University of
READING

Investigation of liquid atmospheric pressure matrix-assisted laser desorption/ionisation mass spectrometry (AP-MALDI MS) for the large-scale analysis of liquid samples.

PhD in Chemistry

Department of Chemistry

Oliver John Hale

September 2018

Abstract

Matrix-assisted laser desorption/ionisation (MALDI) is a soft ionisation technique suitable for ionising biological molecules for analysis by mass spectrometry (MS). MALDI MS is typically performed on vacuum-compatible solid samples and on instruments with intermediate to high vacuum (10^{-2}-10^{-7} mBar) ion sources.

Less commonly, MALDI can be performed at atmospheric pressure (AP). However, AP-MALDI presents useful advantages. For example, compatibility is expanded to include samples with components usually volatile under vacuum conditions. Liquid samples may be analysed with ease. Recent developments in the fieldusing liquid support matrices (LSMs), which incorporate a viscous liquid into a solution of a MALDI matrix compound, have demonstrated the ability to reliably generate stable ion signal over long periods of time.

Interestingly, the combination of AP-MALDI, LSMs and a heated ion inlet on the mass spectrometer enables the predominant generation of multiply charged ions. MALDI MS spectra are usually dominated by signals for singly charged ions, with high charge states generally the domain of electrospray ionisation (ESI). The *m/z* of highly charged proteins enables AP-MALDI analysis on narrow *m/z* range mass spectrometers. Additionally, tandem mass spectrometry (MS/MS) experiments also benefit: collision induced dissociation (CID) results in more fragment ion signals from multiply charged precursor ions and electron transfer dissociation (ETD) is possible, since it requires a minimum 2+ charge state to be performed at all.

Liquid AP-MALDI has potential for mass profiling and sample screening, due to the reliability of the signal. In this thesis, the ability to rapidly analyse liquid samples using a prototype AP-MALDI ion source was demonstrated. Applications included the identification of doubly charged phospholipid ions by MS/MS; profiling of lipids, peptides and proteins for sample classification; and the assessment of milk samples for β-lactamase activity. Barium adduct ions of phosphatidylcholines were successfully analysed by liquid AP-MALDI CID and ETD MS/MS to reveal detailed structural information. Profile analysis of milk extracts was used to classify milk samples. A major section focuses on the rapid and confident detection of bovine mastitis, a disease with a considerable economic impact.

<u>Declaration</u>

I confirm that the content of this thesis is my own work and the use of all material from other sources has been properly and fully acknowledged.

Signed: _____

                 Oliver John Hale

Date: _____

Related publications

Data and figures in this thesis have been published in scientific journals, or were included in poster presentations.

Title: Collision-induced dissociation of doubly-charged barium-cationized lipids generated from liquid samples by atmospheric pressure matrix-assisted laser desorption/ionization provides structurally diagnostic product ions.

Abstract: Obtaining structural information for lipids such as phosphatidylcholines, in particular the location of double bonds in their fatty acid constituents, is an ongoing challenge for mass spectrometry (MS) analysis. Here, we present a novel method utilizing the doping of liquid matrix-assisted laser desorption/ionization (MALDI) samples with divalent metal chloride salts, producing ions with the formula $[L+M]^{2+}$ (L = lipid, M = divalent metal cation). Multiply charged lipid ions were not detected with the investigated trivalent metal cations. Collision-induced dissociation (CID) product ions from doubly charged metal-cationized lipids include the singly charged intact fatty acids $[snx+M-H]^+$, where 'x' represents the position of the fatty acid on the glycerol backbone. The preference of the divalent metal cation to locate on the sn2 fatty acid during CID was found, enabling stereochemical assignment. Pseudo-MS^3 experiments such as in-source decay (ISD)-CID and ion mobility-enabled time-aligned parallel (TAP) MS of $[snx+M-H]^+$ provided diagnostic product ion spectra for determining the location of double bonds on the acyl chain and were applied to identify and characterize lipids extracted from soya milk. This novel method is applicable to lipid profiling in the positive ion mode, where structural information of lipids is often difficult to obtain. Graphical abstract MALDI of liquid lipid samples doped with divalent metal salt (e.g. $BaCl_2$) produces doubly charged lipid-barium ions and enables structural elucidation via MS/MS and MS^3 analysis.

Title: AP-MALDI-Q-IMS-TOF MS as a liquid sample mass profiling platform for food authenticity tests

Citation: Hale, O. J., Morris, M. and Cramer, R. In: 65th ASMS Conference on Mass Spectrometry, June 4 – 8, 2017, Indianapolis, IN, USA.

Abstract: Advances in rapid, direct analysis by mass spectrometry, such as Rapid Evaporative Ionisation Mass Spectrometry (REIMS) and Direct Analysis in Real Time (DART) allow straightforward data acquisition but are typically insensitive. Matrix-Assisted Laser Desorption/Ionisation Mass Spectrometry (MALDI MS) is an established, sensitive technique that has traditionally required ions to be produced under vacuum. Combined with liquid sample preparation at atmospheric pressure (AP) rapid data collection can be achieved.

Presented is a proof-of-concept application, with a previously developed liquid support matrix and ion source, for establishing the authenticity of milk samples. Milk adulteration can lead to an economic advantage for dishonest producers and health problems for those with conditions such as Cow's Milk Allergy (CMA). CMA is estimated to affect 2-3% of infants in their first three years of life and milk substitutes are able to accommodate for this.3 Adulterated milk supplies present a risk to the production of these substitutes, and so to those affected by CMA.

Liquid AP-MALDI MS was used to classify six distinct milk types and determine the adulteration level of mixtures of bovine and caprine milk by multivariate analysis of lipid mass profiles. There is scope to extend the methodology into animal and plant health applications.

Title: Production and analysis of multiply charged negative ions by liquid AP-MALDI MS

Abstract:

RATIONALE:

Liquid AP-MALDI has been shown to enable the production of ESI-like multiply charged analyte ions with little sample consumption and long-lasting, robust ion yield for sensitive analysis by mass spectrometry. Previous reports have focused on positive ion production. Here, we report an initial optimisation of liquid AP-MALDI for ESI-like negative ion production and its application to the analysis of peptides/proteins, DNA and lipids.

METHODS:

The instrumentation employed for this study is identical to that of earlier liquid AP-MALDI MS studies for positive analyte ion production with a simple non-commercial AP ion source that is attached to a Waters Synapt G2-Si mass spectrometer and incorporates a heated ion transfer tube. The preparation of liquid MALDI matrices is similar to positive ion mode analysis but has been adjusted for negative ion mode by changing the chromophore to 3-aminoquinoline and 9-aminoacridine for further improvements.

RESULTS:

For DNA, liquid AP-MALDI MS analysis benefited from switching to 9-aminoacridine-based MALDI samples and the negative ion mode, increasing the number of charges by up to a factor of 2 and the analyte ion signal intensities by more than ten-fold compared to the positive ion mode. The limit of detection was recorded at around 10 fmol for ATGCAT. For lipids, negative ion mode analysis provided a fully orthogonal set of detected lipids.

CONCLUSIONS:

Negative ion mode is a sensitive alternative to positive ion mode in liquid AP-MALDI MS analysis. In particular, the analysis of lipids and DNA benefited from the complementarity of the detected lipid species and the vastly greater DNA ion signal intensities in negative ion mode.

Title: Lipid Analysis by Liquid AP-MALDI-Q-IMS-TOF MS

Citation: Hale, O. J., and Cramer, R. In: 39th BMSS Annual Meeting 2018, September 11 – 13, 2018, Cambridge, UK.

Abstract: MALDI MS is well known for its profiling power in an increasing number of applications such as speciation/biotyping in food analysis or clinical microbiology. However, conventional MALDI predominantly produces singly charged ions, resulting in the need to record spectra over a large *m/z* range if peptides/proteins are analysed and with the disadvantage of matrix (cluster) ions in the lower *m/z* range where lipids are typically detected, both leading to some undesired disadvantages for profiling and subsequent identification of either class of biomolecules.

In this presentation the focus will be on the analysis of lipids using our newly developed liquid AP-MALDI methodologies, demonstrating the power of liquid MALDI profiling as well as its generation of multiply charged lipids for structural elucidation, including identification of double bond locations. We will present the latest in applying classification software for liquid AP-MALDI MS profiles, leading to classification accuracies beyond 95% for distinguishing diseased from healthy farm animals, while new methods in producing doubly charged lipids from real-life milk samples allow their accurate identification, using accurate mass, ion mobility and MS^2/MS^3.

The data presented were acquired on a commercial Q-TOF ion mobility mass spectrometer with a modified ion source, using a heated ion transfer tube.

Abbreviations

3-AQ: 3-aminoquinoline

9-AA: 9-aminoacridine

ACN: acetonitrile

Amm. Phos.: ammonium phosphate

AP: atmospheric pressure

ASAP: atmospheric pressure solids analysis probe

CCS: collision cross section

CHCA: α-cyano-4-hydroxycinnamic acid

DART: direct analysis in real time

DHB: 2,5-dihydroxybenzoic acid

DMA: differential mobility analyser

DT: drift tubes

EG: ethylene glycol

ESI: electrospray ionisation

FA: formic acid

FAIMS: field asymmetric-waveform ion mobility spectrometry

GL: glycerol

GUI: graphical user interface

HIP: Hexane/Isopropanol (3:2) solution

ILM: ionic liquid matrix

IMS: Ion mobility spectrometry

IR: infrared

KMD: Kendrick mass defect

LDA: Linear discriminant analysis

LDI: laser desorption/ionisation

LSI: laserspray ionisation

LSM: liquid support matrix

MAII: Matrix-assisted inlet ionisation

MALDI: matrix-assisted laser desporption/ionisation

MeOH: methanol

MG, DG, TG: mono-, di- and triacyclglycerols

MRM: multiple reaction monitoring

MS: mass spectrometry

OPLS-DA: orthogonal partial least squares discriminant analysis.

PA: Phosphatidic acids

PC: phosphatidylcholines

PCA: Principal component analysis

PE: phosphatidylethanolamines

PG: phosphatidylglycerols

PI: phosphatidylinositol

PLS-DA: Partial least squares discriminant analysis

PRM: parallel reaction monitoring

PS: phosphatidylserines

Q: quadrupole

REIMS: rapid evaporative ionisation mass spectrometry

RF: Radio frequency

SAII: solvent-assisted inlet ionisation

SCC: somatic cell count

SM: sphingomyelin

SPE: solid phase extraction

TCA: Trichloroacetic acid

TEA: Triethylamine

TFA: Trifluoracetic acid

TIMS: trapped ion mobility spectrometry

TOF: time-of-flight

TWIMS: travelling wave ion mobility spectrometry

UV: ultraviolet

Table of Contents

1 Introduction

1.1 Mass Spectrometry – a brief historical overview

Mass spectrometry (MS) is a chemical analysis technique with roots that span back over a century into the investigation of "canal rays" by Wilhelm Wien.[1] In 1908 he discovered that these rays (originally discovered in 1886 by Eugen Goldstein), which we now know to be positive ions, were deflected by magnetic and electric fields.[2] The English scientists Sir Joseph John Thomson (1906 Nobel Prize in Physics winner and the discoverer of the electron) and Francis William Aston (the 1922 Nobel Prize in Chemistry winner) built upon this understanding in 1913 to detect beams of cations, which had passed through electric and magnetic fields, on a photographic plate.[3] Two distinct traces were visible and attributable to Neon isotopes, ^{20}Ne and ^{22}Ne. This was the first discovery of isotopes of a stable element and is the first example of mass spectrometry. In addition to Arthur Jeffrey Dempster's introduction of a mass spectrometer in 1918, Aston was responsible for the further development of his "mass spectrograph" into a true scientific instrument in 1919, with revisions to the design continuing for years after.[4, 5]

In 1940, in what retrospectively is a significant moment in world history, a mass spectrometer was used to demonstrate the separation and collection of uranium isotopes.[6] This led to the development of a 'calutron' for large-scale uranium enrichment during the Manhattan Project, alongside the gaseous diffusion method.[7] This ultimately resulted in the development of ^{235}U-based nuclear fission bombs. The very first nuclear weapon used in warfare was ^{235}U-based and was detonated above Hiroshima, Japan, in August 1945 by the USA.[8]

An important advancement of the mass spectrometer was the proposal of the time-of-flight (TOF) mass analyser, a design which forwent the magnetic fields of its contemporaries. The concept was introduced by Stephens in 1946.[9] A working instrument was demonstrated by Cameron and Eggers in 1948 ("Velocitron") and later by Wolff and Stephens in 1953.[10, 11] The TOF has become a staple mass analyser in contemporary mass spectrometry, incorporated into most of the major manufacturers' product lines. It is found in hybrid quadrupole-time of flight (QTOF) systems, first developed by Micromass (now Waters) in 1996, which are pervasive in biological MS.[12-16]

Advances in the latter half of the 20th century saw recognition, in the form of the 1989 Nobel Prize in Physics, for Hans G. Dehmelt and Wolfgang Paul's development of ion trapping techniques.[17] Most ion traps used for mass spectrometry are Paul traps, making use of radio frequency electric fields from three-dimensional quadrupoles, in linear, cylindrical or planar geometries. Linear ion

traps are commonly used as detectors for gas chromatography and incorporated into triple-quadrupole mass spectrometers. Penning traps have found use in Fourier transform-ion cyclotron resonance (FT-ICR) MS, which enabled high m/z resolving power using technology derived from nuclear magnetic resonance (NMR) spectroscopy.[18] The basis for a specific ion trap mass analyser that has become ubiquitous in modern biological MS, was described as early as 1923, and modified again in 1981.[19, 20] Eventually, this concept of ion trapping was also applied by Alexander Makarov in form of the 'Orbitrap' mass analyser and was demonstrated with a laser desorption ion source before the conclusion of the 20th Century.[21] Since 2005, mass spectrometers incorporating this type of ion trap have been exclusively manufactured and sold by Thermo Fisher Scientific.

