Lab-on-a-tissue : optimization of on-tissue chemistry for improved mass spectrometry imaging

Heijs, B.P.A.M.

Chapter 1
Introduction

1.1 Introduction

In the modern age of personalized medicine and patient stratification, biomolecular markers and accompanying molecular profiling techniques have become deeply embedded in clinical research and molecular pathology (Cristofanilli et al., 2004; Hurvitz et al., 2013; Oviaño et al., 2016; Pecoraro et al., 2016). Specifically analytical chemistry methods are applied to pathological tissues in order to obtain the identity, quantity and localization of disease-specific molecular features. One of these techniques is mass spectrometry imaging (MSI), which is a mass spectrometry-based technique that allows the spatially-correlated analysis of biomolecular ions directly from thin tissue sections (Caprioli et al., 1997). This introduction will cover the basics of the mass spectrometry required for understanding the MSI research described in this thesis and to place it in the context of wider MSI research.

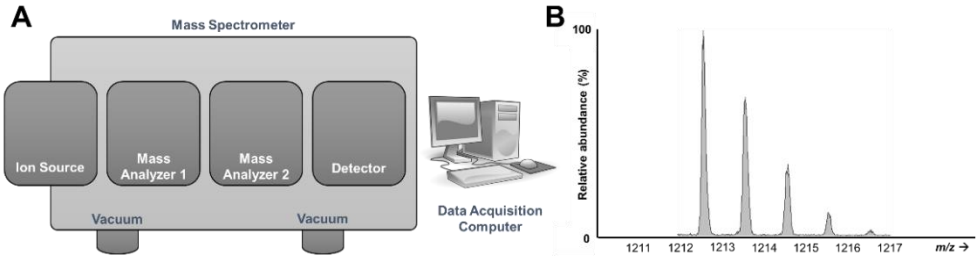

Figure 1-1: (A) A schematic overview of the layout of a mass spectrometer with its core components. From left to right, the ion source, one (or more) mass analyzers, a detector, and a data acquisition computer. **(B)** Example of a mass spectrum, with the horizontal axis representing the mass-to-charge ratio, and the vertical axis representing the (relative) abundance, or signal intensity.

1.2 Mass spectrometry

Mass spectrometry (MS) is an analytical chemistry technique that allows the detection, identification and quantification of ionized molecular species. Since the recording of the first mass spectrum in 1912 by Sir J.J. Thomson (Thomson, 1912), numerous technological advances have made MS one of the most sensitive and fast molecular analysis techniques available with a great variety of application areas including petrochemistry, forensic science, pharmaceutical research, food sciences and biomedical research (te Brake et al., 2015; Dekker et al., 2014; Esteve et al., 2015; Frese et al., 2017; Groeneveld et al., 2015; Maleki et al., 2016; Smith et al., 2008).

Mass spectrometers employ a range of different technologies but share the same basic layout (Figure 1-1A); ion source, one or multiple mass analyzers, and a detector and acquisition computer. In the ion source, the molecular analytes are converted from the liquid or solid phase into gas-phase ions. The mass analyzer separates the ions based on their mass-to-charge (*m/z*) ratio. Finally, the separated ions and their relative abundance, are recorded by the detector. The molecular content of the analyzed sample is visualized in a mass spectrum. Consisting of two perpendicular axes, a mass

spectrum is usually displayed with the x-axis representing the *m/z* or 'mass' in Daltons (Da) per unit charge, and the y-axis representing intensity of the signal (Figure 1-1B) (Glish and Vachet, 2003).

1.2.1 Ion source

A large number of different ionization techniques are available, each with different characteristics that make them more/less suitable to specific application areas. The choice of ionization method can have a large effect on the quantitative and qualitative outcome of the analysis. For example, the analysis of the metabolic content of serum by electrospray (ESI) MS, atmospheric pressure chemical ionization (APCI) MS, and matrix-assisted laser desorption/ionization (MALDI) MS led to the detection of different sets of small molecules (Nordström et al., 2008). Also, the linearity and lower limit of quantification for the same compound were shown to be heavily influenced by the choice of ionization method (Buse et al., 2014). Therefore, a careful understanding of the different ionization methods is essential to determine the application areas most suited to the ionization methods available within any single laboratory. Present day biomolecular MS utilizes soft ionization techniques that are capable of transferring large molecules, such as proteins, from the liquid or solid phase to the gas phase without excessive fragmentation. The two most common ionization techniques in biomolecular MS are ESI, and MALDI (Domon and Aebersold, 2006).

1.2.1.1 Electrospray ionization

ESI, first reported by Yamashita and Fenn in 1984 (Yamashita and Fenn, 1984) and later demonstrated to enable protein analysis by mass spectrometry (Fenn et al., 1989), is a continuous ionization mechanism, based on the nebulization of a dissolved sample in an atmospheric pressure environment. The analyte molecules are dissolved in an aqueous solution containing an organic solvent (e.g. methanol or acetonitrile) and, in the case of positive ion analysis, a small amount of acid (e.g. formic, or acetic acid). The sample solution is introduced into a needle, which is held at a high electrical potential relative to the mass spectrometer's sampling capillary. Upon leaving the needle, the solution experiences the strong electric field between the needle and the mass spectrometer's sampling capillary, and forms a Taylor cone from which small, highly charged droplets are released. The formation of molecular ions from the charged droplets is depicted in Figure 1-2 (Wilm, 2011).

The strong electric field between the needle and the sampling capillary attracts ions to the solution surface, forming the Taylor cone and a steady stream of charge-enriched droplets. Solvent evaporation then takes place (the ratio of the surface area to volume of the droplets is high and so evaporation takes place quickly). During solvent evaporation charge repulsion between the excess charges on the droplets leads to deformation; when localized repulsion exceeds the local surface tension of the droplet (Rayleigh instability), asymmetric fission occurs in which the smaller droplets carry an excess fraction of the charge. This process of solvent evaporation and

Rayleigh instability continues until only very small droplets remain. In the charge residue model, which is believed to be the dominant mechanism for the formation of multiply charged ions of large molecules, this process continues until a single, solvated, and multiply charged analyte ion is produced. In the ion evaporation model, which is believed to explain the formation of ions of small molecules, the (solvated) ions are desorbed out of the small droplets owing to the very strong electric field in the ion source (Awad et al., 2015; Beaudry et al., 1999; Fenn et al., 1990; Nordström et al., 2008).

Multiple charging and multiple charge states, e.g. lysozyme being detected in charge states +5 to +25, are characteristic of ESI (Krusemark et al., 2009; Valentine et al., 1997). This has two distinct benefits; (i) it brings the *m/z* of the protein ions into the *m/z* range that can be analyzed by most mass analyzers with higher mass resolution and sensitivity, and (ii) it enables peptide/protein ion fragmentation and the detection of complementary fragments, both of which promote their structural characterization (Wysocki et al., 2005).

The continuity of the ESI method makes it highly suitable for use in a hyphenated system where the mass spectrometer is coupled to a chromatographic separation, such as liquid chromatography (LC), or capillary electrophoresis (CE).

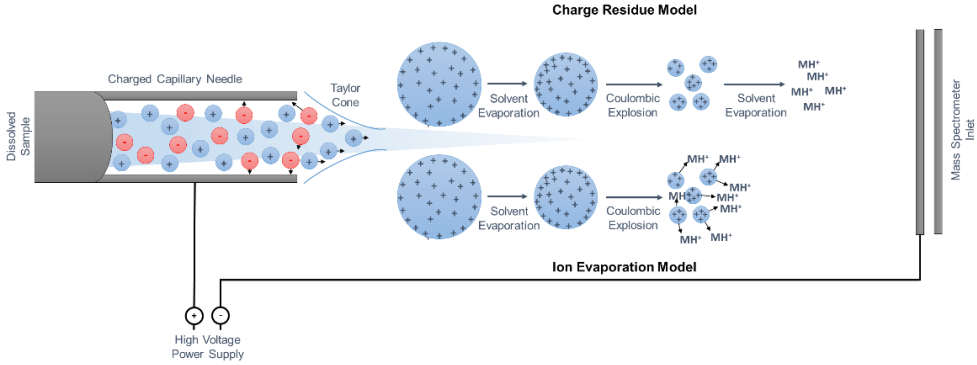

Figure 1-2: Schematic overview of the ion formation in electrospray ionization, visualizing both the charge residue model and the ion evaporation model.

1.2.1.2 Matrix-assisted laser desorption/ionization

In contrast to the continuity of ESI, MALDI is a pulsed ionization process. Developed by Karas and Hillenkamp (Karas and Hillenkamp, 1988; Karas et al., 1987) the MALDI method requires the analyte molecules to be embedded in an excess amount of a chemical matrix. The matrix is typically a small organic acid and a strong absorber of UV light. In a MALDI experiment a solution of the analyte is mixed with an excess of a matrix solution. During solvent evaporation the matrix crystalizes, incorporating the analyte (proteins) or coming in to close contact with the analyte (small molecules). The matrix crystals have been shown to segregate salts away from the analytes, thereby

effectively helping to de-salt the samples and improve detection sensitivity. Irradiation of the analyte doped matrix crystals with a pulsed UV laser leads to efficient production of molecular ions. MALDI involves the super excitation of the laser-absorbing matrix with the laser (MALDI requires laser fluence thresholds of 30 J/m^2 and pulse lengths <5 ns, so that the laser energy does not have sufficient time to dissipate during irradiation). Above laser fluence threshold the super heating causes an almost explosive phase transition, with rapid collective motion of the matrix substrate into the gas phase. The analyte molecules are brought into the gas phase by being entrained within this collective motion. This process of the explosive phase change leads to the formation of neutral and charged matrix and analyte molecules, as well as clusters. All molecules that form gas-phase ions, anions by deprotonation or cations by available protons (H^+) or metals (e.g. Na^+, K^+, Ag^+), are accelerated toward the mass analyzer by means of an electric (Figure 1-3) (Dreisewerd, 2014; Hanton et al., 1999; McDonnell and Heeren, 2007).

