectrometry, 2011. **301**(1–3): p. 166-173.
152. Gillig, K.J., et al., *An electrostatic focusing ion guide for ion mobility-mass spectrometry.* International Journal of Mass Spectrometry, 2004. **239**(1): p. 43-49.
153. Webb, I.K., et al., *Experimental Evaluation and Optimization of Structures for Lossless Ion Manipulations for Ion Mobility Spectrometry with Time-of-Flight Mass Spectrometry.* Analytical Chemistry, 2014. **86**(18): p. 9169-9176.

154. Hoaglund, C.S., et al., *Three-Dimensional Ion Mobility/TOFMS Analysis of Electrosprayed Biomolecules.* Analytical Chemistry, 1998. **70**(11): p. 2236-2242.
155. Lanucara, F., et al., *The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics.* Nat Chem, 2014. **6**(4): p. 281-294.
156. Giles, K., et al., *Applications of a travelling wave-based radio-frequency-only stacked ring ion guide.* Rapid Commun Mass Spectrom, 2004. **18**(20): p. 2401-14.
157. Thalassinos, K., et al., *Ion mobility mass spectrometry of proteins in a modified commercial mass spectrometer.* International Journal of Mass Spectrometry, 2004. **236**(1–3): p. 55-63.
158. Giles, K., J.P. Williams, and I. Campuzano, *Enhancements in travelling wave ion mobility resolution.* Rapid Commun Mass Spectrom, 2011. **25**(11): p. 1559-66.
159. Pringle, S.D., et al., *An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument.* International Journal of Mass Spectrometry, 2007. **261**(1): p. 1-12.
160. Shvartsburg, A.A. and R.D. Smith, *Fundamentals of Traveling Wave Ion Mobility Spectrometry.* Analytical Chemistry, 2008. **80**(24): p. 9689-9699.
161. Ruotolo, B.T., et al., *Ion mobility-mass spectrometry analysis of large protein complexes.* Nat Protoc, 2008. **3**(7): p. 1139-52.
162. Takebayashi, K., et al., *Application of ion mobility-mass spectrometry to microRNA analysis.* J Biosci Bioeng, 2013. **115**(3): p. 332-8.
163. Silva, J.C., et al., *Simultaneous qualitative and quantitative analysis of the Escherichia coli proteome: a sweet tale.* Mol Cell Proteomics, 2006. **5**(4): p. 589-607.
164. Silva, J.C., et al., *Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition.* Mol Cell Proteomics, 2006. **5**(1): p. 144-56.
165. Geromanos, S.J., et al., *Using ion purity scores for enhancing quantitative accuracy and precision in complex proteomics samples.* Anal Bioanal Chem, 2012. **404**(4): p. 1127-39.
166. Buryakov, I.A., et al., *A new method of separation of multi-atomic ions by mobility at atmospheric pressure using a high-frequency amplitude-asymmetric strong electric field.* International Journal of Mass Spectrometry and Ion Processes, 1993. **128**(3): p. 143-148.
167. Guevremont R., P.R., *FAIMS apparatus and method with ion diverting device*, in *National Research Council*. 2004: Canada.
168. Canterbury, J.D., et al., *Assessing the dynamic range and peak capacity of nanoflow LC-FAIMS-MS on an ion trap mass spectrometer for proteomics.* Anal Chem, 2008. **80**(18): p. 6888-97.
169. Cumeras, R., et al., *Review on ion mobility spectrometry. Part 1: current instrumentation.* Analyst, 2015. **140**(5): p. 1376-90.
170. Park, M.A., Kim, T., Stacey, C., Berg, C., *Ion guide for mass spectrometers*. 2008.
171. Park, M.A., *Apparatus and method for parallel flow ion mobility spectrometry combined with mass spectrometry*. 2010.
172. Michelmann, K., et al., *Fundamentals of trapped ion mobility spectrometry.* J Am Soc Mass Spectrom, 2015. **26**(1): p. 14-24.
173. Silveira, J.A., M.E. Ridgeway, and M.A. Park, *High resolution trapped ion mobility spectrometery of peptides.* Anal Chem, 2014. **86**(12): p. 5624-7.
174. Benigni, P. and F. Fernandez-Lima, *Oversampling Selective Accumulation Trapped Ion Mobility Spectrometry coupled to FT-ICR MS: Fundamentals and Applications.* Anal Chem, 2016.
175. Meier, F., et al., *Parallel Accumulation-Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device.* J Proteome Res, 2015. **14**(12): p. 5378-87.
176. de Godoy, L.M., et al., *Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast.* Nature, 2008. **455**(7217): p. 1251-4.