A plethora of ionisation techniques suitable for coupling to mass spectrometers were developed in parallel with mass analyser advancements. Most of these were only able to ionize atoms or small molecules, for example electron ionisation (EI) and chemical ionisation (CI), now typically found on gas chromatography-mass spectrometry (GC-MS) systems. Perhaps the most important ion sources for biological applications, electrospray ionisation (ESI) and matrix-assisted laser desorption/ionisation (MALDI), came into widespread use in the 1990s after development through the 1980s. Both are forms of 'soft' ionisation, resulting in little fragmentation of fragile molecules, and their application to biological macromolecules was partly awarded the 2002 Nobel Prize in Chemistry.[22-25]

1.2   Quadrupole-time of flight (QTOF) mass spectrometry

All mass spectrometers used throughout this project were of the QTOF design, thus particular focus will be paid to explaining their functionality. QTOF mass spectrometers are hybrid instruments, at the minimum combining a linear scanning quadrupole mass analyser (Q) followed by an orthogonal-acceleration TOF mass analyser (oaTOF). Figure 1 shows a schematic for an early Waters QTOF instrument. Additional abbreviations may be encountered, for example QqTOF ('q' indicating the second, transmission-only quadrupole for fragmentation experiments) and Q-IM-

TOF (the IM indicating an ion mobility analyser situated between the resolving quadrupole and TOF analyser). For brevity, 'q' and 'oa' are often omitted now that the technology has matured.
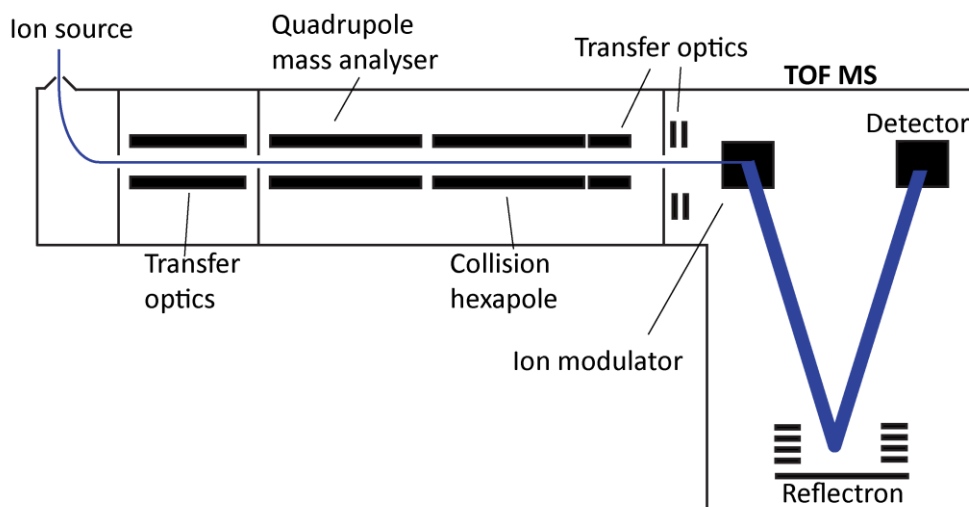
Figure 1: Simplified schematic of a QTOF MS; the blue line represents the ion beam from source to detector.

The quadrupole mass analyser

A QTOF's quadrupole is rarely used as a mass analyser beyond instrument installation qualification and tuning, as its performance, in terms of mass accuracy, resolving power and *m/z* range, compared to the TOF analyser is poor. Modern systems do not include a detector post-quadrupole. The scanning required for detection of all *m/z* within the desired acquisition range introduces inefficiencies in ion transmission, giving a low duty cycle. The reduced Mathieu equations describe the AC field component, often referred to as the radiofrequency (r.f.) component, and DC field component when operating as a mass analyser[26]:

$$q_M = \frac{4eV}{(m/z)\omega^2 r_0^2}$$
Equation 1: Mathieu equation (r.f. component)

$$a_M = \frac{8eU}{(m/z)\omega^2 r_0^2}$$
Equation 2: Mathieu equation (DC component)

Where *e* is elementary charge, *V* is the r.f. voltage amplitude, ω is the r.f. angular frequency, *U* is the DC voltage and r_0 is the inscribed radius of the quadrupole. In analyser mode, the *m/z* transmission range is narrow. The duty cycle is significantly reduced since ions outside of the scan window are lost. However, the ability to 'park' a Q mass analyser on a specified *m/z* window does make it a useful device for MS/MS experiments (discussed in 1.2.1). Operating in transmission mode (without a DC component, i.e. 'r.f.-only'), the quadrupole can act as a high mass filter, with stepping of the r.f. voltage required to transmit wider mass ranges. Duty cycle is reduced to a

3

lesser degree than in scanning mode but stepping still results in losses. The depth of the effective potential well is inversely proportional to m/z, so transmission is reduced at higher m/z in both modes. Many hybrid quadrupole MS instruments use an 'MS profile' to enhance ion transmission over a greater m/z range.

Q mass analysers also have a much poorer mass resolving power relative to a TOF. Higher resolution for a Q mass analyser comes at the expense of scanning speed, meaning further duty cycle reduction.

Other ion guides, operating under the same principles as a quadrupole, include other multipoles (hexa-, octa-, etc) and ring ion guides. Each features a characteristic pseudopotential well profile. Quadrupoles generate a V-shaped pseudopotential well, which enables narrow m/z window selection and ion focusing at the expense of transmission efficiency. Multipoles of higher number produce a more U-shaped pseudopotential well suited for higher transmission efficiency. Ring ion guides have the most U-shaped pseudopotential well of all the above, with a flat bottom and steep sides.[27] Ion transmission is achieved by alternating the electrostatic potential polarity of two alternately-connected sets of stacked rings. Stacked ring optics are found in Waters QTOFs as collision cell and ion mobility optics, where a mass-resolving capability is not necessary, but trapping and transmission performance is required.
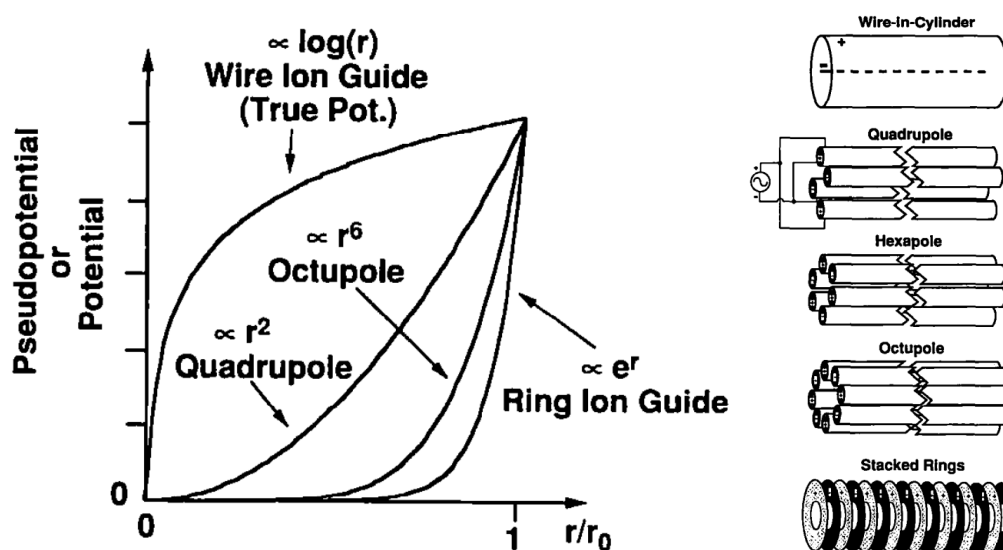
Figure 2: Description of potential wells for different ion guides (where r_0 is the distance to the outermost electrode surface from the ion guide central axis), and diagrammatic representation of each ion guide. Adapted from Figures 1 and 3 as presented by Guan and Marshall.[27]

The time-of-flight mass analyser

Fundamentally, the TOF analyser relies on the conservation of energy; specifically, the conversion of potential energy (e_p) to kinetic energy (e_k).

$$e_p = Q.U \qquad \text{Equation 3: Electric potential energy}$$

And

$$e_k = \frac{1}{2}mv^2 \qquad \text{Equation 4: Kinetic energy}$$

Where Q is the elementary charge, U is voltage, m is mass and v is velocity. These are the parameters on which a TOF MS operates. A diagram of a combined linear/reflectron MALDI-TOF MS is shown in Figure 3. TOF MS typically uses constant energy acceleration (CEA); an accelerating potential, U, is applied to ions before the field free region, thus providing the same e_k to all. The alternative, constant momentum acceleration has not found widespread use.[28] The flight time of ions through the field free region is recorded upon contact with a detector, with ions of greater m/z ratio taking longer than those with a smaller m/z ratio.
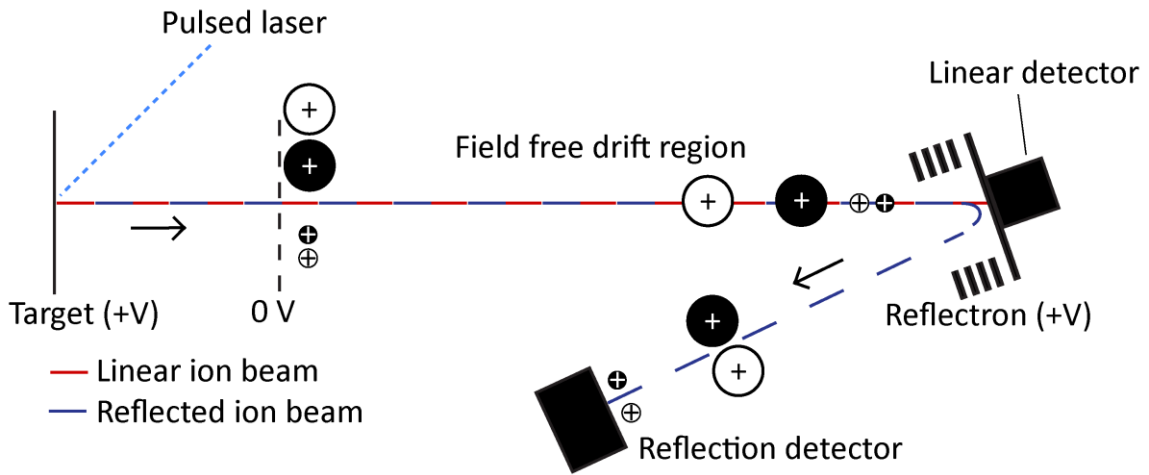
Figure 3: Diagram of a linear/reflectron TOF MS. Low mass (small circle) ions arrive at the detectors faster than high mass ions (big circles) of the same charge. The reflectron reduces the flight time spread between high (black) and low (white) energy ions of the same m/z ratio, providing higher mass resolution than linear TOF mode.

Since velocity is simply = *distance/time*, the *m/z* ratio is calculable from the following relationship:

$$t = k\sqrt{\frac{m}{z}} \qquad \text{Equation 5: Ion flight time}$$

Where *t* is the flight time and *k* is a constant derived from the flight distance *d* and the electric potential difference *U,* applied to accelerate ions:

$$k = \frac{d}{\sqrt{2U}} \qquad \text{Equation 6: Instrument parameter-derived constant}$$

The TOF of a QTOF is positioned orthogonally to the ion beam exiting the quadrupole and transfer optics (see Figure 1). As a result, the beam must be redirected inside the TOF to enter the field-free drift region, whereas the ion beam from a linear/reflectron TOF is produced axially to the

analyser. An assembly of electrostatic grids, culminating as the 'pusher/puller' or 'ion modulator', sample the incident ion beam and trigger the beginning of flight time measurement. The pusher pulse frequency is varied to accommodate the acquisition *m/z* range, since ions of greater *m/z* drift more slowly than low *m/z* ions in the field-free region between modulator and detector, as described by Equation 5.