Careful selection of the matrix is key to the success of a MALDI-MS analysis, as the physiochemical properties of the MALDI matrix will affect its co-crystallization with the analytes, their co-desorption on laser irradiation and the efficiency of the ionization process (Dreisewerd, 2003, 2014; Zenobi and Knochenmuss, 1998). Over the years many compounds have been empirically determined to be suitable MALDI matrices; some for very specific applications, such as 1H-pteridine-2,4-dione (lumazine), 2-aminobenzaminde, 2-aminobenzoic acid, β-carboline alkaloids, and graphene (Calvano et al., 2010; Dong et al., 2010; Nonami et al., 1997; Shroff and Svatoš, 2009; Smargiasso et al., 2012), while other matrices are more generally applicable. For example, α-cyano-4-hydroxycinnamic acid (CHCA), 2,5-dihydroxybenzoic acid (DHB), and sinapinic acid (SA) are matrices that are commonly used for the analysis of peptides, lipids, and proteins in positive ion mode. 9-aminoacridine (9-AA) is commonly used for the negative ion analysis of small molecules.

The experimental conditions of the ion source and the MALDI sample preparation determine the accessible ion yield of the analyte. The ion source instrumentation includes how the laser interacts with the sample. Optimization of the irradiation conditions can affect the ion flux towards the mass analyzer; too low of a laser fluence (energy per illuminated surface) will result in poor MALDI signals, too high of a laser fluence will cause extensive in-source fragmentation of the analyte molecules (Dreisewerd et al., 1995). The focus of the laser is also important, in that it must be tuned along with the laser fluence to ensure that the on-target laser fluence is above threshold but not too high that it leads to extensive fragmentation. Furthermore, the MALDI matrix crystals of the sample preparation can lead to heterogeneity in the sample preparation, mostly observed in dried-droplet sample preparations. In order to effectively average this crystal-derived heterogeneity older MALDI instruments used larger spot sizes (Dreisewerd et al., 1995; McDonnell and Heeren, 2007).

In the last decade, MALDI instruments have moved toward fast, tightly focused laser beams, and even include structured laser spots. Qiao *et al.* demonstrated that with careful control of the laser fluence tightly focused laser beams delivered more signal per unit area (Qiao et al., 2008), and Holle *et al.* demonstrated that increased signal intensity and reduced analyte depletion was observed if a larger laser beam was converted into an array of tightly focused point-illumination beams (Holle et al., 2006). These technologies have now combined to offer high speed, high sensitivity MALDI ionization systems. Part of the driving force for the development of such high speed, high sensitivity, tightly focused laser beam systems has been MALDI-MSI, in which the laser beam focus, sensitivity and throughput helps determine the achievable spatial resolution.

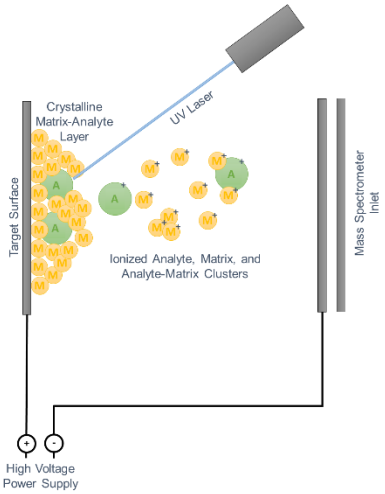

Figure 1-3: A schematic overview of the matrix-assisted laser desorption/ionization mechanism.

1.2.2 Mass analyzer and ion detection

The mass analyzer separates the ions based on their *m/z* ratio, after which the ions are detected at their *m/z*. In mass spectrometry two terms are often used to define the quality or the performance characteristics of a mass analyzer: (i) the mass accuracy refers to the error between the measured *m/z* and the calculated, true *m/z* of a compound, and (ii) the mass resolving power, or resolution, defines the ability of the mass analyzer to distinguish between two peaks with very similar *m/z*. A broad range of technologies are available that combine different characteristics in terms of mass accuracy, mass resolving power, sensitivity (ion transmission and detection efficiencies), *m/z* range, speed, tandem mass spectrometry, ease of hyphenation with e.g. liquid separations, ease-of-use, cost (purchase & maintenance). Such mass spectrometry systems include: (i) quadrupole mass analyzers, (ii) time-of-flight (TOF) mass analyzers, (iii) magnetic or electrostatic sector mass analyzers, (iv) quadrupole ion trap mass analyzers, (v) Orbitrap mass analyzers, and (vi) Fourier transform (FT)

ion cyclotron resonance (ICR) mass analyzers. For a detailed comparison of the systems the interested reader is directed toward the review by Himmelsbach (Himmelsbach, 2012). Here we will focus on the technologies mostly used in MALDI-MSI.

1.2.2.1 Quadrupole mass analyzer

The quadrupole mass analyzer consists of a set of four parallel metal rods to which a combination of radio frequency (RF) and direct current (DC) voltages are applied. The stability of ion motion in a quadrupole is governed by the Mathieu parameters

$$a = \frac{8zU}{mr_o^2 f^2} \; q = \frac{4zV}{mr_o^2 f^2}$$

where U, V and f are the DC potential, RF potential and RF frequency, respectively, and r_o is the quadrupole's field radius (Kero et al., 2005).

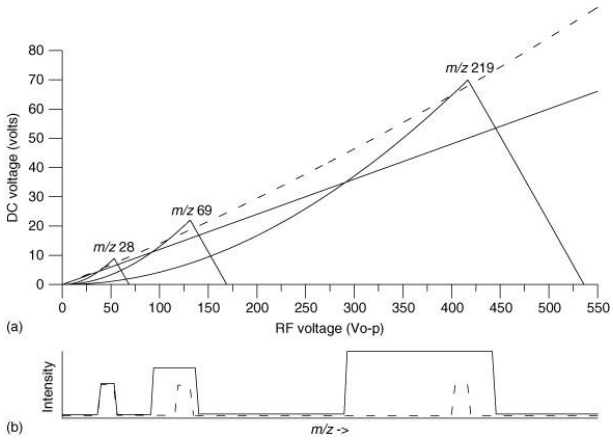

Figure 1-4: (A) Mathieu stability diagram for ions of *m/z* 28, 69, and 219. **(B)** Simulated mass spectra showing these ions as the RF and DC voltages are scanned along the dotted and solid scan lines (in A). Figure recreated from Kero *et al.* (Kero et al., 2005).

The quadrupole is essentially a mass filter, in which the applied RF and DC voltages determine which ions have a stable trajectory through the quadrupole. Figure 1.4 shows the stability diagram for ions of *m/z* 28, 69 and 219. It can be seen that when the DC voltage (Mathieu parameter *a*) is set to zero, ions of a wide mass range have a stable motion and are passed through the quadrupole. In this mode the quadrupole acts as an ion guide. With increasing DC voltage, the stability region decreases in size, and thus the *m/z* range of ions with a stable trajectory decreases. By ramping the RF and DC voltages synchronously it is possible to perform a mass analysis, by sequentially only permitting the transmission of selected ions (and in which the DC voltage determines the mass resolution of the mass analysis). The solid and dashed lines in Figure 1.4 represent two such scan lines, a lower resolution and higher

resolution mass scan. The figure shows idealized mass spectra resulting from these two scans. It can be seen that the values of the RF and DC voltage can be selected to transmit a single ion of interest, which is used extensively in quadrupole-time-of-flight (Q-TOF) type instruments for selection of the precursor ion (Kero et al., 2005).