177. Smith, G.P., *Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface.* Science, 1985. **228**(4705): p. 1315-7.
178. Fields, S. and O. Song, *A novel genetic system to detect protein-protein interactions.* Nature, 1989. **340**(6230): p. 245-6.
179. Dunham, W.H., M. Mullin, and A.C. Gingras, *Affinity-purification coupled to mass spectrometry: basic principles and strategies.* Proteomics, 2012. **12**(10): p. 1576-90.
180. Blagoev, B., et al., *A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling.* Nat Biotechnol, 2003. **21**(3): p. 315-8.
181. Ranish, J.A., et al., *The study of macromolecular complexes by quantitative proteomics.* Nat Genet, 2003. **33**(3): p. 349-55.
182. Ewing, R.M., et al., *Large-scale mapping of human protein-protein interactions by mass spectrometry.* Mol Syst Biol, 2007. **3**: p. 89.
183. Malovannaya, A., et al., *Analysis of the human endogenous coregulator complexome.* Cell, 2011. **145**(5): p. 787-99.
184. Leitner, A., et al., *Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics.* Mol Cell Proteomics, 2010. **9**(8): p. 1634-49.
185. Chen, Z.A., et al., *Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry.* EMBO J, 2010. **29**(4): p. 717-26.
186. Ficarro, S., et al., *Phosphoproteome analysis of capacitated human sperm. Evidence of tyrosine phosphorylation of a kinase-anchoring protein 3 and valosin-containing protein/p97 during capacitation.* J Biol Chem, 2003. **278**(13): p. 11579-89.
187. Huttlin, E.L., et al., *A tissue-specific atlas of mouse protein phosphorylation and expression.* Cell, 2010. **143**(7): p. 1174-89.
188. Kim, S.C., et al., *Substrate and functional diversity of lysine acetylation revealed by a proteomics survey.* Mol Cell, 2006. **23**(4): p. 607-18.
189. Choudhary, C., et al., *Lysine acetylation targets protein complexes and co-regulates major cellular functions.* Science, 2009. **325**(5942): p. 834-40.
190. Ong, S.E., G. Mittler, and M. Mann, *Identifying and quantifying in vivo methylation sites by heavy methyl SILAC.* Nat Methods, 2004. **1**(2): p. 119-26.
191. Humphrey, S.J., S.B. Azimifar, and M. Mann, *High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics.* Nat Biotechnol, 2015. **33**(9): p. 990-5.
192. Rifai, N., M.A. Gillette, and S.A. Carr, *Protein biomarker discovery and validation: the long and uncertain path to clinical utility.* Nat Biotechnol, 2006. **24**(8): p. 971-83.
193. Antman, E.M., et al., *Cardiac-specific troponin I levels to predict the risk of mortality in patients with acute coronary syndromes.* N Engl J Med, 1996. **335**(18): p. 1342-9.
194. Barnidge, D.R., et al., *Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-Ms/MS using protein cleavage and isotope dilution mass spectrometry.* J Proteome Res, 2004. **3**(3): p. 644-52.
195. Oe, T., et al., *Quantitative analysis of amyloid beta peptides in cerebrospinal fluid of Alzheimer's disease patients by immunoaffinity purification and stable isotope dilution liquid chromatography/negative electrospray ionization tandem mass spectrometry.* Rapid Commun Mass Spectrom, 2006. **20**(24): p. 3723-35.

5. ACKNOWLEDGEMENT

I would like to deeply thank the following people, who contribute to this work:

First I would like to thank Prof. Dr. Matthias Mann for the great support, fruitful discussions and motivating words during my whole PhD. It is a pleasure working in your group. I wish you and your group all the best for the future!

Thanks to all my thesis advisory committee members Prof. Dr. Klaus Förstemann and Dr. Annette Michalski for your support and scientific input.

Next, I would like to thank my examination board members Prof. Dr. Jürgen Cox, PD Dr. Dietmar Martin, Prof. Dr. Karl-Peter Hopfner and Prof. Dr. Thomas Carell for being part of my thesis committee and for your support.

Thanks to Florian Meier for all the great discussions and help. I wish you all the best for your PhD and future career.

Thanks to Dan Itzhak for proof reading my thesis.

I want to thanks Eva Keilhauer for all the scientific and non-scientific talks and inputs. You are a great mother and Anna can be proud of you.

Thanks to all my old and new office mates Mario Oroshi, Philipp Geyer and Sophia Doll for the very nice and familiar atmosphere.

Thanks to Korbinian Mayr, Igor Paron and Gaby Sowa for great technical support.

Thanks to Alison Dalfovo and Theresa Schneider for all the administrative support.

Thanks to all members of the Mann group. I had a great time with you at work, conferences and at our retreats. There are a lot of memories, I never will forget. I wish you all the best for your future!

I also would like to thank a lot of people at Bruker, in particular Markus Lubeck, Oliver Raether, Niels Goedecker, Stephanie Kaspar-Schoenefeld for the great technical support.

Thanks to my friend from university Cornelia Brönner, Patricia Kammerer and Andrea Künzel for numerous nice lunch breaks. I am pretty sure all of you will finish up, soon.

Thanks to my little sister Kerstin Koehler for being there for me. You are a great sister, mother and artist. I am very proud of you!

Thanks to my daughter Emilie Koch. You make me happy and proud every day. I really enjoy the time with you. You are such a happy person that makes me smile every day.

Thanks to Heiner Koch for your love, support and motivation. Until now we had an amazing time and I am looking forward to the next 100 years spending with you!

Last, I would like to thank my parents for the great support and for always finding the right words. Without you this thesis would never exist.