For instance, the flight time of two ions, [A+H]$^+$ and [B+H]$^+$, of mass 1000 Da and 10,000 Da respectively and unit charge of 1$^+$, for an example flight distance of 1.5 m and accelerating potential of 10 kV:

$$t([A + H]^+) = \frac{1.5\,m}{\sqrt{2 \times 10000\,V}} \sqrt{\frac{(1000\,Da \times 1.66 \times 10^{-27}\,kg)}{1.602 \times 10^{-19}\,C}} = 3.42 \times 10^{-5} s$$

$$t([B + H]^+) = \frac{1.5\,m}{\sqrt{2 \times 10000\,V}} \sqrt{\frac{(10000\,Da \times 1.66 \times 10^{-27}\,kg)}{1.602 \times 10^{-19}\,C}} = 1.08 \times 10^{-4} s$$

Thus, a wider *m/z* range reduces the number of times the pusher can trigger per second without causing ions to overlap, thus reducing TOF duty cycle. For example, the Synapt G2-Si QTOF mass spectrometer employed for most analyses of this project, with an acceleration potential of 10 kV, was used with a pusher interval of 69 μs for *m/z* 100-2000, and 110 μs for *m/z* 100-5000. Additionally, the benefit of multiply charged ions for increasing the duty cycle when analysing large molecules can be demonstrated with the example of [B+10H]$^{10+}$:

$$t([B + 10H]^{10+}) = \frac{1.5\,m}{\sqrt{2 \times 10000\,V}} \sqrt{\frac{(10009\,Da \times 1.66 \times 10^{-27} kg)}{(10 \times 1.602 \times 10^{-19}\,C)}} = 3.42 \times 10^{-5} s$$

Multiply charged ions also possess greater kinetic energy because of the relationships described in Equations 3 and 4. The flight time is shorter for higher charged ions of an analyte of mass *m,* because the velocity is higher, resulting in a greater amount of kinetic energy imparted upon hitting the detector. TOF analysers in a QTOF have multichannel plate detectors (MCPs) connected to a time-to-digital converter (TDC) or analogue-to-digital converter (ADC). A fast ADC is beneficial for applications requiring high dynamic range. Each channel on the MCP acts as electron multiplier, in the presence of a strong electric field. An impacting ion produces a cascade of electrons, which are amplified and recorded. An ion of charge 2$^+$ ejects approximately double the number of electrons from the MCP as a 1$^+$ ion of the same *m/z*.[29] Ultimately this manifests as greater sensitivity to multiply charged ions, in addition to the duty cycle improvement from the narrower acquisition *m/z* range. For these reasons, a doubly charged ion signal tends to

overrepresent the number of ions detected when compared to the singly charged ion of the same analyte. This is also true for MALDI-axialTOF mass spectrometers, but without the pusher duty cycle factor, since the *m/z* measurement is initiated by the firing of the laser, rather than a pusher pulse.

The mass resolving power of a CEA TOF analyser is described by:

$$\frac{m}{\Delta m} = \frac{t}{2\Delta t}$$ Equation 7: CEA TOF mass resolving power

Where *m* is mass and, *Δm* is the peak width at half its maximum intensity, *t* is time-of-flight and *Δt* is the time difference in arrival times for the two ion packets. Slight energy differences in otherwise identical ions entering the field free region are a limiting factor for TOF resolving power since this produces the peak width.[30] The $\frac{t}{2\Delta t}$ dependency leads to a decrease in resolving power with increasing *m/z* since the kinetic energy distribution for heavier ions is higher than for lighter ions.[28] Orthogonal acceleration in a QTOF instrument alleviates some of the kinetic energy differences, so the resolving power for heavier ions tends to be better than for axial TOF. Modern commercial reflectron oaTOFs typically have a resolving power of 10,000 – 60,000, at approximately *m/z* 1000, depending on their mode of operation (which sacrifice sensitivity for improved peak resolution). Mass resolving power is typically quoted for an analytical standard at a defined *m/z*.

The variance in flight times for ions of identical *m/z* is reduced by incorporating an ion mirror, which effectively doubles the ion beam path length, reduces energy spread and results in more highly resolved peaks in the mass spectrum.

$$t = \frac{d}{\sqrt{2U}}\sqrt{\frac{m}{z}} + \frac{2L_m\sqrt{2U}}{U_m}\sqrt{\frac{m}{z}}$$ Equation 8: Time of flight with a single-stage reflectron

Where *t* is ion flight time, L_m is the length of the ion mirror, *U* is pusher potential difference, U_m is the ion mirror potential difference and other parameters are as described in Equations 5 and 6. Dual-stage reflectrons may be used to improve resolving power performance. These feature two electric fields of differing field strengths; the first, strong to rapidly decelerate ions and the second weaker.[31] Dual-stage reflectrons can correct kinetic energy distributions to the second order, limited to just the first for the single-stage variety, hence the improved resolving power and are typical of the reflectrons found in oaTOFs today.[32]

An oaTOF usually incorporates a reflectron, producing a V-shaped ion path, for improved resolution compared to a linear TOF. Some commercial instruments incorporate multiple reflectrons or reflectron passes to boost resolution further, for instance, with W and N-shaped paths (examples in Figure 4). Additional novel TOF geometries are also reported in the literature.[33-35]
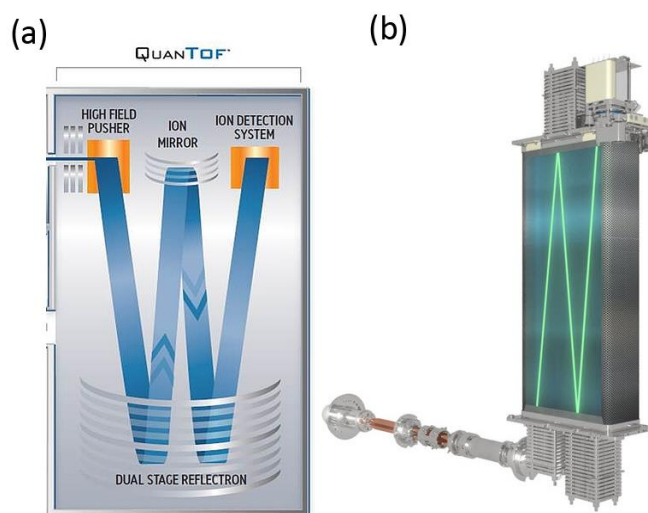
Figure 4: Schematics of (a) W-optics on a Waters Synapt G2-Si and (b) N-optics of a Sciex X500R QTOF.

1.2.1 Tandem mass spectrometry

QTOF mass spectrometers are capable of tandem mass spectrometry (MS/MS) experiments. MS/MS involves an ion fragmentation step experiment performed between two MS stages. With knowledge of an analyte's precursor and MS/MS product ions, detailed structural information can be determined, and confident analyte identity assigned. MS/MS is also described as MS^n, where n >1. For example, fragmentation of MS^2 products is an MS^3 experiment. Multiple stages (>3) of MS^n are achievable on high performance ion traps such as FT-ICR, Paul traps and Orbitrap instruments featuring a linear ion trap.[36, 37]

With respect to QTOF MS, the most routine MS/MS experiment is collision induced dissociation (CID) MS/MS, which involves the fragmentation of quadrupole-selected precursor ions through collisions with an inert gas (argon, nitrogen, etc.) at elevated kinetic energy.[38] Energy transferred to the precursor ions in these collisions is distributed throughout the ion and causes bonds to break and rearrange. The product ions of these collisions are transmitted to the TOF for analysis. Since the precursor ions are selected by a narrow *m/z* window, typically only transmitting the precursor ions' isotopologues, fragments are indicative of the precursor ions' structure. Reconstruction of these fragments can enable identification of the precursor ions. This is routinely used for small molecules in selected or multiple reaction monitoring (SRM, MRM) experiments;

the 'transition' from precursor ion to one or more product ions is monitored to confirm the identity of the analyte.

The study of the proteins of a biological system, "proteomics", has been a significant application of MS for the last two decades. In proteomics, CID MS/MS is commonly used for identification of peptides/proteins in a 'bottom-up' workflow; first, proteins are digested by a highly specific enzyme (usually trypsin); then the resulting peptides are analysed by LC-MS/MS. Since peptide ions that undergo CID generally fragment to produce a predictable series of 'b' (N-terminal) and 'y' (C-terminal) ions through cleavage of peptide bonds (see Figure 5), confident identification of thousands of ions per experiment is possible with a fast mass spectrometer and the appropriate software. The nomenclature for peptide fragment ions was developed in the mid-1980s.[39, 40] 'Top-down' analysis with intact proteins is also possible, although not all MS/MS techniques are effective for protein ion fragmentation.[41] Top-down proteomics analyses are often performed by MALDI-TOF MS after offline sample fractionation.
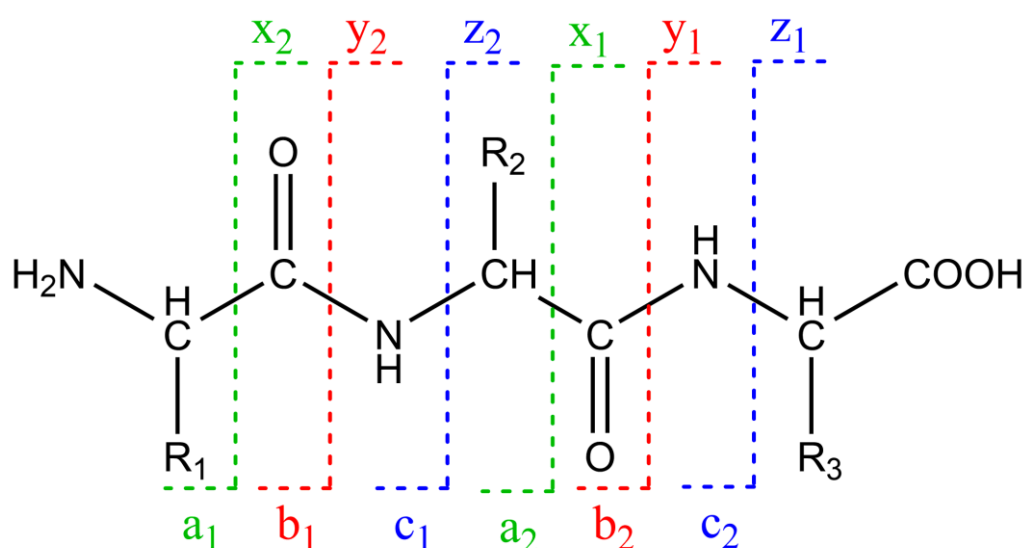
Figure 5: Schematic of peptide fragmentation and nomenclature. a, b and y ions are usually observed from low energy CID and ISD/PSD; c, z+1 and z+2 ions are typical of electron-mediated techniques.

The Waters Synapt instruments are capable of CID pre- and post- ion mobility since they feature 'trap' and 'transfer' collision cells. Thus, drift time measurements on either precursor or product ions is possible. An additional MS^n fragmentation scheme called time-aligned parallel (TAP) maintains the drift time of fragment ions so that they match the precursors' (example in Figure 6). This method can be used as a pseudo-MS^3 technique for identifying fragments of fragments by their alignment on a mobiligram (*m/z* versus drift time). In addition to TAP MS^3, pseudo-MS^3 can be achieved using in-source fragmentation (ISD). On a QTOF instrument with an API inlet, this simply involves increasing the inlet cone voltage. Product ions are like those produced in CID.

When a known analyte is being analysed, this allows a second stage of MS/MS to be performed on its fragments. It is not a suitable MS/MS technique for complex mixtures of similar components, since ISD does not offer precursor ion selectivity.

(a) CID in trap cell

(b) CID in transfer cell
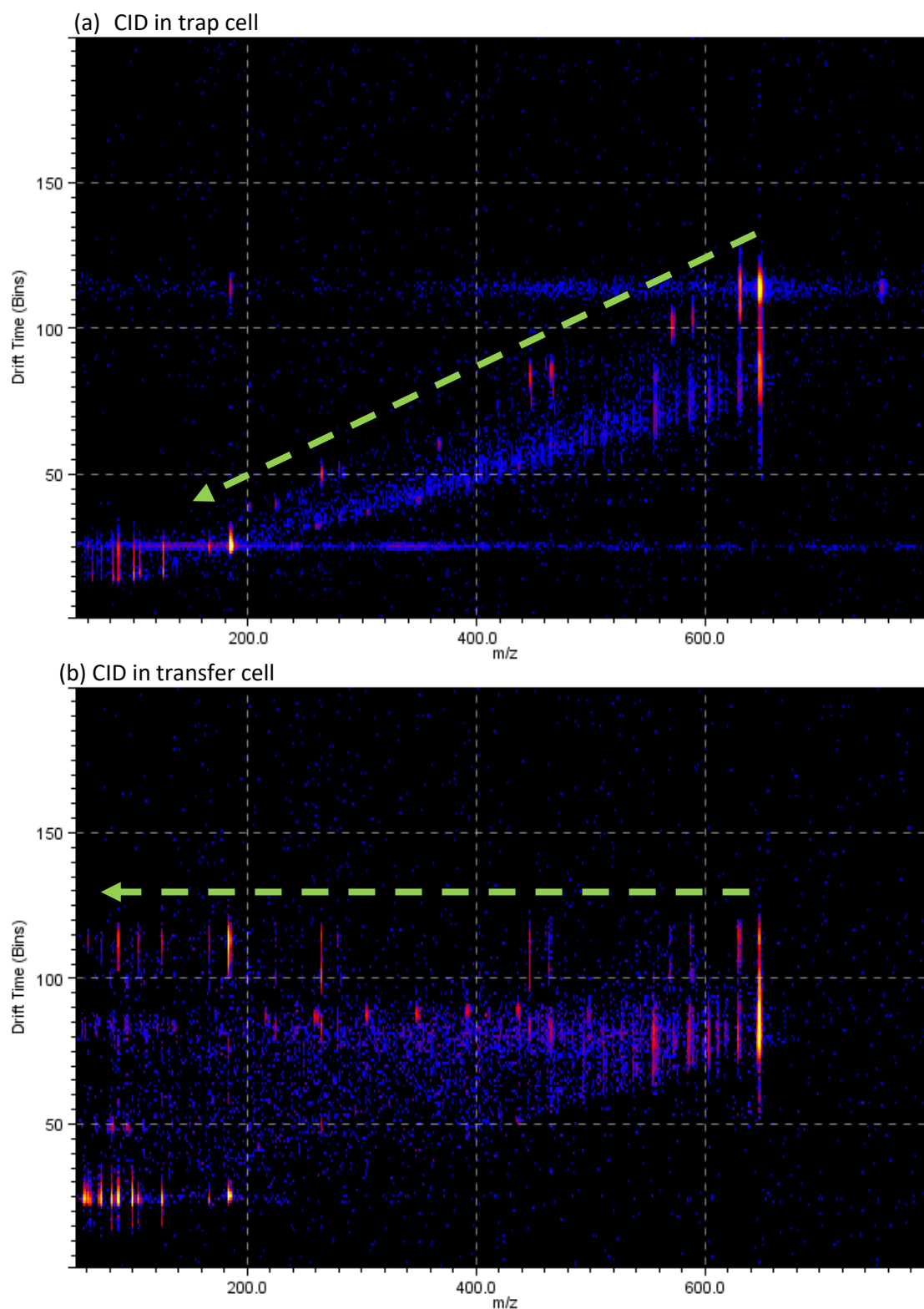
Figure 6: 2-dimensional m/z vs DT plots for CID-MS/MS of a sphingomyelin (m/z 647.5) compound in the trap (a) and transfer (b) cells of a Synapt G2-Si. Fragmentation in the transfer cell keeps product ions 'time-aligned' with the precursor ion as indicated by the green arrow.