1.2.2.2 Time-of-flight mass spectrometry

A TOF mass analyzer (Figure 1.5A) is a pulsed system and so has been extensively coupled to MALDI ion sources. It is also used with ESI, in which pulsed ion gates are used to introduce packets of ions from the continuous ESI beam. In a MALDI-TOF system the ion source consists of a target plate (containing the MALDI sample) separated by a short distance from a series of counter electrodes. A strong electric field between the target plate and the first counter electrode, 20-30 kV over a distance of a few mm is common, means any ion formed by the MALDI process is immediately accelerated into the mass spectrometer. MALDI generated ions are normally singly charged and so the kinetic energy of the ion (in eV) is simply equal to the accelerating potential of the MALDI-TOF ion source. All ions then enter the time-of-flight region with the same kinetic energy. The time-of-flight (T_f) is defined as the time between the MALDI laser pulse and the moment the ion hits the detector. In its most simple form this can be written as:

$$T_f = \frac{L}{v} \tag{1}$$

$$E_k = \frac{mv^2}{2} = zV \tag{2}$$

where L is the length of the flight tube, v is velocity, E_k is the kinetic energy, m is mass, z is charge, and V s the accelerating potential of the ion source. A substitution results in formula (3), which shows that when L and V are constant the time-of-flight is proportional to the square root of the m/z ratio. Higher molecular weight molecules will have longer flight times compared to small molecules (Cotter, 1992; Guilhaus et al., 1997).

$$T_f = L\sqrt{\frac{m}{z}}\sqrt{\frac{1}{2V}} \propto \sqrt{\frac{m}{z}} \tag{3}$$

Such a simple TOF mass spectrometer results in a moderate mass resolving power (Radionova et al., 2016). Over the last decades multiple developments have further increased the achievable mass resolution. These include:

(i) pulsed extraction – otherwise known as time-lag focusing or delayed extraction, uses an additional pulsed electric field in the ion source to time focus the ions on to the detector (Brown and Lennon, 1995; Wiley and McLaren, 1955).

(ii) reflectron – effectively an electrostatic mirror that reflects ions onto the detector. Faster ions move deeper into the mirror before being reflected, and thus travel a longer distance. Careful design of the reflectron field enables the ions to be time-focused onto the detector, thereby improving mass resolution (Mamyrin et al., 1973).

(iii) longer flight tubes – as can be seen in equation (3) a long flight tube theoretically leads to longer flight times, and increased time between ions of different mass (so greater mass resolution). However, this is offset by dispersion of the ions during the time-of-flight. In order to increase the flight path without losing mass resolution, a focusing electrostatic lens was introduced in the flight tube (Vestal and Juhasz, 1998).

(iv) orthogonal acceleration – orthogonal acceleration TOF systems decouple the mass analysis from the ion generation event. In this manner the variability in initial position and energy of ion formation, which compromises the theoretical performance of axial TOF systems, has less effect on the performance of an orthogonal acceleration TOF system (Dawson and Guilhaus, 1989).

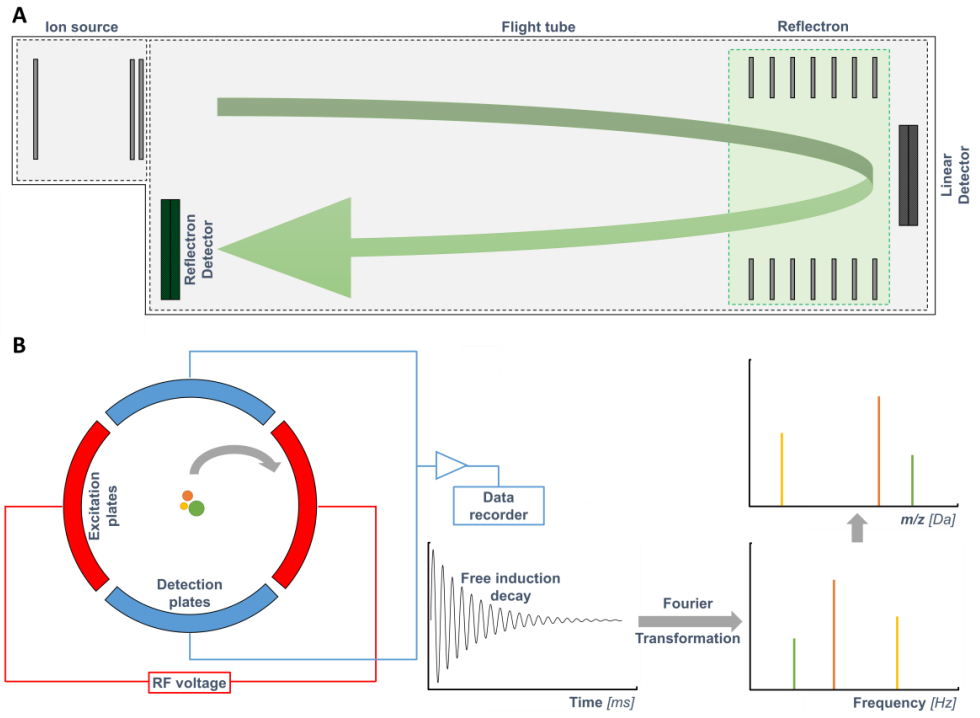

Figure 1-5: A schematic overview of **(A)** a linear TOF mass analyzer including a reflecton and reflectron detector for ion focusing, and **(B)** the process of FTICR mass analysis and signal conversion.

Modern TOF systems have many characteristics favorable to MSI, including rapid scan speeds, high sensitivity, good mass resolution and mass accuracy (up to several ppm and 50k resolving power for thin layer samples using a highly focused laser beam), wide mass range, and good dynamic range ($10^2 – 10^3$).

1.2.2.3 Fourier Transform mass spectrometry

While TOF-MS will provide decent performance, based on mass accuracy and resolving power, for some applications higher performance is required. Fourier Transform MS (FTMS) is capable of providing ultra-high mass resolution and accuracy, as well as a much higher dynamic range. It may not be necessary for every application to analyze samples using ultra-high resolution and accuracy, however, in some applications in which the samples consist of complex analyte mixtures containing many near-isobaric species, FTMS is essential in order to obtain accurate and confident molecular identifications (Scigelova et al., 2011).

In FTMS, the most common mass analyzers are the Fourier Transform Ion Cyclotron Resonance (FTICR) (Comisarow and Marshall, 1974) and the Orbitrap (Makarov, 2000). And, although their appearance is very different, both FTICR and Orbitrap analyzers use an image current detection system and the application of a mathematical Fourier transformation to convert the time domain transients, produced by the image current, into a mass spectrum (Scigelova et al., 2011). FTICR is based on the rotational (cyclotron) motion of an ion in a magnetic field. In FTICR-MS (Figure 1-5B), the ions are guided into a Penning trap placed in the center of a strong magnetic field (usually provided by a superconducting magnet). There, the ions are trapped by an electric field provided by two trapping electrodes at both ends of the trap. The ions trapped in the magnetic field rotate in a plane perpendicular to the magnetic field, at their cyclotron frequency (minus a small frequency due to the trapping field and the presence of other ions in the trap). At room temperature the cyclotron frequency of an ion is very small, e.g. in a 9.4 T field an ion of *m/z* 1000 has a cyclotron radius of 0.077 mm (Marshall et al., 1998). In order to detect the ions, their cyclotron motion is excited by the application of a radio-frequency (RF) electric field at the ion's cyclotron frequency. All ions of the same *m/z* ratio are excited, coherently, within the ICR trap, increasing their cyclotron radius. After the excitation field is switched off, the ICR-motion-excited ions remain at their excited radii. The coherent motion of these ions induces an image charge as they move past a detector plate; using a pair of opposing detector plates leads to the creation of an image current oscillating between the plates, at the ion's cyclotron frequency. This image current is converted to a voltage, amplified and digitized, and forms the free-induction-decay (FID) of the FTICR experiment. A Fourier transform is used to convert the ICR signal from the time domain to the frequency domain. In the absence of an electric field or the presence of other ions (space charge) an ion moving freely in a magnetic field, rotates with a characteristic cyclotron frequency, given by

$$\omega = \frac{q}{m}B \tag{4}$$

in which ω is the ICR frequency, q represents charge, m is the mass of the ion and B is the magnetic field strength (Nikolaev et al., 2016). In the presence of space charge the ion moves with two distinct circular motions, the reduced cyclotron motion (at a frequency slightly below that of the free cyclotron frequency) and the magnetron motion (a slow and low frequency secular motion). A number of different equations have been reported for accurate calibration of FTICR datasets (Barry et al., 2013; Marshall, 1998). The two most common calibration equations are those reported by Francl and Ledford (Francl et al., 1983; Ledford et al., 1984).

$$Francl: \frac{m}{z} = \frac{A}{\omega + B}$$

$$Ledford: \frac{m}{z} = \frac{A}{\omega} + \frac{B}{\omega^2}$$

The advantage of the Francl equation is that the numerator is dependent only on the magnetic field strength, and B is essentially a correction factor for the near-constant shift in frequency due to space charge. Accordingly, FTICR mass spectra that have been calibrated using the Francl equation can be recalibrated using a single internal lock mass to correct for differences in space charge from scan to scan (Barry et al., 2013; Easterling et al., 1999).

Ultra-high resolution and mass accuracy of FTICR mass spectrometers are highly dependent on the strength of the magnetic field provided by the superconductive magnets. For this reason, progress in FTICR-MS resolution/mass accuracy has benefited from developments in superconductive magnet technology (Nikolaev et al., 2016).

1.2.3 Tandem mass spectrometry

The availability of a wide variety of ionization, mass analysis and detection methods has made mass spectrometry a powerful tool for the characterization of biomolecules. In order to assign definitive identities to the detected ions it is necessary to fragment the ions into structurally informative fragments. Structural information is obtained by selecting and isolating single ion species from the sample, cause them to fragment, and analyze the masses of the fragments (de Hoffmann, 1996). This is commonly referred to as tandem mass spectrometry (MS/MS), as it involves at least two stages of mass analysis. The first stage of mass analysis is used to select and isolate the precursor ion. Following the fragmentation event (discussed in more detail below) the fragment ions undergo the second stage of mass analysis where they are separated, based on their

m/z, before reaching the detector. Depending on the instrumental setup MS/MS can be performed either in space by coupling two mass analyzers; in time by performing the two events in the same mass analyzer, utilizing an ion storage device; or in a combination of space and time, in so-called hybrid instruments (de Hoffmann, 1996).

Table 1-1: Characteristics of commonly applied fragmentation methods for tandem MS of peptides and proteins.

Fragmentation method	Peptide/protein fragment ions	Satellite ions	Retains labile PTMs	MALDI / ESI	Instru-ments	References
collision-induced dissociation (CID)	b, y		(-)	Both	TOF, FTMS, IT	Frese et al., 2011; Wells and McLuckey, 2005
higher-energy collisional dissociation (HCD)	a, b, c, x, y, z	d, v, w	(-)	Both	TOF, FTMS, IT	Frese et al., 2011; de Graaf et al., 2011
electron-transfer dissociation (ETD)	c, y, z+1, z+2	w	(+)	ESI	FTMS, IT	Han et al., 2008; Qi and Volmer, 2017
negative electron transfer dissociation (NETD)	a, x		(+)	ESI	FTMS, IT	Coon et al., 2005; Riley et al., 2015
electron-detachment dissociation (EDD)	a, x		(+)	ESI	FTMS, IT	Budnik et al., 2001; Ganisl et al., 2011
infrared multi-photon dissociation (IRMPD)	b, y		(-)	Both	FTMS	Reilly, 2009; Vasicek et al., 2011

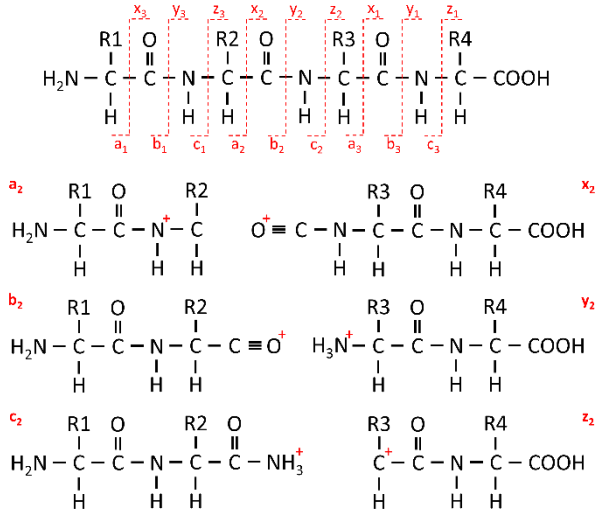

Figure 1-6: Overview of the different fragment ions produced during peptide fragmentation. The top panel shows the fragmentation of the peptide backbone, the lower panel shows the ions formed during fragmentation.

A common example of MS/MS in space is performed using a Q-TOF type instrument, in which the precursor ion is isolated with a quadrupole mass analyzer. The isolated ions are then accelerated into a collision cell (typically another multipole in a separate vacuum chamber held at higher pressure) where they will collide with an inert gas (e.g. N_2). The collision converts some of the kinetic energy of the precursor ion into internal energy. If the increase in internal energy is sufficient it will cause the precursor ions to fragment in a process termed collision-induced dissociation (CID), pioneered by Jennings, and McLafferty in the late 1960s (Jennings, 1968; McLafferty et al., 1973). In a Q-TOF setup, the newly formed fragments enter a TOF section for mass analysis and detection. By switching the DC component of the quadrupole off/on, a Q-TOF instrument can be rapidly switched between MS and MS/MS. Such switching is now routinely performed for data dependent MS/MS, in which the most intense ions detected in the MS scan are selected for isolation in the quadrupole, and fragmented by CID before fragment ion analysis.

MS/MS in time is normally employed in trap-based MS platforms (e.g. ion traps and FTICR). The fragmentation step occurs following a chain of events: (i) all ions are guided to the mass analyzer, (ii) all but the selected ions are ejected from the trap, (iii) the ion's flight path is excited by using an auxiliary RF field at the frequency of the ion's secular motion in the trap, (iv) a collision gas is introduced into the trap, collisions between the excited ions and gas molecules cause fragmentation, (v) mass analysis of the fragments (de Hoffmann, 1996).

Many fragmentation methods are currently available, each with its specific set of fragmentation characteristics and application areas, nevertheless the basic principle of fragmenting selected precursor ions in order to obtain structural information remains the same. In the tandem MS analysis of peptides, the choice of fragmentation method will affect the type of fragment ions detected in the MS/MS spectrum. Roepstorff & Fohlman proposed the nomenclature for the different fragment ions which is still used to date, albeit after modifications by Johnson *et al.* (Johnson et al., 1987; Roepstorff & Fohlman, 1984) (Figure 1-6 & Table 1-1).

1.3 Mass spectrometry imaging

The mass spectrometric analysis of sample surfaces was made available with the introduction of secondary ion MS (SIMS). However, the introduction of MALDI-MS provided the breakthrough for the analysis of large biomolecules. Both SIMS and MALDI-MS are capable of acquiring mass spectra from discrete locations, which makes them suitable technologies for imaging applications. The first *macromolecular* mass spectrometry imaging (MSI) analysis of biological samples was performed in 1997 by Caprioli *et al.* (Caprioli et al., 1997). This first publication showed the distribution of several proteins and peptides in various biological tissues following analysis by spatially correlated MALDI-MS. While the early work in the Caprioli lab focused on the development of software to automate the spatially correlated tissue analysis by MALDI-

MS, it was quickly recognized that the MSI technology was of great value for biomedical and clinical research (Chaurand et al., 1999; Stoeckli et al., 1999, 2001).

The principle behind every MSI analysis is the acquisition of spatially correlated mass spectra from discrete spots on a sample surface (Figure 1-7). The single mass spectra, acquired following a Cartesian coordinate system, are assembled into a data cube in which the X- and Y-axes represent the X- and Y-coordinates of the image, the Z-axis represents the *m/z* ratio, and the values at each voxel the intensities. By selecting a specific *m/z*, representing an analyte, one can plot the intensity as a function of position, to display the relative abundance of the analyte throughout the analyzed sample (McDonnell and Heeren, 2007).

1.3.1 MSI modalities

Depending on the application and the analytical requirements, a variety of MSI modalities are available (Figure 1-8). Secondary ion mass spectrometry (SIMS, Figure 1-8B) is available for the very high spatial resolution (~50 nm) analysis of atomic ions and small molecules (Slodzian et al., 1992). In SIMS, a focused (primary) ion beam is fired at the sample surface. Upon impact, part of the primary ions' kinetic energy is transferred to the atoms and ions on sample surface. If their resulting kinetic energy exceeds the binding energy of the substrate, the atoms and ions are released (Delcorte and Garrison, 2000). The analysis of larger molecules by SIMS suffers from poor sensitivity, as a result of the decrease in secondary ion emission for higher *m/z* species (Touboul et al., 2004). In desorption electrospray ionization (DESI, Figure 1-8C), the first MSI modality capable of analysis under ambient conditions, features of both ESI and desorption ionization methods are combined. An electrospray needle, directed at the sample, is used to create a focused electrospray by applying a high electrical potential to the spray solution. Like in a conventional ESI source, droplet formation, desolvation and acceleration is assisted by a gas flow (Takáts et al., 2004). During the brief interaction of the charged droplets with the target surface, analytes present on the surface are transferred to the mass spectrometer inlet following the droplet pick-up model (Venter et al., 2006). DESI-MSI is very well suited for small molecule and lipid analysis, but similar to SIMS, has a low sensitivity analyzing large biomolecules from tissue samples. One of the major advantages of DESI, and other ambient ionization methods, is that it does not require sample preparation for the analysis of analytes from either solid, frozen, liquid, or gaseous samples (Takáts et al., 2004, 2005). The success of DESI-MSI was followed by the rapid development of many novel ambient ionization methods. A thorough review of a large number of ambient ionization methods is provided by Wu *et al.* (Wu et al., 2013). Here a short description of the more established techniques is provided.

In laser ablation electrospray ionization (LAESI, Figure 1-8D) mass spectrometry, a focused mid-infrared laser irradiates discrete spots on the sample surface exciting OH vibrations of water molecules in the sample, causing desorption of the analyte

molecules from the sample surface. The desorbed molecules are post-ionized by capturing them in an electrospray jet, positioned 1-3 cm above the sample surface. LAESI is performed under atmospheric pressure in ambient conditions, and does not require extensive sample preparation as it uses water in the sample as the laser-absorption matrix. This MSI modality has been applied for the analysis of a multitude of analyte classes, ranging from small molecules to peptides and proteins (Nemes and Vertes, 2007; Wu et al., 2013).