Another fragmentation technique which has become advantageous in biological mass spectrometry is electron transfer dissociation (ETD). It is one of many electron-mediated MS/MS methods and has been successfully implemented on QTOF, Orbitrap and quadrupole ion trap

11

(QIT) mass spectrometers. A caveat is that it requires positive ions possessing at least 2 units of charge to function, since the core mechanic is an ion-ion reaction involving the transfer of an electron from a radical reagent anion to the analyte cation, resulting in a reduction of the analyte charge by 1. A negative analogue of ETD 'nETD' also exists, but typically has a lower fragmentation efficiency due in part to competition from proton transfer reactions.[42] Electron transfer with no dissociation (ETnoD) is another related technique, used for investigation of protein structural rearrangement by the addition of a single electron.[43] Similarly, electron capture dissociation (ECD) and electron induced dissociation (EID) use electrons of low (1-5 eV) and high (>10 eV) energy to promote fragmentation and are most often found on FT-ICR MS.[44] ECD is similar to ETD, but does not involve the interaction of a cation and anion, but rather a free electron with the cation, proposed to be a non-ergodic process.[45] EID produces complementary fragments to CID through collisions with electrons. ETD is a valuable technique for fragmenting precursors where CID is ineffective, finding implementation for proteomics in particular.[46, 47] The product ions from peptides and proteins are even electron c-ions and odd electron z$^\bullet$-ions which can be reconstructed to divulge structural information in much the same way as for CID. ETD is favourable for post-translational modification (PTM) analysis, since the fragmentation mechanism leaves modifications intact.[41] ETD's application to smaller molecule analysis is less frequent, although there are reports of ETD of multiply charged lipid complexes.[48, 49] The implementation of ETD on the Synapt mass spectrometers involves ionising the reagent by negative mode glow discharge, simultaneously with the generation of cations from the sample. The ion optics up to and including the TriWave trap (see Synapt G2-Si diagram Figure 16) are rapidly switched in polarity to transfer and trap positive and negative ions together, with precursor selection by the quadrupole if desired. A DC potential in the trap keeps reagent and analyte ions separated from one another until reaction is desired, at which point the potential barrier is lowered. The ETD reaction occurs, and fragments are transferred to the TOF through the remaining TriWave cells.[50]

Fragmentation by photon-induced techniques is a third alternative. Two prominent examples include ultraviolet photodissociation (UVPD) and infrared multiphoton dissociation (IRMPD). A commercial UVPD implementation exists for the Orbitrap platform but reports for QTOF instruments featuring UVPD also populate the literature, but it is in relative infancy.[51-53] UVPD and IRMPD have featured on Fourier transform-ion cyclotron resonance and modified linear quadrupole mass spectrometers.[54-56]. IRMPD is a common, analogous fragmentation method to CID, but can be performed within the FT-ICR cell without affecting the ultra-high vacuum since a gas is not introduced. UVPD has the potential to increase specificity and fragment yield beyond

other fragmentation methods.[57] The wavelength of the laser can be chosen to enhance specificity, for example the production of d-ions from carbon-sulfur bonds with 213 nm photons.[58] UVPD also offers a route for characterising lipid structures in detail, where other methods fail.[59, 60] UVPD was not available on the mass spectrometers in this project, but it is an upcoming disruptive technology and has relevance to lipid analysis discussed later in this thesis.

1.2.2   Ion mobility-mass spectrometry (IM-MS)

Ion mobility (IM) devices have been incorporated into commercial mass spectrometers for around a decade. These include technologies such as travelling wave ion mobility spectrometry (TWIMS, on Waters Q-IM-TOFs), trapped IMS (TIMS, on Bruker timsTOF instruments), drift tubes (DT), field asymmetric waveform IMS (FAIMS, able to be retrofitted to the mass spectrometer API inlet) and differential mobility analysers (DMA). Amongst the commercially available implementations, TIMS has the highest mobility-resolving power, with $\Omega/\Delta\Omega = 100\text{-}250$ achievable at room temperature.[61] Current commercially available TWIMS has $\Omega/\Delta\Omega = 40\text{-}50$, although performance has been demonstrated to be better than the resolution value would suggest.[62, 63] Since only TWIMS was used in this project, details of the other techniques are omitted.

TWIMS has featured in every iteration of Waters' Synapt product line as the middle cell of the TriWave device (see Figure 16 later in the thesis), and as a separate cell prior to the resolving quadrupole on the Vion QTOF. The TWIMS cell is operated at elevated pressure with a drift gas, typically nitrogen. Helium is used to cool ions entering the high-pressure region. The stacked ring assembly is pulsed with an r.f. voltage, the phase on each ring being opposite that of the one before. This produces an r.f. wave that carries ions through the drift gas (Figure 7).
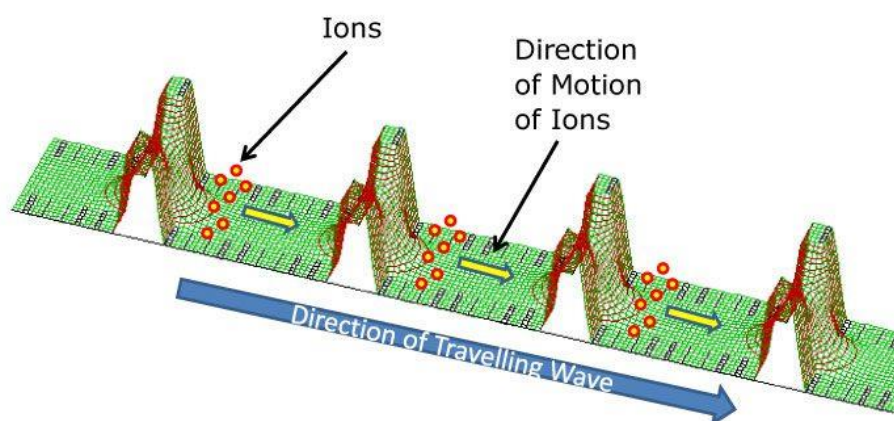
Figure 7: The travelling wave IMS concept, as presented on the Waters website.[64]

With regards to TWIMS use for mass spectrometry, the separation of ions is not based on their *m/z* ratio but on their physical interaction with the drift gas and gives an additional, orthogonal

dimension of information. The travelling time confusingly also called 'drift time' (*cf.* TOF drift time), can allow ions of differing unit charge, but identical *m/z* to be identified. Ions of different molecular species (for example, species of phospholipids) are distinguishable by their drift time.[65] If calibrated, the TWIMS cell can also provide information on the collision cross section (CCS) of the ions. This physical property can reveal information about the ion structure and is essential for characterising ions across multiple TWIMS instruments, since small variations in instrument conditions and manufacturing tolerances result in variations in the raw drift time measurement.

1.3 Matrix-Assisted Laser Desorption/Ionisation (MALDI) and other ionisation techniques

1.3.1 MALDI Development and Applications

Laser Desorption/Ionisation (LDI) MS was first reported during the late 1970s and early 1980s, with multiple publications making use of the Laser Microprobe Mass Analyser (LAMMA) instrument.[66-70] This instrument was developed for biomedical purposes, a field which drives MS developments to this day.[71] Studies of cations in body tissues were reported[67, 68] but also small organic compounds;[70] the analysis of amino acids and dipeptides was reported in 1985.[72] Application of the technique to organic molecules of up to *m/z* 100,000 was demonstrated in 1988, in publications by Tanaka *et al.,* and Karas and Hillenkamp.[23, 73] The approach in the former used a matrix of fine cobalt powder mixed with glycerol and organic solvents; the latter demonstrated MS analysis with a nicotinic acid matrix compound. The matrices enabled controlled absorption of laser energy, protecting delicate molecules and enhancing desorption. While the use of glycerol-based matrices has continued in some applications and Tanaka was awarded part of the 2002 Nobel Prize in Chemistry, organic acid matrices have become the standard for routine MALDI MS analysis.[22] Particularly, the organic acids 2,5-dihydroxybenzoic acid (DHB) and α-cyano-4-hydroxycinnamic acid (CHCA) are staples for peptide mass fingerprinting (PMF), intact protein ionisation, polymer analysis and biotyping. Matrices beneficial to the analysis of specific molecular families have been reported, for example 2,5-dihydroxyacetophenone for proteins, dithranol for lipids and small molecules and 2,4,6-trihydroxyacetophenone for lipids.[74-76] The matrix 4-chloro-α-cyanocinnamic acid (Cl-CCA) was designed from first principles for MALDI analysis of peptides and has been shown to significantly improve analyte signal intensity versus CHCA.[77] However, the improvement was only possible with 337 nm lasers because of the absorbance properties of the compound. For negative ion analysis, matrices with high gas phase basicity may be chosen to provide improvements over acidic matrices.[78] UV-absorbing organic bases such as 3-aminoquinoline (3-AQ), 9-aminoacridine

(9-AA) and 1,8-bis(dimethyl-amino)naphthalene (DMAN) are examples.[79-81] The latter has been demonstrated for sensitive analysis of small metabolites with low background signal. While MALDI analysis of small molecules (<1000 Da) is possible, it is generally hindered by chemical noise contributed by matrix-related ions. Matrix suppression modes are available on some commercial MALDI mass spectrometers.

In addition to the matrix compounds themselves, matrix additives can provide additional analysis potential. Weak acids like formic and trifluoroacetic acid are commonly added to MALDI samples for positive mode analysis to provide additional protons. Metal salts are another additive and have been used for lipid and synthetic polymer analysis. Lithium salts are frequently reported in the literature, resulting in lithium-cationised analyte ions. Divalent metal salts of alkaline earth or transition metals can be used in the same way, and fragmentation of ions cationised like this can reveal additional structural information about the analyte.[76, 82, 83]

MALDI is most commonly initiated with a pulsed UV laser; either nitrogen gas lasers that emit at 337 nm, or solid state, frequency-tripled Nd:YAG lasers with λ=355 nm. Both typically feature a pulse width in the 1-10 ns range, application of lasers with pulse widths in the pico and femtosecond range is an area of active research. Nitrogen lasers typically have lower pulse repetition rates (1-30 Hz) than solid state lasers (commercially available at 10 kHz) and their performance degrades over time. However, many MALDI matrices were developed with the 337 nm wavelength and thus it is still prevalent today. Less commonly used for UV-MALDI are frequency-quadrupled Nd:YAG lasers (λ=266 nm) and frequency-tripled Nd:YLF lasers (λ=349 nm). Visible wavelength MALDI can be conducted with frequency-doubled Nd:YAG lasers (λ=532 nm).[84] Infrared (IR) MALDI is less common than UV-MALDI, but was demonstrated for analysing biological molecules relatively soon after MALDI's introduction.[85] The lasers usually feature a longer pulse width in the region of 100 ns. Nd:YAG lasers (λ=1064 nm) may be used, although Er:YAG lasers (λ=2940 nm) are vastly superior as water and other OH-containing molecules strongly absorb this radiation, thus can be exploited as a matrix.[86-88] Thus, aqueous samples do not require a matrix chromophore to be added, as with MALDI at other wavelengths. Sample consumption in IR-MALDI is high, relative to UV-MALDI. Tuneable wavelength pulsed lasers are also used in some MALDI studies across a wide range of wavelengths, although they are not typical of commercial ion sources.[89]

MALDI TOF MS is suitable for clinical analysis, allowing the identification of microbes present in fluids and tissue.[90-93] To date, two mass spectrometry systems, the Bruker MALDI BioTyper CA System (Bruker, Bremen, Germany) and the VITEK MS (BioMerieux, Marcy-l'Étoile, France) have received FDA approval for analysis and identification of microorganisms. Rapid data acquisition,

the high sensitivity and resolution of TOF instruments and relatively good tolerance to sample contaminants have all contributed to MALDI's establishment as a routine method. Each system uses a vacuum source and a UV laser. Both MS biotyping systems feature a linear TOF MS, marketed as providing sufficient resolution and sensitivity for microbiology applications. In practice this also results in easier to maintain, cheaper, and smaller footprint systems. More capable TOF analysers would increase costs and are largely unnecessary for the application.

1.3.2 Ionisation

MALDI is perhaps most distinct from other soft ionisation methods in that singly charged ions are predominantly produced, even for molecules over 100,000 Da.[94] A laser pulse is delivered to the sample, where it is absorbed and the energy distributed, causing ablation of material (Figure 8).
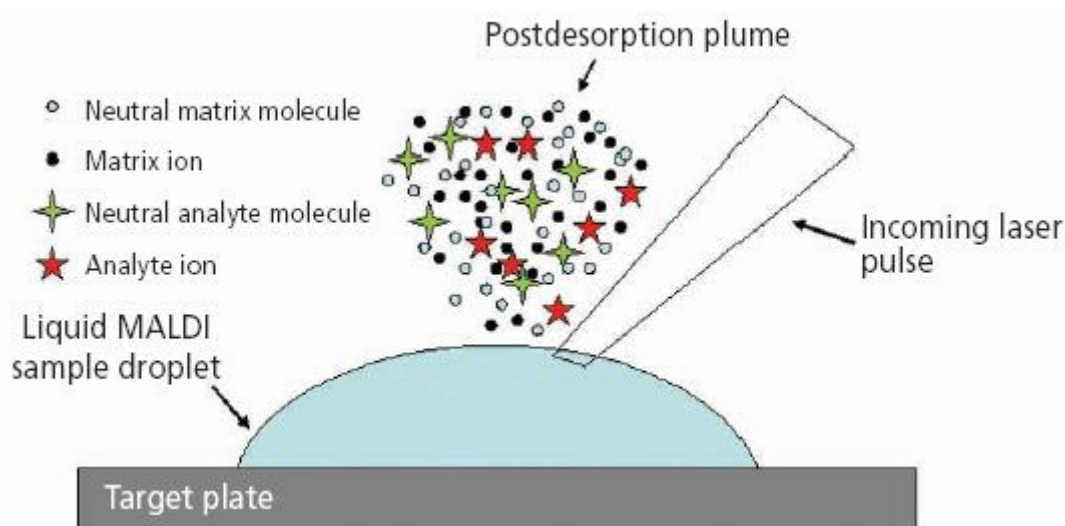
Figure 8: Illustration of the MALDI process from a liquid droplet (adapted from Towers and Cramer, 2010 [95])

It is generally accepted that ionisation occurs by multiple convoluted pathways, as evidenced by the variety of models proposed in the literature, which continue to refine understanding of the processes.[96-101] The 'lucky survivor' and 'cluster' models suggest that detected ions are the result of surviving neutralisation events in the MALDI plume.[102] However, aspects of the theory, such as ions preformed in the sample, are now less favoured explanations.[103][104] An exciton-based photochemical model, thermal and droplet-based models have all been proposed.[105, 106] Nevertheless, there is still no complete model that is accepted by the community.

Explanations of the ionisation processes in ESI are somewhat less convoluted and better understood.[107] They also potentially have a greater relevance to AP-MALDI of liquid samples than traditional MALDI models, since it is expected that droplet desolvation is a critical factor (discussed in the following section).[108, 109] Increasingly, it is becoming evident that soft ionisation methods are closely related. For liquid samples, it has been proposed that there are essentially

two processes; neutralisation of the counter ion and separating ions. An ion source may exploit either one or both by multiple strategies, for example the application of heat, shockwave or voltage. [110] Taking the example of ESI, a high voltage is applied to the sample emitter, which neutralises ions of the opposite polarity. Ions of the same polarity are ejected in solvent droplets which rapidly evaporate to leave analyte ions or eject ions in a charge equilibration event. ESI probes are often fitted with heated desolvation and nebulising gases, and heated MS inlets are present on many commercial and modified mass spectrometers. Heated inlets promote additional desolvation through the ESI processes described in Figure 9.[111] The ion evaporation model (IEM) suggests Coulombic repulsion causes ions to leave an evaporating droplet, while the charge residue model (CRM) predicts that charges remain on an analyte molecule left behind after solvent evaporation. Finally, the chain ejection model (CEM) mostly applies to amphiphilic molecules, with the hydrophobic portion aiding ejection of the ion from the droplet. Recent reports of LDI MS of liquid samples suggest an API ion source's heated inlet is key to achieving ionisation since a conventional matrix compound was not used.[89, 112]. This is largely in agreement with reports on solvent- and matrix-assisted inlet ionisation (SAII and MAII respectively), which demonstrated ESI-like mass spectra with sample introduction directly into heated API inlets.[113, 114] High charge state ions comparable to ESI have recently been reported in vacuum MALDI instruments, through use of high volatility matrices originally used for laserspray ionisation and MAI. In addition, it was shown that lower volatility matrices were capable of similar results should sufficient desolvation energy be provided by the laser.[115]
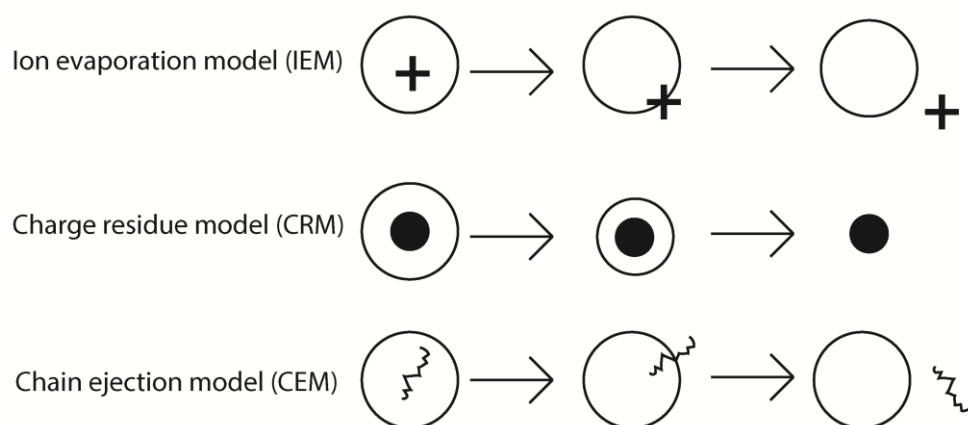
Figure 9: ESI processes resulting in free ions. The CEM is mostly applicable to amphiphilic molecules like proteins.

1.3.3   MALDI with solid and liquid samples

Typical sample preparation for MALDI MS experiments results in a crystalline sample. MALDI samples prepared as such have found wide adoption in applications including top-down proteomics, molecular imaging of body tissues and bacterial speciation.[116-118] A variety of strategies for applying samples to the MALDI target are employed. For example, sample and

matrix may be premixed and then spotted and dried on target, or each may be spotted into the same sample well without premixing. Variations on the ubiquitous 'dried droplet' method are reported throughout the literature.[119, 120]Thin films and matrix 'seed-layers' have been proposed to enhance sample homogeneity.[121, 122]However, there are multiple disadvantages to the solid sample state. Sample heterogeneity is a considerable issue, with so-called 'sweet spot hunting' a necessity to achieve analyte signal.[123-125] The samples require time for drying, a problem which increases with sample spot volume. In addition, drying conditions like humidity and temperature can substantially affect crystal morphology, which in turn affects the analytical performance.[126, 127] There are devices for alleviating these problems, like controlled environment cabinets and matrix spraying robots, but each has additional costs, maintenance and performance issues of their own. As an alternative, the liquid sample state can be employed. Although it is not as widely established as the solid state there are various advantages. There are two categories of liquid MALDI matrices; ionic liquid matrices (ILM) and liquid support matrices (LSM). MALDI ILMs combine a traditional MALDI organic acid matrix, an organic base, such as 3-aminoquinoline or triethylamine, and solvents.[128] LSMs integrate an organic acid or base solution and a viscous liquid such as glycerol.[108, 129-131]

For many liquid MALDI samples, a pseudo-continuous ion beam similar to the spray-type ion sources can be achieved since the liquid samples exhibit self-healing properties, eliminating the requirement to move the laser as with solid samples. This property allows straightforward tuning of the mass spectrometer settings (like in ESI MS) and stable analysis of a sample over time. The stable beam also has utility for MS/MS of low abundance analytes, where minutes of data may be required to achieve useful spectra.

Both ILMs and LSMs have some limited compatibility with (high-)vacuum MALDI sources, if the volatile components of the droplets fully evaporate under these conditions. Atmospheric pressure (AP) or in an external vacuum chamber are alternatives for these cases. However, this means a drying stage is not omitted. Without drying outgassing can occur within the mass spectrometer, resulting in contamination across the target and internal components and stress on the vacuum system. This is a significant problem with a highly populated target plate because even a small amount of outgassing per sample adds up to the point where evacuation of the vacuum source may fail. In addition, the sample longevity within vacuum sources is low. A sample becomes highly viscous after a short period of time and behaves as if it is a solid, i.e. the analysis of a whole target will be impractical.

AP-MALDI is a true alternative to typical vacuum MALDI that can take advantage of liquid samples to their fullest. Samples can be loaded immediately after spotting without the risk of

contamination or damage to the vacuum system. Low sample consumption allows the analysis of a single sample for up to 30 minutes or more at a time.[108, 132] Collisions with atmospheric molecules enhance desolvation compared to vacuum MALDI. However, AP-MALDI sources are uncommon with few commercial ion sources being available. MassTech Inc. have had most significant interest in this type of ion source since the turn of the century.[133-135] There are also home-built sources of varying design used by research groups worldwide.[89, 108, 136, 137]

1.3.4   Ambient Ionisation.

At present, there is considerable interest in atmospheric pressure and ambient ionisation techniques. Techniques like rapid evaporative ionisation (REI)MS, atmospheric pressure solids analysis probe (ASAP), direct analysis in real time (DART) and liquid extraction surface analysis (LESA) provide a range of complementary techniques. All have a focus on limited sample preparation and rapid analysis.

REIMS particularly has captured the attention of mainstream media as the 'iKnife' device; a superficially simple concept introduced by Professor Zoltan Takats that was acquired by Waters for further development into a commercial product. Currently, the source features an electrosurgical tool which ablates material from tissue samples. The vapours are transferred to the mass spectrometer inlet by a Venturi pump where ionisation occurs by impacting on a heated surface. The 'impactor' is key as it acts to rapidly desolvate the surgical vapours, releasing abundant, polar molecules, such as phospholipids, for analysis. Multivariate statistical models are built from the mass spectra for online classification of samples. Attention has been paid to its application in cancer surgery, where effective removal of as much tumour tissue as possible is essential.[138] The source design features an inlet for introduction of a lock mass solution. A patent claims that solvents sprayed with the lock mass compound improved REIMS ion yield.[139] Laser-based and ultrasound-based REIMS are in development, with the potential for greater sample compatibility.[140, 141] REIMS has yet to be demonstrated for analysing larger biological molecules; the desorption and ionisation conditions may simply be too energetic for peptides and proteins to survive (the impactor operates at approximately 900 °C.).

REIMS and other ambient ionisation techniques are also being developed for food applications, which are explored in Section 1.5.1.

1.4   Lipid Characterisation

While characterising molecules by the fragmentation techniques described in Section 1.2.1 is routine for many applications, lipids are a challenging class of molecules to fully characterise. The

field of lipidomics (the large-scale analysis of lipids from cells and body tissues[142]) is a rapidly developing area of mass spectrometry thanks to technological advancements, such as accurate mass MS and IM-MS, and database improvements.[117, 143-145] Lipids are important as structural molecules in body tissues, but also perform signalling roles in cells. Typical tandem mass spectrometry methods do not yield the same depth of information that they are able to reveal for peptides. There are many classes of phospholipids, which possess different headgroups (Figure 10) and other classes such as sphingophospholipids (Figure 11). With CID of lipid ions $[L+H]^+$ or $[L-H]^-$ important structural properties like head group identity and side chain mass can be inferred from neutral losses or detection of headgroup ions. However, information such as double bond location has required more exotic techniques. For example, phospholipids cationised by alkali metals (Alk = Li, Na) in the form of $[L+Alk]^+$ produce different CID spectra compared to $[L+H]^+$.[146] Neutral loss of fatty acid side chains was recorded with $[L+Li]^+$, whereas $[L+H]^+$ predominantly produced the headgroup ion. Lithium has been used in many lipid structural studies since, although it provides limited double bond information without many successive stages of MS^n.[82, 147-151] Phospholipid complexes in the form of $[L_n+M]^{2+}$ and individual phospholipids in the form of $[L+M]^{2+}$ cationised with alkaline earth metals have been produced by ESI and liquid AP-MALDI, respectively.[152, 49, 112] The latter was able to identify the double bond location in phosphatidylcholines. The stereochemistry of the fatty acids was identified in both cases. ETD of $[L+2Na]^{2+}$ produced by ESI has enabled tentative assignment of double bond locations.[48]
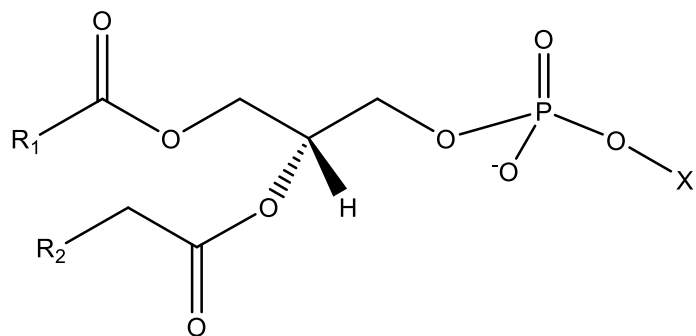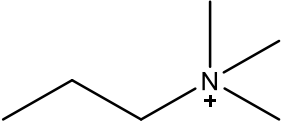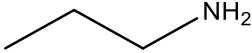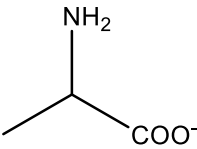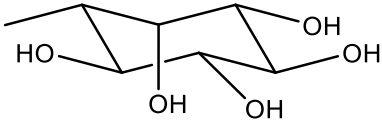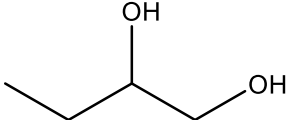
Headgroup Substituent (X)	Type	Abbreviation
H	Phosphatidic acid	PA
	Phosphatidylcholine	PC
	Phosphatidylethanolamine	PE
	Phosphatidylserine	PS
	Phosphatidylinositol	PI
	Phosphatidylglycerol	PG