Nano-DESI (Figure 1-8E) and liquid extraction surface analysis (LESA, Figure 1-8F) are closely related MSI modalities as they both use liquid microjunction surface-sampling (LMJ-SSP) to transfer analytes from the sample surface to the MS. LMJ-SSP is based on a set of capillaries that are in close proximity of each other and to the sample surface. A solvent flow connects the capillaries and sample surface through a liquid microjunction, allowing the localized extraction of analytes from the sample surface (Van Berkel et al., 2002). The configuration of the capillaries is what sets apart nano-DESI and LESA. In a typical LESA setup, a concentric capillary is placed just above the sample surface in a perpendicular fashion. Solvent will flow through the outer capillary, forming a liquid junction with the sample surface, allowing analyte extraction from the sample surface. The inner capillary is used to aspirate the dissolved analytes and guide them to an ESI source where the analytes are ionized and enter the MS, as illustrated in Figure 1-8F (Kertesz and Berkel, 2010). The nano-DESI setup is based on the same principles, but uses two separate capillaries (one solvent supplying capillary, and one aspirating capillary) that move over the sample in conjunction, allowing lower flow rates and smaller liquid bridges (resulting in better lateral resolution) (Lanekoff et al., 2013). Both these MSI modalities operate under atmospheric pressure and ambient conditions and find their applications in the analysis of small molecules, pharmaceuticals, lipids, and (proteolytic) peptides (Van Berkel et al., 2008; Eikel et al., 2011; Rao et al., 2013).

Due to its versatility, the most popular MSI modality is MALDI-MSI (Figure 1-8A); while most ambient MSI modalities are limited to the analysis of small molecules and lipids, MALDI-MSI is routinely used for the analysis of small molecules, lipids, *N*-linked glycans, (proteolytic) peptides, intact proteins, and more recently also proteoforms (Dilillo et al., 2017; Ellis et al., 2016; Kriegsmann et al., 2016; Liu et al., 2013a; Lou et al., 2016a; Powers et al., 2013). Most ambient MSI modalities have in common that, in order to analyze a sample, hardly any sample preparation is required. In MALDI-MSI the class of analytes that is eligible for analysis is determined mostly by the sample preparation method.

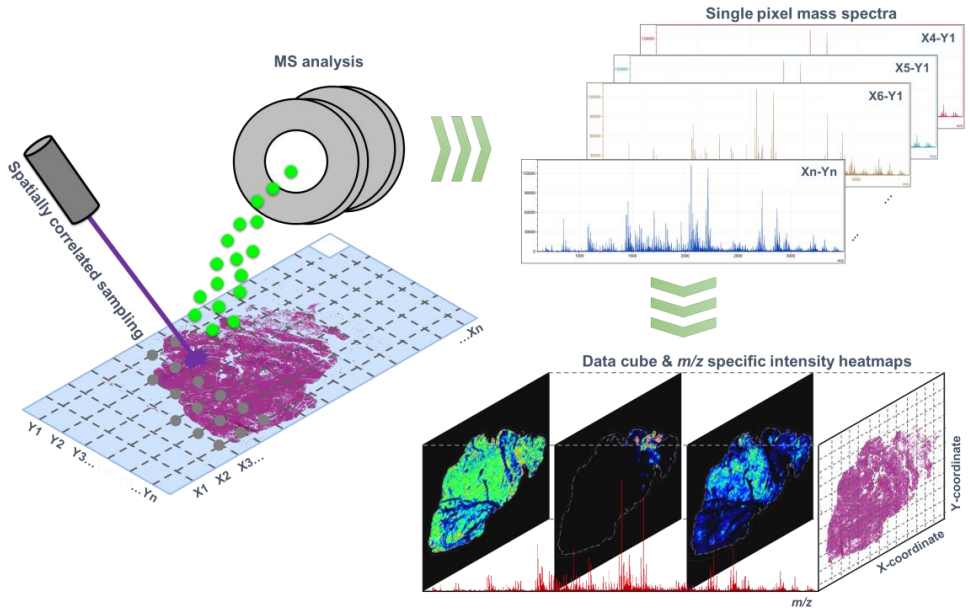

Figure 1-7: Schematic overview of the principle of mass spectrometry imaging (MSI).

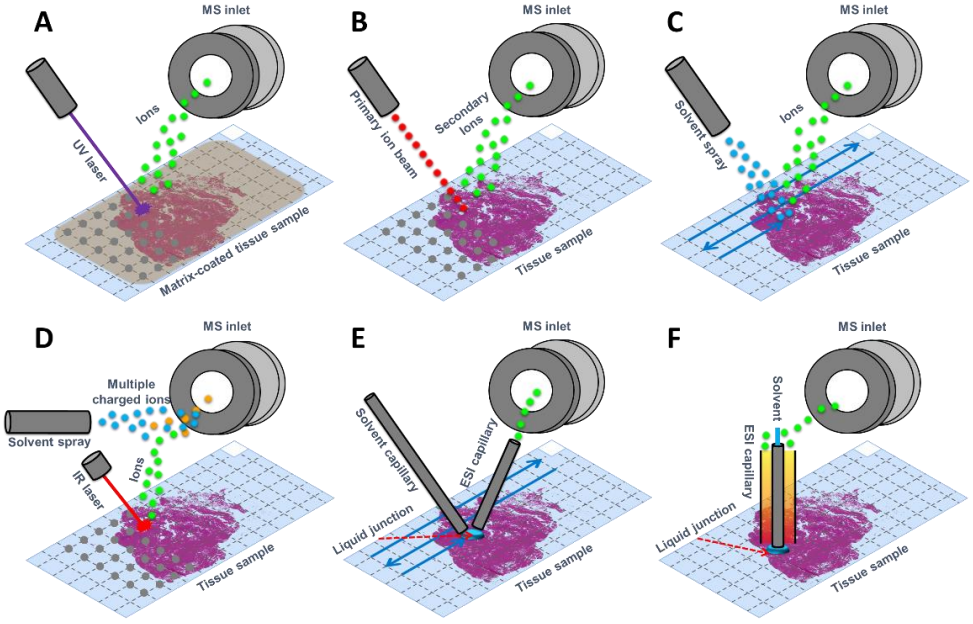

Figure 1-8: Schematic representations of various MSI modalities: **(A)** matrix-assisted laser desorption/ionization (MALDI), **(B)** secondary ion mass spectrometry (SIMS), **(C)** desorption electrospray ionization (DESI), **(D)** laser-assisted electrospray ionization (LAESI), **(E)** nano-flow desorption electrospray ionization (nano-DESI), and **(F)** liquid extraction surface analysis (LESA). Adapted from Addie *et al.* (Addie et al., 2015).

1.4 Sample preparation for MALDI-MSI

1.4.1 Tissue fixation and sampling

MALDI-MSI is commonly used for the analysis of biological tissue samples. While MALDI-MSI analysis of insect and plant tissue have been reported (Grassl et al., 2011; Kaftan et al., 2014; Kaspar et al., 2011; Khalil et al., 2015; Peukert et al., 2012; Sturtevant et al., 2016), MALDI-MSI is more commonly used for biomedical research involving the analysis of human or animal model tissues (Balluff et al., 2010; Carreira et al., 2015; Dekker et al., 2014). MALDI-MSI is usually performed on thin tissue sections of thickness 4-20 µm, obtained from either fresh frozen, heat stabilized, or chemically fixed (most commonly using formalin) tissues (Goodwin, 2012; Thomas and Chaurand, 2014). In order to maintain tissue integrity during the process of tissue sectioning, tissues are often embedded in an embedding material. For different applications, a great variety of embedding materials are available; paraffin, optimal cutting temperature (OCT), aqueous solutions of gelatin, and/or carboxymethyl cellulose (CMC) (Casadonte and Caprioli, 2011; Chen et al., 2009a; Schwartz et al., 2003; Stoeckli et al., 2007). The different tissue fixation and embedding methods each come with their specific set of benefits and limitations (Tables 1-2 & 1-3).

For MALDI analyses, performed in vacuum or intermediate pressure ion sources, the tissue samples are mounted on a conductive surface (the conductive surface is needed for the definition of the electric extraction field, which transfers ions to the mass analyzer). The conductive surface also helps dissipate the charge formed by photoelectrons after illumination with the UV laser. Ineffective dissipation of the charge will result in local distortions of the electric field in the sampled area, and cause mass shifts of the detected ions (Frankevich et al., 2003; Ibáñez et al., 2007; Knochenmuss, 2004). Steel target plates were used in early MALDI-MSI experiments. However, mounting tissues on glass coated with a transparent and conductive material (e.g. indium tin-oxide (ITO)) is now routine as it enables the post-analysis co-registration with histological images (Amstalden van Hove et al., 2010; Chaurand et al., 2004; Chughtai et al., 2012). A common problem in histopathology is the detachment of tissue sections from the glass slides during staining procedures in which the samples are repeatedly submerged in a series of staining solutions. In order to increase the adherence of the tissue sections to the glass slide, the slide is coated using a strong polycationic polymer of L-lysine residues, poly-L-lysine. The cationic sites in the L-lysine polymer interact with the anionic sites in the tissues, strengthening the adherence between the glass and the tissue. This method, first implemented by Zavalin *et al.*, is commonly implemented in the workflow for protein analysis by MALDI-MSI (Zavalin et al., 2012).