Figure 10: Glycerophospholipid general structure and headgroup variations.
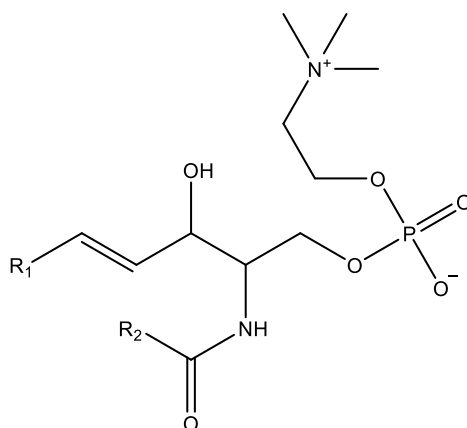
Figure 11: The general structure of sphingomyelins, a type of sphingophospholipid.

Other methods involve moving beyond a simple sample additive. The recently developed ozone-induced dissociation (OzID) is promising for lipidomics since it has been demonstrated to be functional on a chromatographic timescale.[153, 154] MS plays a similar, critical role as in proteomics; in detection and identification of analytes, often by LC-MS/MS. In OzID, ozonolysis at double bonds was performed on ions trapped in a high pressure region in the MS, enabling the identification of their location through understanding the resultant product ion spectra.[155-157] OzID has also been combined with CID to enhance fragmentation.[158-160] ESI in an ozone-rich environment (OzESI) for double bond ozonolysis has produced similar results to OzID.[161] Since essentially the only change required to standard instrumentation is to replace the typical trap gas with ozone, this method could see wide adoption. Safety with handling ozone and the necessity to generate it on-site are matters which must be considered if implementing the technique.

Taking advantage of the Paternò-Büchi (PB) photochemical reaction, carbonyl compounds can react with double bonds by activation in UV light. Subsequent CID of the reaction products provides diagnostic product ion spectra for determining the double bond location.[162 163] PB-MS has been demonstrated online with LC-ESI MS/MS. Little instrument modification would be required beyond the UV light source to take advantage of PB reactions in MS.

Charge transfer dissociation (CTD) currently requires substantial instrument modification since no commercial apparatus is available. However, the structural information produced from the charge-increased ions $[L+H]^{2+\bullet}$ is considerable.[164] Double bond *cis/trans* isomerism has been determined by differential mobility spectrometry-electron impact excitation of ions from organics (DMS-EIEIO).[165] *Cis/trans* isomerism was assigned from fragments resulting from electron irradiation and separation by DMS.

Areas of application for lipidomics, whether utilising mass profiling, structural characterisation or quantitation, include food analysis and disease detection, both of which were investigated during this project.

1.5   Food Adulteration

Modern food production is a global industry. With raw materials sourced from all over the world, it is paramount that quality standards are maintained. The health of the consumer is the primary consideration; mischaracterised and adulterated products may cause medical issues such as allergic reactions, malnutrition or poisoning. Product authenticity is also important from an ethical and financial perspective. Economically motivated adulteration (EMA) can put dishonest producers at an unfair competitive advantage.

The importance for ensuring food authenticity and quality was demonstrated in cases which have occurred over the last two decades. The 'Melamine Scandal' was a high-profile case in 2007-2008. Melamine-adulterated milk products entered the supply chain, affecting 300,000 people with six juvenile fatalities in China. This led to a ban on the import of milk products from China by the USA where the pet food supply chain was compromised.[166, 167] Other countries followed suit, with consequential economic effect. Melamine was added to boost the apparent protein content of products, a property on which quality was assessed, but resulted in severe kidney damage for those who consumed it. EU member states were not affected in this case, however, the EU horse meat scandal occurred only a few years later (2013). This was another case of fraudsters benefiting financially from adulterating products; beef cost €3900 per tonne versus €800 for horse meat, so there was a considerable monetary incentive.[167]

Throughout this thesis, proof-of-concept AP-MALDI MS liquid sample profiling has been applied to milk speciation. Milk (along with olive oil, honey and saffron) is one of the most common targets of EMA according to a 2012 report covering the period 1980-2010.[168] Cow milk accounts for 85% of global milk production.[169] Higher value milk, for example that from goats (3.4%) and sheep (1.4%), is liable to EMA using cheaper cow milk.[170, 171] Animal milk products may also be subject to adulteration with cheap, plant-based proteins.[172] Many plant proteins are derived from sources determined as allergens; soy, wheat and almonds. Conversely, plant-based milk products are often consumed by those with lactose intolerance, so their adulteration with animal milk presents a similar problem. People following modern diets and lifestyles that omit animal products, such as veganism, also suffer in these cases.

1.5.1   Mass spectrometry profiling in food adulteration testing

As the technology has matured and become more affordable, food fraud identification has become a focus for MS. Methods amenable to routine, high-throughput characterisation of samples are possible routes to improve food safety and quality. Classification of food and food products has been demonstrated for multiple ion sources on mass spectrometers. A non-exhaustive list of publications using mass profiling in this area are shown in Table 1. Most of these

examples required sample preparation to produce viable mass spectra. For instance, organic solvent extraction was necessary for DART of TAGs in milk and milk products.[173] The literature also includes many studies which took advantage of LC-ESI MS/MS for separation and identification of sample components. However, each sample analysis is performed over hours, rather than seconds and so these are of less relevance to this project.

Table 1: MS profiling publications with a focus on food adulteration detection.

Ion source	Sample	Analytes	Reference
OPSI with APCI	Vegetable oils	TG, DG	[174]
DART	Milk/milk products	TG	[173]
MALDI	Vegetable oils	TG	[175]
	Cooking oils	TG	[176]
	Milk powder	TG	[177]
	Animal milk	Phospholipids	[178]
	Olive oil	Phospholipids	[179]
	Animal milk	Peptides (tryptic digest)	[180]
	Animal milk	Proteins	[181]
	Cheese	Proteins	[182]
	Donkey milk	Proteins	[183]
	Animal milk	Peptides, proteins	[184]
REIMS	Meat	Phospholipids	[185]
	Pork	Phospholipids	[186]
	Pork	Phospholipids	[187]
	Fish	Phospholipids	[188]
Flow Injection (FI) ESI	Milk	Proteins	[189]
	Milk	Oxidised lipids	[190]
ASAP	Food from jars	Plasticisers	[191]
Direct Sample Analysis (DSA)	Fruit juices	Organic acids, TG	[192]

Lipids are prominent analytes for determining sample classification as they are highly abundant in crude biological samples. However, they are not species-specific unlike peptides and proteins, and classification relies on their relative abundance in the mass spectrum. Additionally, the biological relevance of lipid classes is only beginning to be understood. This is a limiting property of many current techniques such as REIMS and DART, which have yet to demonstrate the analysis of larger biological molecules, like peptides or proteins. For REIMS, it has been repeatedly demonstrated that with sufficiently intense spectra, rapid and accurate classification models for food analysis can be produced with just phospholipid data. The ease and rapidity with which the data can be acquired, without sample preparation, is unmatched.[185-188] REIMS has yet to be reported for direct

analysis of milk for routine food testing despite the dairy industry being an attractive market, suggesting milk is a problematic medium for the current sampling probes.

Although requiring some sample preparation before analysis, MALDI-TOF MS profiling has the wide analyte range compared to other profiling techniques. MALDI's good salt tolerance also makes it more suitable than direct injection ESI methods, which suffer from considerable signal suppression and background interference in crude samples. Adulteration detection based on protein content of the samples has been demonstrated with solid state MALDI samples.[181-184] The specificity of proteins to the original organism provided powerful distinguishing features in the mass spectra. This is not immediately transferable to the MALDI-QTOF platform without some compromise. Traditional MALDI over an extended m/z range is not entirely suitable for the QTOF platform because of the predominant formation of singly charged ions. The QTOF quadrupole must be capable of transmitting ions of greater m/z values than required for ions generated by ESI. Typical 4k quadrupoles will transmit up to m/z 4000. MALDI-capable QTOF instruments may be equipped with 8k (up to m/z 8000) or even 32k (up to m/z 32000) quadrupoles for high mass analysis.[193] It is also necessary to consider that a greater m/z range results in a lower duty cycle for the TOF analyser, ultimately meaning poorer sensitivity. This may be somewhat negated by ion optics that gate ions efficiently into the pusher region. The applicability of a MALDI source capable of predominantly producing multiply charged ions for peptides and proteins thus offers greater compatibility with the QTOF platform by reducing the analyte ion mass-to-charge ratio. The use of highly charged ions generated by AP-MALDI for milk speciation models will be explored later in this thesis. Another method for protein-rich samples is enzymatic digestion of the raw sample, and analysis of the resulting peptides by MALDI-TOF MS. The challenges for compatibility with the QTOF platform are also reduced by producing the lower mass digest products. This strategy has been employed for determining levels of adulteration in milk samples from a tryptic digest.[180] Cow and goat milk peptide markers were detectable to a 5% adulteration level. These samples are also amenable to LC-MS/MS; however, this is most useful for confidently identifying peptides and proteins rather than rapid sample speciation. A similar approach separating proteins by 2D gel electrophoresis (2-DE) followed by MALDI-TOF MS was able to achieve adulteration detection at the 0.5% level.[194] This technique is relatively labour intensive and is thus less suited for high-throughput analysis.

Some efforts in MS for detection of food product adulteration are focused on the monitoring of the relative intensity ratios of pairs of ions.[178, 179] However, by focusing on just two ions for determining speciation, there is the potential for adulteration of samples on a molecular level to circumvent this type of test. It has already been reported that fraudsters are able to manipulate

the isotopic abundance in synthetic vanilla to 'fool' nuclear magnetic resonance (NMR) analysis into classifying it as a natural variety.[195]

1.6   Analysis of Milk

1.6.1   Molecular Composition

Milk is a complex biological fluid containing lipids, peptides, proteins and carbohydrates, amongst a variety of other biologically important nutrients and minerals.[196] The breadth of molecules means milk is a significant source of biological information. For the example of bovine milk, around 80% of proteins are caseins; phosphoproteins with poor aqueous solubility. The remaining fraction is broadly termed 'whey protein', and comprises β-lactoglobulin and α-lactalbumin, amongst other lower concentration globular proteins. The lipid content includes an abundance of free fatty acids, mono-, di- and triacyclglycerols (MG, DG, TG), phospholipids (PL) and sphingolipids. Lactose represents the major component of carbohydrates in animal milks (cow, goat, sheep, etc.) with an approximate concentration of 5 g/100 g.

Plant milks are similarly complex but have hugely varied composition depending on the source organism. For example, soya milk contains storage and whey proteins totalling 2.9-3.7% of the total composition, whereas rice milk has only 0.1-0.2% protein.[197] Calcium, vitamin B_{12}, B_2, D and E are generally less available in plant milks, so they may be fortified when produced as a food product. Plant milks are generally lower in saturated fats with respect to animal milk but tend to have greater sugar and carbohydrate content (although are lactose-free).