1.4.2 Tissue washing

The chemical complexity of a tissue section can lead to poor sensitivity in MALDI-MSI analyses. Tissue washes were introduced to reduce the chemical complexity and

improve detection sensitivity by selectively removing interfering compounds from the tissue section, whilst retaining the spatial distribution of the analytes of interest. Tissue washing was first introduced for the MALDI-MSI analysis of intact proteins, and improved measurement sensitivity by removing physiological salts and a large part of the lipid content present in the tissue (Lemaire et al., 2006; Schwartz et al., 2003). Early tissue washing procedures for protein analysis usually contained several steps by

Table 1-2: Various tissue fixation methods and relevant characteristics for MSI.

Tissue fixation methods	In situ/ Ex vivo	Ischemia time	Enzymatic activity	Effect on histology	References
Post-mortem freezing	Ex vivo	Minutes	Paused	Freezing artefacts, post-mortem degradation	Hattori et al., 2010
In-situ/funnel freezing	In vivo	None	Paused	Freezing artefacts, post-mortem degradation	Hattori et al., 2010; Mulder et al., 2016
Focussed microwave irradiation	In vivo	None	Terminated	Bubble formation; increased tissue fragility	Sugiura et al., 2014
Heat-induced stabilization	Ex vivo	Minutes	Terminated	Force induced deformation / flattening	Mulder et al., 2016; Sturm et al., 2013; Svensson et al., 2009
Formalin fixation	Ex vivo	Minutes-hours	Terminated	Isotope masking	Thavarajah et al., 2012

Table 1-3: Various tissue embedding compounds and relevant characteristics for MSI.

Embedding media	Chemical basis	Dissolves in	Fixation methods	Storage conditions	Potential MSI artefact	References
Optimal Cutting Temperature (OCT)	poly-ethylene polymers	Water	Cryo	Frozen	Signal suppression / Polymer contamination	Cazares et al., 2015; Groseclose and Castellino, 2013
Gelatin	Collagen	Water	Cryo	Frozen	Interference in high mass analysis	Altelaar et al., 2005; Dong et al., 2016
Carboxymethyl cellulose (CMC)	Poly-saccharide	Water	Cryo	Frozen	Polymer contamination in high mass analysis	Kawamoto, 2003
Paraffin	Alkane/ Organic	Aromatic organic solvents	Formalin fixed	Room temp.	Loss of metabolites & lipids	Porter et al., 2017
Poly[N-(2-hydroxy-propyl)-methacryl-amide] (pHPMA)	HPMA polymer	Water	Cryo	Frozen	Polymer contamination in high mass analysis	Strohalm et al., 2011

either rinsing or submersing the samples in a variety of organic solutions (e.g. ethanol, isopropanol, chloroform, acetone, or hexane). Later, also washes with water, or aqueous buffers were added to increase the sensitivity of the analysis of hydrophobic compounds (e.g. membrane proteins, specific classes of lipids), and proteins after on-tissue proteolytic digestion (Angel et al., 2012; Enthaler et al., 2013a; Grey et al., 2009; Groseclose et al., 2007; Nicklay et al., 2013; Thomas et al., 2013). Also, water-based tissue washes were successful in the clean-up of tissue sections by removal of polymers (e.g. polyethylene glycol (PEG), and polyethylene vinyl (PEV)) originating from embedding media, such as OCT (Cazares et al., 2009).

The early tissue washing protocols were developed on (mouse or rat) brain tissue. The brain contains many distinct regions with profound difference in their molecular composition, which simplifies the assessment of molecular delocalization, or lateral diffusion. However, with the introduction of MALDI-MSI in clinical research, studies were set up with tissues of different origin (e.g. other organs, tumors) and methods optimized for brain tissue proved insufficient or incompatible. Several tissue treatment optimization studies exemplify the improvement in analysis quality that can be gained from the tedious effort of method optimization (Enthaler et al., 2013a; Lou et al., 2016b; Martin-Lorenzo et al., 2015).

Tissue washing is usually omitted when it comes to analyzing small molecules, as the tissue washing procedures are commonly used to remove small molecules as interferants in the analysis of other molecular classes. However, Shariatgorji *et al.* have proposed a tissue cleanup procedure based on pH-controlled buffers for the targeted analysis of small molecule drugs by MALDI-MSI. In this particular study, tissues were washed in a pH-controlled, and concentration adjusted buffer in which the target compounds were insoluble and the interfering endogenous compounds were removed, improving the measurement sensitivity for various pharmaceutical compounds (Shariatgorji et al., 2012).

1.4.3 On-tissue digestion

In bottom-up proteomics, proteins are identified by analyzing and fragmenting proteolytic fragments, produced by an in-solution digestion with a proteolytic enzyme (Altelaar et al., 2012). Similar to this approach, proteolytic enzymes can be applied directly to tissue sections to digest the proteins *in-situ*. This on-tissue digestion approach was first explored by Shimma *et al.* in 2006, who applied trypsin to a fresh frozen mouse brain section using a chemical inktjet printer and performed on-tissue MALDI-MS/MS to identify proteolytic peptides directly from the tissue (Shimma et al., 2006). In 2007, the on-tissue digestion approach was applied by Groseclose *et al*. to perform MALDI-MSI of proteolytic peptides (Groseclose et al., 2007). In this pioneering work the analysis of proteolytic peptides by MALDI-MS/MS was used to identify the proteolytic peptides in a similar fashion presented by Shimma *et al.*. Moreover, the distributions of the proteolytic peptides in the tissue, obtained by MALDI-MSI of the

proteolytic peptides, was used to assign identities to *m/z* features detected in an intact protein MALDI-MSI analysis. MALDI-MS/MS directly from tissue sections has, since then, been successfully performed by a number of groups (Groseclose et al., 2007; Schober et al., 2011; Shimma et al., 2008). However, the lack of chromatographic separation and the small amount of tissue present in each pixel make it a very challenging approach. Ion mobility separation (IMS) MALDI-MSI has been performed to compensate for the lack of chromatographic separation, and has proven itself a useful tool for on-tissue feature identification (McLean et al., 2007; Stauber et al., 2010; Trim et al., 2008). However, the greatest number of peptide identity assignments have been achieved through mass matching the MALDI-MSI peak list with peptide identifications obtained through LC-MS/MS analysis of tissue extracts (Casadonte et al., 2014; Schober et al., 2011). High-resolution, accurate mass (HRAM) MS is essential in order to obtain accurate and confident peptide identifications from the LC-MS/MS analysis and to resolve the complex mixture of (near) isobaric proteolytic peptides created by the on-tissue proteolytic digestion (Groseclose et al., 2007; Schober et al., 2012). Despite the challenge that the complexity of the sample imposes on the experiment, the analysis of proteolytic peptides does enhance the proteomic information gained from a single tissue section, by enabling the analysis of larger proteins that are otherwise difficult to analyze by MALDI-MSI (Groseclose et al., 2007). An additional benefit of on-tissue proteolytic digestion is that it allows the analysis of sections originating from FFPE tissues, and as many clinical biobanks consist of FFPE tissues, the ability to obtain spatially correlated proteomic information from these samples is invaluable. A (partial) reversal of the formalin fixation by an antigen retrieval step, followed by a proteolytic digestion 'releases' the peptides and makes them eligible for MALDI-MSI analysis (Djidja et al., 2009b; Groseclose et al., 2008; Lemaire et al., 2007; Ronci et al., 2008).

Trypsin is often the enzyme of choice for bottom-up proteomics and on-tissue digestion MALDI-MSI analyses, nevertheless other proteases, such as pepsin, elastase, Arg-C, Lys-C, and Lys-N, as well as combinations of proteases, such as trypsin/Lys-C, have also been reported and have proven to be beneficial for proteome coverage (Angel et al., 2012; Enthaler et al., 2013a; Grey et al., 2009; Groseclose et al., 2007; Nicklay et al., 2013; Thomas et al., 2013). Although the majority of reports concerning on-tissue digestion MALDI-MSI describes the use of proteolytic enzymes, the pioneering work by Groseclose *et al.* also demonstrated the ability of the endoglycosidase Peptide *N*-glycosidase F (PNGase F) to perform *in-situ* protein deglycosylation (Groseclose et al., 2007). Later, Powers *et al.* demonstrated the ability to perform *N*-linked glycan MALDI-MSI after on-tissue digestion with PNGase F on both fresh frozen and FFPE tissues (Powers et al., 2013, 2014). As *N*-linked glycans are not affected by the chemistry of the formalin fixation, *N*-glycan MALDI-MSI can be performed from FFPE tissues both with and without the prior treatment with an antigen retrieval method (Holst et al., 2016; Briggs et al., 2017; Powers et al., 2013, 2015).

In bottom-up proteomics disulfide bond reduction in conjunction with alkylation is routinely used to help denature the proteins and thereby increase their accessibility to the proteases. Green-Mitchell *et al.* have adapted this approach for direct use on tissue sections, in order to improve and confirm the detection of insulin in pancreatic tissue sections (Green-Mitchell et al., 2011).