Protein constituents of milk have been extensively investigated by LC-ESI MS[198-200] and MALDI-TOF MS[181-183, 194, 201] to achieve sample characterisation and adulteration detection, but typically require significant sample preparation prior to MS analysis. Studies of the lipid content have been performed with MALDI-TOF MS[177, 202] and DART MS[173] and there are numerous reports of milk lipid analysis methods for LC-MS(/MS).[203-206]

1.6.2   Mass Profiling for Disease Detection: Bovine Mastitis

The molecular composition of milk may differ depending on a cow's breed, lactation number, age, diet and health.[196] The latter is of interest as a source of information for disease diagnosis, as an alternative to other body fluids like blood, saliva and urine. An advantage is that dairy cows' milk is typically collected twice daily, allowing non-invasive, routine sampling. Modern milking machines allow online sampling of multiple animals in tandem, providing an opportunity for high-throughput monitoring of each animal's health. One of the most economically impactful diseases

for dairy farms is mastitis, an infection response to the invasion of an udder teat by a foreign microorganism.[207] While infected, milking is uncomfortable for the animal and milk is unsuitable for consumption due to the presence of clots and pus. An infected animal is isolated from the milk supply chain at the earliest opportunity to avoid reducing the quality of the bulk tank milk. Milk from infected animals usually contains higher concentrations of somatic cells, the levels of which are a measure of the milk grade. Should bulk tank milk Somatic Cell Count (SCC) rise above the regulated value (for example, in the EU, no more than 400,000 cells/mL) the value of the milk is decreased.[208, 209] Antibiotic treatment is administered per animal to combat the infection; however, antibiotics and metabolites must fall below maximum residue limits before milk can be sold. For example, no more than 4 µg/kg of benzyl penicillin is allowed by EU regulation, thus further financially impacting the dairy.[210] Particularly severe or repeat cases can result in the animal being deemed economically non-viable and being destroyed. There is considerable difficulty in determining subclinical mastitis cases as they tend not to produce visual symptoms (reddening of the udder teat, abnormal milk appearance) until they manifest as clinical cases. This delay to diagnosis can allow the disease to be spread through the herd. Subclinical cases (loosely defined at a SSC of approx. >200,000 cells/mL) also lower milk yield from the cow, but a concentration of somatic cells up to approximately 100,000 cells/mL is not deemed to be economically impactful.[208] The development of tests for earlier diagnosis through biological, chemical and physical means is thus an area of research with potential.

Existing detection methods range from "cow-side" to extensive, lab-based, biological tests. Bacterial cultures are typically used to determine the underlying cause of a mastitis infection. This test is performed over a period of days, introducing extra delay before treatment although some farms are equipped with on-site testing facilities.[211] Advantageously, bacteria species can be identified enabling targeted administration of antibiotics. MALDI MS is already established for bacteria identification, and is used in at least one commercial UK laboratory.[212, 213] Faster methods of mastitis detection include online electrical conductivity (EC) assessment, somatic cell count (SCC) and the cow-side California mastitis test (CMT).[214] However, these tests have been reported with low sensitivity (25%, 39.8% and 60.1% respectively) and specificity (87.1%, 84.4%, 62.7%) that shows a significant proportion of false positives.[215] False positives are problematic because they may introduce greater economic burden, as healthy animals may be removed from milking and treated unnecessarily with antibiotics or destroyed. More acceptable rates for both sensitivity and specificity have been published in other studies, but there is clearly significant variability in

the performance of these tests.[216] Visual determination of clots and discolouration may also be performed as a crude test to identify the most severe cases.

Since the data produced on a mass spectrometer are computerised, automatic statistical assessment can allow unbiased, operator-free classification in addition to providing a record of analysis. Classification may be possible from complex LC-MS/MS data or from mass profiles of the unfractionated sample. MS offers the potential for mastitis diagnosis by molecular biomarkers. A validated biomarker for diagnosing mastitis has yet to be described fully, however some have been proposed.[217] Milk exhibits proteins over a large concentration range, making sample preparation and analysis for detecting low abundance proteins difficult.[218] Hogarth *et al.* reported that many of the most abundant bovine milk proteins were absent from 2-DE gels of clinical mastitis samples.[219] Reasons for this were suggested: the proteins may be subject to degradation by proteinases present in the udder from bacteria or as an immunological response, or a reduction in the production of the proteins. The former was disregarded as whey proteins are resistant to proteolysis, and peptides were not identified in the 2-DE gel. However, it is known that endogenous proteases, such as plasmin, will hydrolyse whey proteins.[220]

There has yet to be a peer-reviewed publication reporting direct sampling and mass profiling for mastitis detection. A limited sample preparation for detection of mastitis-causing bacteria without requiring culturing has been reported recently, but with the conclusion that bacterial count was highly influential on the ability to perform correct identification.[221] Initial results for liquid AP-MALDI MS analysis of milk extracts for mastitis detection from this project were presented at the British Mass Spectrometry Society Ambient Ionisation special interest group (BMSS AI SIG) meeting in January 2016. The following year, presentations by Dr Simon Cameron at the same meeting and the American Society for Mass Spectrometry's Annual Conference on Mass Spectrometry and Allied Topics (2017) briefly introduced initial research with a laser ablation(LA)-REIMS probe. Both AP-MALDI and LA-REIMS require a short sample preparation procedure to produce diagnostic spectra; experiments with the former are the basis of much of the Results chapter.

1.7   Project Aims

The overarching aim of the project was to develop AP-MALDI for improved analysis of liquid samples on a high-throughput scale. This was to take the form of AP-MALDI ion source modifications, the use of modern MS instrumentation capabilities, sample preparation developments, investigation of method applications and the use of new software techniques. The vision was to develop proof-of-concepts, taking advantage of a combination of the above-mentioned points, to demonstrate the applicability of liquid AP-MALDI MS for solving analytical problems. No application was prescribed by the project but rather were to be developed with reference to contemporary literature, research group interests and the direction of the industry in general. The main area of interest was rapid mass profiling, especially since the FDA approval of two MALDI-TOF MS systems (Section 1.3.1) for clinical use showed the potential for commercialisation of MALDI as a clinical diagnostic method. In addition, the attention mainstream media were paying to the iKnife drove my interest in a similar type of analysis workflow. Harnessing the ability to produce multiply charged ions was not stipulated, but when considering applications our ability to produce these ions was always considered.

2   Materials, Instrumentation and Methods

2.1   Materials

2.1.1   Chemicals

Water, methanol (MeOH), acetonitrile (MeCN), isopropanol (IPA) and trifluoracetic acid (TFA) (all HPLC-grade) were bought from Fisher Scientific (Loughborough, UK). Lipid standards PC (16:0/18:0) and SM (d18:1/12:0) were purchased from Avanti Polar Lipids (Alabaster, AL, USA). Sodium iodide calibration solution and ETD reagents were bought from Waters (Manchester, UK). Peptide Calibration Standard II was purchased from Bruker (Coventry, UK). HPLC-grade n-hexane, all other matrix components, reagents, calibrants and analytical standards were acquired from Sigma-Aldrich (Gillingham, UK).

2.1.2   Milk Samples

Bovine milk (Tesco British skimmed and whole, pasteurised and homogenised), whole caprine milk (St Helen's, Yorkshire, UK, pasteurised and homogenised) and unsweetened soya milk (Alpro, Wevelgem, Belgium) were obtained from Tesco. Raw and pasteurised ovine milk was kindly donated by The Sheep Milk Company (Lancashire, UK). Each milk type was aliquoted and stored at -80 °C until analysis. Existing research has reported that milk from the same species may exhibit variation depending on the source, however no clustering was observed in PCA of samples collected over four years from a range of producers.[173] Thus, the milk samples here are justified as representative of each species.

For the mastitis investigation, raw bovine milk was sampled from individual dairy cows at the Centre for Dairy Research (CEDAR), a facility operated by the University of Reading. Samples were taken from a single quarter of the udder, unless collected on the milking carousel. These were aliquoted and stored at -80 °C until analysis. Samples were collected intermittently from November 2015 (initial samples collected on the carousel) to December 2017. Mastitis diagnosis was determined by experienced personnel by cow side evaluation and should be considered 'clinical' cases. Milk was visually assessed by fore milking each teat prior to attachment of milking apparatus. If clots were present, the cow was determined to have mastitis. Additional microbiological characterisation was undertaken by Quality Milk Management Services Ltd (QMMS), a UK laboratory equipped with MALDI MS for identifying microorganisms cultured from milk. Figure 12 shows examples of milk from mastitis-diagnosed cows displaying significant visual differences.
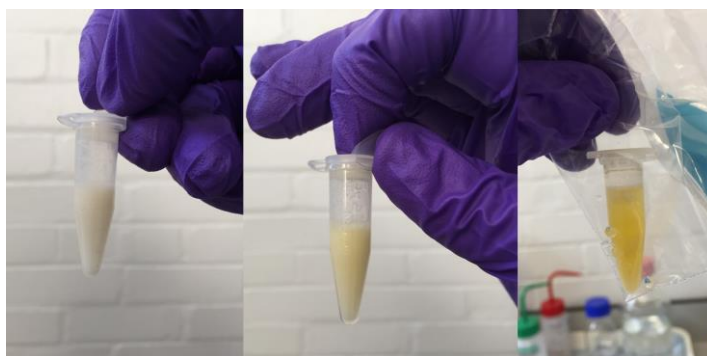
Figure 12: Three examples of milk samples from mastitis-diagnosed cows.

2.2 Liquid support matrix (LSM) composition

MALDI LSMs used throughout the project are detailed in Table 2.

Table 2: LSMs used throughout the project.

ID	Chromophore	Concentration (mg/mL)	Solvents (v:v)	Liquid Support (µL added to 100 µL chromophore solution)	Applicable Ion Mode
A	DHB	25	MeCN/H$_2$O (7:3)	EG (60)	+/-
B	DHB	25	MeCN/H$_2$O (7:3)	GL (60)	+/-
C	CHCA	5	MeCN/Amm. Phos. (7:3)	EG (60)	+
D	CHCA/TEA	100/100 (µL)	MeOH/H$_2$O (1:1)	GL (20)	+
E	3-AQ	10	MeCN/Amm. Phos. (7:3)	EG (60)	-
F	9-AA	10	MeCN/Amm. Phos. (7:3)	EG (60)	-
G	None	N/A	MeCN/H$_2$O (7:3)	EG (60)	+/-

Abbreviations: DHB; 2,5-dihydroxybenzoic acid, CHCA; α-cyano-4-hydroxycinnamic acid, TEA; triethylamine, 3-AQ; 3-aminoquinoline, 9-AA; 9-aminoacridine; EG; ethylene glycol, GL; glycerol.

Matrix **A** was used frequently, as it was broadly applicable to a range of applications and provided good analyte signal intensity in both positive and negative ion mode. **B** was also used broadly but was superseded by **A** in most cases. The only difference between **A** and **B** was the type of liquid support, ethylene glycol (EG) and glycerol (GL), respectively. EG-based LSMs were recently shown to provide greater ion signal intensity than the GL-based equivalents.[222] **C** was used for enhanced high mass protein analysis.[222] **D** was used on the QTOF Ultima Global (see Section 2.3.1), where the ion source did not feature a heated inlet. No other investigated LSMs were found to be compatible with this mass spectrometer's AP-MALDI source. **E** and **F** outperformed other matrix preparations, in terms of signal intensity and higher charge state production, in the negative ion mode, particularly in the analysis of peptides and proteins.[223] **G** did not contain a UV-absorbing matrix chromophore and was required to ionise sodium iodide for extended-range (> *m/z* 2000) TOF calibration (see Section 2.3.3). There is potential for this LSM to be applied further, but it is somewhat more difficult to use than the other LSMs, as discussed in Section 3.3.3.

2.3   Instrumentation

2.3.1   Mass spectrometers and ion sources.

Three mass spectrometers were used throughout the project. Due to the prototype nature of the
AP-MALDI ion sources many instrumental procedures differed from standard operation.
Preliminary work was carried out on a QTOF Ultima Global (Waters, Wilmslow, UK), retrofitted
with an API inlet and an in-house developed AP-MALDI source. The internals of this instrument
are akin to those depicted in Figure 1. On project commencement, the condition of the
instrument was largely unknown, and initial performance checks with leucine-enkephalin solution
infused by ESI showed a peak with approx. 50 Da width at the baseline. The TOF pusher power
supply was replaced, which brought the instrument to a usable state. The first stacked-ring ion
guide also required substantial cleaning to reduce spontaneous in-source analyte fragmentation
due to contamination (Figure 13). The 'Global' ion block was also replaced with an 'API' ion block
because of a broken isolation valve mechanism and degraded seals.
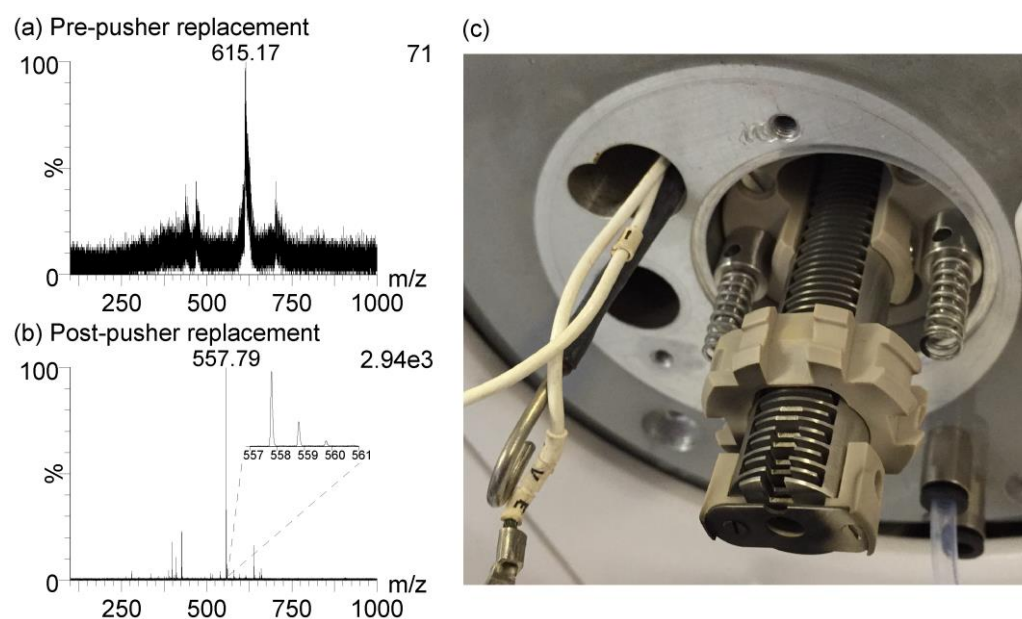
*Figure 13: Uncalibrated QTOF Ultima ESI spectra for leucine-enkephalin (a) before and (b) after pusher power supply
replacement. The first stacked ring ion guide (c) showed significant contamination.*