1.4.4 On-tissue derivatization

Chemical derivatization, the alteration of the chemical properties of a molecule by adding a (functional) chemical group, is often applied in MS-based research to increase the measurement sensitivity for the analysis of compounds that suffer from poor ionization (Liu and Hop, 2005; Melikian et al., 1999). More recently, this strategy has been translated to enable the *in-situ* analysis of derivatized analytes by MALDI-MSI (Chacon et al., 2011; Franck et al., 2009). Chemical derivatization of small molecules for MALDI-MSI analysis is typically not only used to improve the ionization efficiency, but also to increase the analyte mass to prevent suppression, or masking, by MALDI matrix peaks that are commonly found in the same *m/z* range. On-tissue derivatization is commonly used in MALDI-MSI for the analysis of amino acids, neurotransmitters, and amino metabolites, and has been applied in studies on cancer and neurodegenerative disorde rs (Shariatgorji et al., 2014; Esteve et al., 2016; Toue et al., 2014).

Whilst on-tissue derivatization is mostly used for the analysis of small molecules, an *in situ* derivatization method for protein and peptide N-termini has also been reported. This method was reported to improve the *in-situ* identification of on-tissue digested proteins (Franck et al., 2009). Furthermore, the *in-situ* derivatization of sialylated *N*-glycan species was reported by Holst *et al.*, and resulted in: (i) stabilization of sialic acids during MALDI-MSI analysis, (ii) reduction of the ionization bias induced by the negative charge of the sialic acids in the positive ion analysis, and (iii) differentiation between α2,3- and α2,6-linked sialic acid, which are implied in tumor progression and resistance to chemotherapy (Holst et al., 2016).

1.4.5 Matrix application

A robust and homogeneous MALDI matrix deposition is key to a high quality MALDI-MSI analysis. The choice of matrix, solvent system, and the deposition method determine which the analyte class may be analyzed, as well as the sensitivity and spatial resolution with which they may be examined (Goodwin, 2012; Thomas and Chaurand, 2014b). In the first MALDI-MSI analyses, described by Caprioli *et al.*, matrix was applied by rinsing the tissues in a saturated matrix solution, resulting in severe lateral diffusion of the analyte molecules. The same publication describes matrix deposition using electrospray resulting in an improvement of the determination of the analyte location (Caprioli et al., 1997). Since these initial experiments many matrix deposition methods have been developed, which can be divided into three groups: (i) spotting methods, (ii) spray-based methods, and (iii) solvent-free methods.

The spotting methods are based on the droplet matrix application approach that is commonly used in sample preparation protocols for regular MALDI-MS. An array of pico-liter-sized droplets is spotted on to the tissue surface using a chemical printer, resulting in matrix spots with a diameter of ~200 µm. The small spotting volumes, e.g. ~10 pL per spot per iteration, and the inter-spot spacing decreases the risk of molecular delocalization. Other than the lower risk of molecular delocalization, spot-based methods are known for their superior S/N ratio, and number of detected analytes compared to other matrix deposition methods. One of the major limitations of the spot-based methods is the relatively low spatial resolution, which is limited by the droplet pitch and is defined during the matrix deposition (Aerni et al., 2006; Végvári et al., 2010).

Spray-based matrix application has the advantage that the matrix is deposited in a continuous manner, which enables the definition of the spatial resolution based on the size of the matrix crystals (~20 µm). Several automated sprayer systems are available, either based on a pneumatically-assisted spray (TM Sprayer (HTX Technologies, Carrboro, NIC, USA), SunCollect (SunChrom GmbH, Friedrichsdorf, Germany), iMatrix spray (Stoeckli and Staab, 2015)), or vibration-based nebulization (ImagePrep (Bruker Daltonics, Bremen, Germany) of the matrix solution (Thomas and Chaurand, 2014).

More recently, solvent-free matrix deposition by sublimation was introduced for the MALDI-MSI analysis of small molecules and lipids (Hankin et al., 2007; Puolitaival et al., 2008). The absence of solvent results in very small matrix crystal sizes, which allows for extremely high spatial resolution analyses by MALDI-MSI (~5 µm) (Chaurand et al., 2011). Initially, sublimation was only used for small molecules and lipids due to the limited molecular extraction capabilities of the method. However, sublimation in combination with a matrix recrystallization step allowed the analysis of intact proteins up to 30 kDa at spatial resolutions as low as 10 µm (Yang and Caprioli, 2011).

1.5 Clinical mass spectrometry imaging

In the two decades since its introduction, MSI has established itself as an invaluable tool in many scientific areas, including pharmaceutical research, forensic science, and clinical research (Aichler and Walch, 2015; Atkinson et al., 2007; Balluff et al., 2010; Groeneveld et al., 2015; Reyzer et al., 2003; Wolstenholme et al., 2009). One of the key characteristics of MALDI-MSI, resulting in its widespread application in clinical research is the non-destructive nature of the method, allowing post-MSI histological staining and co-registration of the histological image with the MSI data. This multimodal MSI-histological imaging approach allows a direct investigation of how the underlying molecular composition varies with the tissue's histology.

1.5.1 Biomarker discovery

The discovery of disease- or phenotype-specific biomolecular markers or profiles is a research area in which the capabilities of MSI are commonly exploited. Especially in

cancer, the histological landscape is extremely complex and heterogeneous and often contains a multitude of different cell types, including tumor cells with variable levels of differentiation, surrounding normal tissue such as adipose or connective tissue, tumor vasculature, and infiltrating immune cells. Moreover, when comparing tissue specimens from multiple patients the inter-patient variation will add to the complexity, as the cellular composition of the tumor microenvironment and the differentiation state of the tumor cells will vary with each patient. The multimodal MSI-histological imaging data enables a retrospective, digital extraction of molecular profiles from specific regions of interest, a process called virtual microdissection. By carefully selecting the regions of interest by expert pathologists, based on the clinical question, the virtual microdissection reduces the histological heterogeneity, and thereby enables a meaningful comparison of the molecular profiles from histologically homogeneous regions within and between patients. Once specific molecular features have been identified for the relevant areas they can be compared with clinical patient data (e.g. disease prognosis, treatment response, metastasis) in specialized statistical tests (Addie et al., 2015; Aichler and Walch, 2015; Lou et al., 2016b). The described methodology has been used to identify that the absence of the functional mitochondrial polypeptide cytochrome C oxidase 7A2 (COX7A2) serves as a prognostic marker for poor survival in cohort of esophageal adenocarcinoma (fresh frozen, n = 39). A result that was confirmed through a validation study applying immunohistochemical (IHC) stainings of the identified protein biomarkers to an independent tissue set (FFPE, n = 102) (Elsner et al., 2012). Moreover, an MSI-based follow-up study revealed a positive correlation between the absence of functional COX7A2 and response to neo-adjuvant chemotherapy with cisplatin. This finding was later confirmed in work on esophageal cell lines treated with a COX7A2-specific siRNA, which resulted in defective mitochondria causing an increased sensitivity to the cisplatin therapy (Aichler et al., 2013). These studies are excellent examples of the capabilities of MALDI-MSI in clinical research, although it must be noted that these studies were based on the analysis of intact proteins from fresh frozen tissue sections. The analysis of fresh frozen tissue samples often hampers the size of the analyzed cohorts as most tissue archives consist of FFPE tissues.

A tissue microarray (TMA) consists of an assembly of needle core biopsies, in which the selection of relevant and comparable regions from different patients is performed during TMA construction, and thus prior to MSI data acquisition. The advantages of TMA analysis by MALDI-MSI include the increased throughput and a reduced experimental variance. Several examples of large scale clinical studies including TMA analyses by MALDI-MSI are available. These include an early study by Djidja *et al.*, in which peptide classifiers for various tumor grades of pancreatic adenocarcinoma were constructed, based on the on-tissue digestion MALDI-MSI of a TMA of FFPE biopsies from 30 patients (Djidja et al., 2010). More recently, Kriegsmann *et al.* were successful in subtyping patients diagnosed with non-small cell lung cancer (NSCLC) into

adenocarcinoma (ADC) and squamous cell carcinoma (SqCC), an important classification for treatment stratification. The classifier was built using the data from a training TMA that contained needle core biopsies of 110 ADC and 98 SqCC patients, and utilized 339 distinct *m/z* features. The application of the classifier to the MALDI-MSI data from a validation TMA containing 58 ADC and 60 SqCC samples resulted in the correct classification of 117 cases (accuracy 99.1%). Several differential expressed proteins, including known proteins used in current IHC classification (CK5, CK7), as well as novel markers (CK15, HSP27), were identified and validated by MS/MS and IHC (Kriegsmann et al., 2016).

1.5.2 Molecular histology

It is well established that the progression of some cancers can be described as a process of clonal evolution that starts with a single aberrant cell that, driven by (epi)genetic and/or metabolic changes, ultimately evolves into a heterogeneous tumor containing various subclones of the initial tumor clone (Dalerba et al., 2011; Gerlinger et al., 2012; Greaves and Maley, 2012). Typically, most biomarker discovery studies are based on the molecular characterization of homogenized tissue samples, ideally after microdissection of tumor cell-rich regions. While some of the tumor subpopulations may have distinct morphological features, others cannot be discerned on the basis of histology alone due to similar or overlapping morphologies. As these subpopulations are thought to drive tumor progression and influence the patient's prognosis, their elucidation is of high importance to increase our knowledge on the evolution of cancer (Greaves and Maley, 2012; Wu et al., 2010). MSI has been shown to be able to uncover the molecular intratumor heterogeneity, otherwise obscured by the lack of characterizing morphological features. In this case tissue annotation is not performed by visual inspection of the tissue's histology, but by a statistical segmentation based on the similarity of the tissue's biomolecular content (Deininger et al., 2008; Jones et al., 2011). A more recent publication by Balluff *et al.* displays the capability of MALDI-MSI to unravel the biomolecular intra-tumor heterogeneity and define distinct tumor subpopulations purely on the basis of molecular information obtained in the MSI experiment. Correlation to clinical patient information resulted in the identification of clinically relevant tumor subpopulations (Balluff et al., 2015).

1.6 Sarcoma

Soft tissue sarcoma (STS) is a class of uncommon tumors that arise in mesenchymal tissues, such as bone, cartilage, and connective tissues (e.g. muscle and fat). While STS can occur anywhere in the human body, approximately 75% occurs in the extremities (commonly in the thigh), 10% in the trunk wall, and another 10% in the retroperitoneum. The combination of their limited occurrence, 1% of all malignant tumors, and a large number of histological subtypes (> 50) often with overlapping morphology, results in challenges regarding diagnosis and treatment (Fletcher et al., 2013; Taylor et al., 2011). Firstly, their rarity results in a very low probability for a

general pathologist to encounter many of the STS subtypes. Secondly, the overlapping morphology, demonstrated intra-tumor heterogeneity and the only sparse availability of discriminating diagnostic markers complicate tumor subtyping (Willems et al., 2010a).

Previously, different sarcoma subtypes have been studied using MALDI-MSI. Caldwell *et al.* were the first to publish the potential of MSI in sarcoma research. MALDI-MSI of proteins was performed on a variety of STS tissues, including "fibrosarcoma", leiomyosarcoma, liposarcoma, undifferentiated pleomorphic sarcoma, malignant peripheral nerve sheath tumor and synovial sarcoma, in order to establish biomolecular differences between low-grade STS, high-grade STS and skeletal muscle control tissue. Several proteins, including previously characterized markers for other tumor types (calgizzarin, calcyclin, macrophage inhibitory factor, and calgranulin) as well as multiple histones, were found to differentiate high- and low-grade STS. Other proteins, including several myosin light-chains, appeared completely absent in tumor tissues, while vimentin was exclusively detected in the STS tissues. Five of the markers established by MSI, were successfully validated by IHC on FFPE tissues and showed similar expression patterns compared to MSI results (Caldwell et al., 2005). Additionally, Caldwell *et al.* studied the capability of MSI to establish tumor margins of an undifferentiated pleomorphic sarcoma. Using MALDI-MSI of intact proteins they were able to extend the tumor margin beyond the capabilities of histological examination through the detection of differential, disease-like, molecular signatures in regions that were considered non-diseased after histological examination. IHC validation of the MSI markers revealed a mixture of non-diseased cells and transforming tumor cells, morphologically indistinguishable from the non-diseased cells, stretching beyond the histologically determined tumor margins (Caldwell et al., 2006). Willems *et al.* studied the lipid and protein content of myxofibrosarcoma (MFS) and myxoid liposarcoma (MLS) using MALDI-MSI. Lipid MSI analysis revealed that the presence of various triglycerides (TG) was associated with low grade MLS, and the presence of different phosphatidylcholines (PC) was indicative of high grade MLS. These findings confirmed previous NMR-based findings that demonstrated the ability of histology-specific lipid analysis to classify different liposarcoma subtypes (Willems et al., 2010a). Additionally, biomolecular intratumor heterogeneity of intermediate-grade MFS was demonstrated and confirmed the multinodular tumor pattern, based on clonal selection previously observed on the genetic level in low-grade MFS (Jones et al., 2011; Willems et al., 2010a). These methods for the investigation of molecular intratumor heterogeneity were applied in a biomarker discovery study on chondrosarcoma, and demonstrated that different histological variants of central and peripheral chondrosarcoma were distinguishable by specific sets of proteins, including Vimentin, a well-established generic marker of sarcomas (Jones et al., 2013). Lou *et al.* published a study in which the ability of MALDI-MSI to distinguish between four types of high-grade sarcoma (osteosarcoma (OS), undifferentiated pleopmorphic sarcoma (UPS), myxofibrosarcoma (MFS), and leiomyosarcoma (LMS)) was demonstrated. A total of

twenty proteins were found as diagnostic markers for the different tumor types, an additional nine proteins showed prognostic value as they were associated with overall patient survival (Lou et al., 2016a). A similar study was performed investigating a metabolic marker for STS. Inositol cyclic phosphate and carnitine were established as potential generic prognostic markers for STS patients (Lou et al., 2017).

1.6.1 Myxoid liposarcoma

MLS is a malignant myxoid STS, which accounts for 5% of all adult cases of STS. Representing a third of the cases, it is the second most common type of liposarcoma (LPS) (Fletcher, 2014). MLS is often found in the extremities of young adults, and is treated by surgical resection (Antonescu et al., 2001; Fletcher, 2014). The histology of MLS consists of a mixture of cells, including uniform round mesenchymal cells, oval shaped primitive, nonlipogenic mesenchymal cells and small signet-ring lipoblasts in a myxoid ECM with a characteristic capillary vasculature (Fletcher et al., 2013). Specific, highly dense regions containing undifferentiated round cells with little cytoplasm and no ECM, termed 'round cell' regions are associated with high metastatic potential and a general poor prognosis. Other indicators for poor prognosis are the presence of necrotic tissue, the male gender, high age, and multifocal disease (Antonescu et al., 2001; Nishida et al., 2010; de Vreeze et al., 2010; Hoffman, et al., 2012; Fletcher et al., 2013). The myxoid ECM, clearly visible in conventional light microscopy, consists of a variety of secreted proteins and polysaccharides, and determines the physical properties of the tissue (Graadt van Roggen et al., 1999). The myxoid ECM is characteristic for various types of epithelial and mesenchymal tumors. In epithelial tumors a myxoid ECM is often a secondary phenomenon, while in mesenchymal tumors it is often considered an intrinsic part of the tumor (Willems et al., 2010a). Albeit that the myxoid ECM is an intrinsic part of the tumor entity it is not considered directional in diagnosis. Myxoid STSs contains a variety of morphologically similar tumors with varying clinical behavior, ranging from truly benign to highly malignant (Graadt van Roggen et al., 1999; Willems et al., 2010b).

The chimeric transcription factor oncogene, FUS-DDIT3 is characteristic to MLS (>95% of all cases), and is the result of a reciprocal translocation t(12;16)(q13;p11). In this translocation, the 5' half of the *FUS* gene is fused with the complete reading frame of the *DDIT3* gene. The remaining 5% of the cases harbor a similar translocation, t(12;22)(q13;p12), which fuses the *EWSR1* gene to the *DDIT3* gene. The exact mechanism of oncogenesis remains unknown, although it is hypothesized that the chimeric fusion gene acts as an aberrant transcriptional factor, stimulating cell proliferation whilst inhibiting adipogenic differentiation (de Graaff et al., 2016; Powers et al., 2010).

1.7 Aims of the thesis

MSI is a useful tool for the investigation of pathologies in which established histopathological methods may not always be conclusive. In this thesis, the aim was to develop novel enzyme-based on-tissue chemistry methods for an improved characterization of the biomolecular landscape of tissues. These methods were developed to further implement MSI in clinical research in general, and were applied to aid the biomolecular characterization of tumor progression in MLS.

While MSI has been proven to be a valuable tool in clinical research, questions remain about the technical reproducibility and applicability of this technology in *routine* clinical research. Furthermore, the use of proteolytic enzymes has raised questions in the field of in-solution bottom-up proteomics, regarding the influence of the enzyme incubation time on the repeatability of the analysis. However, enzyme incubation time and repeatability had not been addressed for on-tissue digestion MALDI-MSI. **Chapter 2** of this thesis addresses the question: What is the influence of the enzyme incubation time on the repeatability and spatiotemporal dynamics of on-tissue digestion?

The use of on-tissue proteolytic digestion can improve proteome coverage in a MALDI-MSI analysis, compared to the analysis of intact proteins. Nevertheless, the information obtained from a single tissue section is still limited. **Chapters 3 and 5** of this thesis address the question whether the use of multiple enzymes can increase the molecular information obtained using on-tissue digestion MALDI-MSI on both fresh frozen and FFPE tissues.

The field of MSI is currently moving toward high-resolution analyses, both in mass resolution and image resolution. The application of high resolution MSI in clinical research is compromised by the long data acquisition times and the large data load which characterize the high-resolution measurements. The virtual microdissection normally performed to focus the data analysis on specific/comparable cell types means that, in many instances, only a fraction of the MSI data is used in the subsequent statistical evaluation. Thus, throughput may be increased, measurement time decreased, and dataload reduced, if the MSI data acquisition could be focused on these histological regions of interest. **Chapter 4** of this thesis describes an automated histology-guided MSI platform, to facilitate high-resolution MSI analysis in clinical research. The functionality and performance of the histology-guided MSI platform are described using a histologically well-characterized selection of FFPE MLS tissues.

Chapter 6 describes a clinical biomarker discovery study to establish a molecular marker for MLS tumor progression based on the application of the multimodal MSI approach described in **Chapter 5** on a previously constructed TMA containing FFPE tissue material from 32 MLS patients.