The source featured a 355 nm laser (FlareUV, Lumanova) with a max pulse rate of 1 kHz and later,
because of a laser head failure, a VSL-337-ND-S (Spectra-Physics, Mountain View, CA, USA)
337 nm nitrogen laser that was recovered from a decommissioned instrument. A schematic of the
laser optical setup is presented in Figure 14. The MS inlet was not modified, i.e. the standard API
sampling cone was in place and the MALDI sample plate positioned within 5 mm of the cone
(shown in Figure 15). Since the source was modified without much of the original housing, a
resistor that modified the 'Source ID' voltage to 2.9 V for the source interlock was added to allow
instrument operation as if a nanoESI source was present. Push switch interlocks were defeated as

required. For this and all other instrument modification work, the appropriate health and safety (H&S) assessment and documentation were in place. The MALDI sample plate was fitted to a thumbscrew-controlled micrometre XYZ-stage by a polytetrafluoroethylene (PTFE) adaptor, and a high voltage line (providing up to ±4 kV) was connected. The TOF was operated in 'V-mode' for optimal sensitivity. Matrix ions were detected from solid state CHCA samples, prepared from a saturated solution in acetone, in the first demonstration of the ion source. Only a single LSM (**D**) of those tested was found to be compatible with this source, possibly due to the lack of additional desolvation. Unfortunately, one quadrupole RF generator was defective meaning MS/MS was not possible with this instrument as the quadrupole could not isolate ions with a specific *m/z*. Despite its problems, the QTOF Ultima Global served as a substantial learning experience and a platform for developing the first proof-of-concept method for this project.

Figure 14: Optical diagram for the AP-MALDI source mounted to the QTOF Ultima. Components are not to scale. Ghost beams are omitted for clarity.
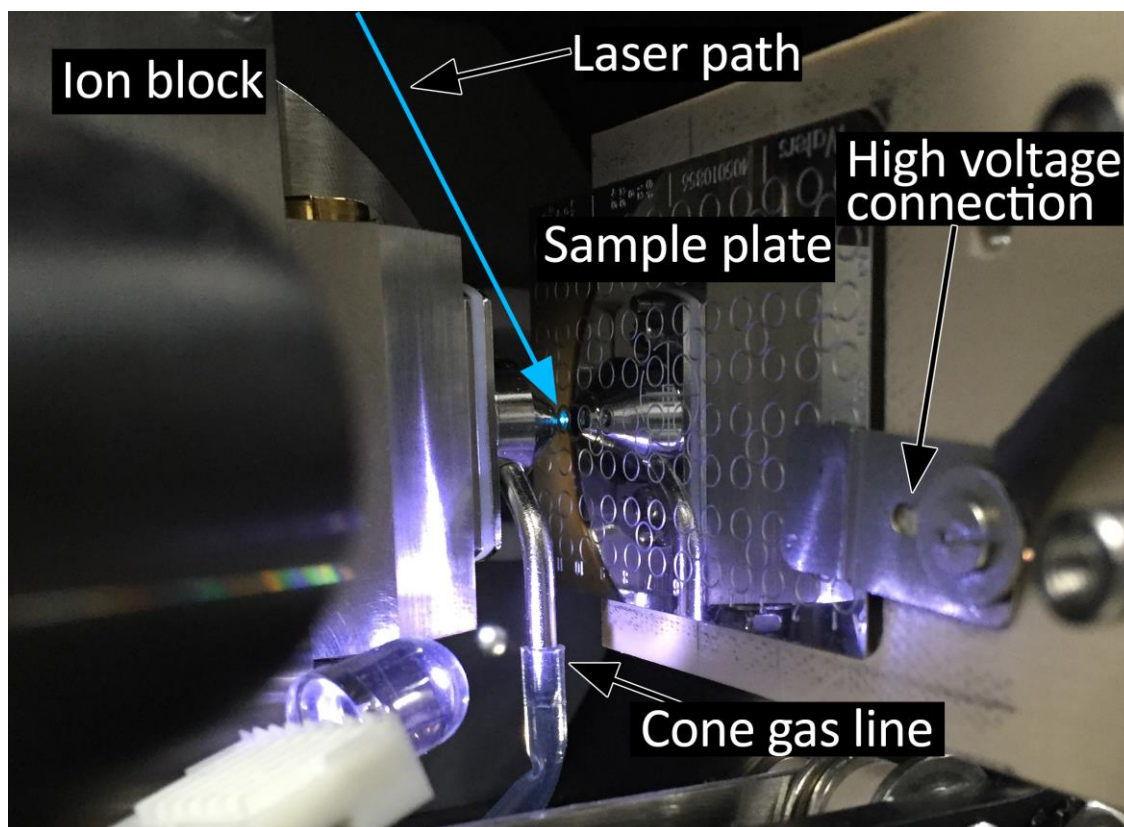
Figure 15: Photograph of the QTOF Ultima AP-MALDI source assembly. The DHB-containing sample is shown fluorescing upon exposure to the UV laser beam.

The majority of MS analysis was performed on a Synapt G2-Si HDMS (Waters), a state-of-the-art, Q-IM-oaTOF MS that was installed in the laboratory six months prior to this project's commencement. Important features were the StepWave ion guide, for effective removal of neutral species from the ion beam; an 8K quadrupole for transmission of an extended *m/z* range (up to approx. *m/z* 8000); the TriWave device, allowing for ion mobility measurements, ETD, and two stages of CID; and a high-resolution TOF analyser (mass resolving power of 10,000 – 40,000 measured on the [M+2H]$^{2+}$ ion of [Glu^1]-Fibrinopeptide B, mode dependant) for sensitive, accurate mass (<1 ppm) measurements. The instrument was typically operated in ion mobility-TOF mode, with 'sensitivity' ion optics in the desired polarity.
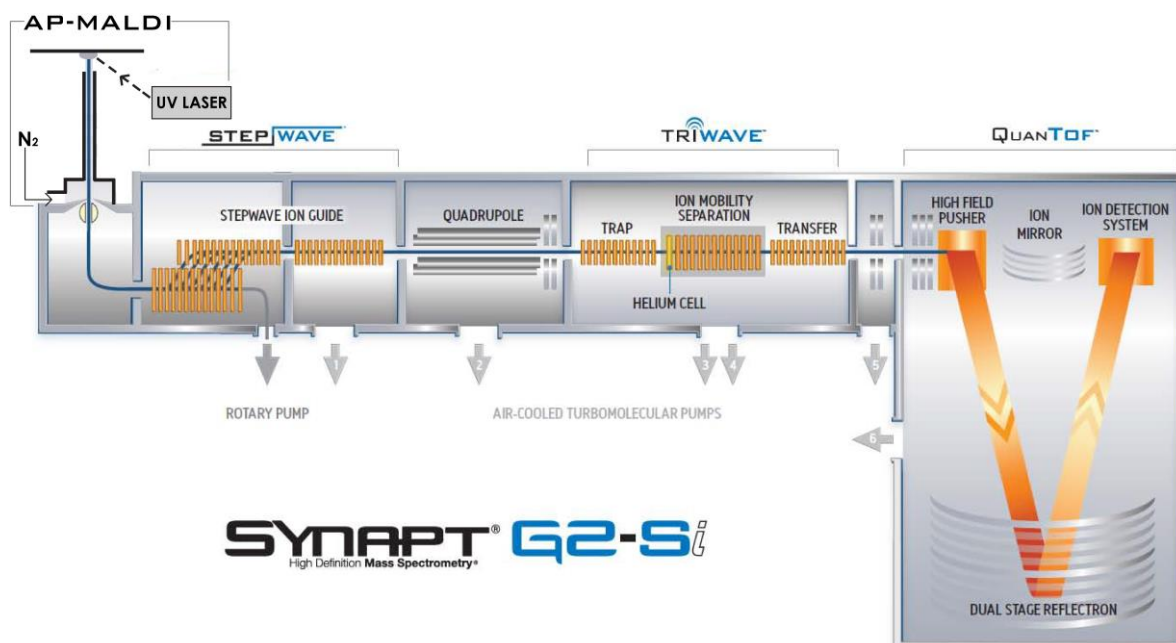
Figure 16: Schematic of the Synapt G2-Si modified to show the newly developed AP-MALDI source attached.[109] The TOF ion beam is depicted for sensitivity optics. Original schematic adapted from a presentation by Eleanor Riches[224]

The 'Synapt' was also modified using a custom-built AP-MALDI ion source, differing from that of the QTOF Ultima Global's in that a custom heated ion transfer tube was fitted to the API inlet as described previously.[109] The main laser was an MNL100 nitrogen laser (LTB Lasertechnik Berlin, Berlin, Germany) and was operated at a repetition rate of 1-30 Hz with approximately 20 μJ/pulse. A second laser, a 355 nm Nd:YAG FlareUV 200-100 Waters A1.0 (InnoLight GmbH, Hannover, Germany) with a pulse rate of ~0.65-1 kHz was added towards the project's end. The lasers were mounted to an optical breadboard atop the MS, with optics transmitting the beams to the ion inlet, as depicted in Figure 15. UV optics, optical breadboard and kinematic components were purchased from Thorlabs (Newton, NJ, USA). The laser pulse energy was measured with a J9LP pyroelectric Joulemeter (Coherent, Santa Clara, CA, USA).
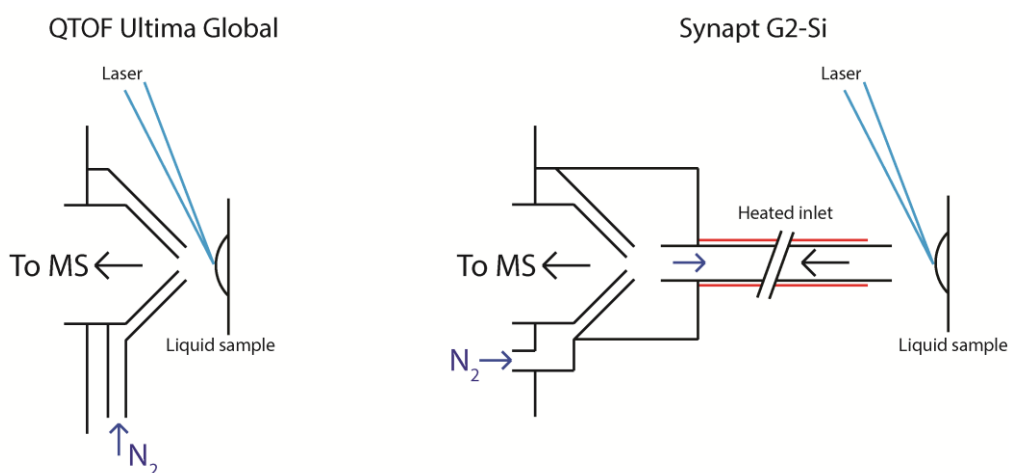
Figure 17: schematic comparison of the inlets of the QTOF Ultima Global and Synapt G2-Si. The heated inlet promoted ESI-like desolvation of liquid droplets aiding the formation of multiply charged ions.
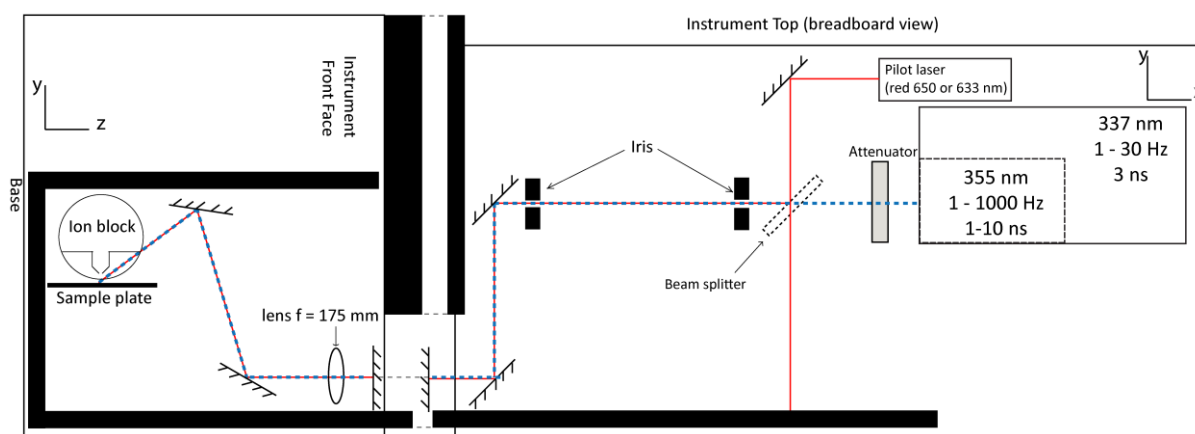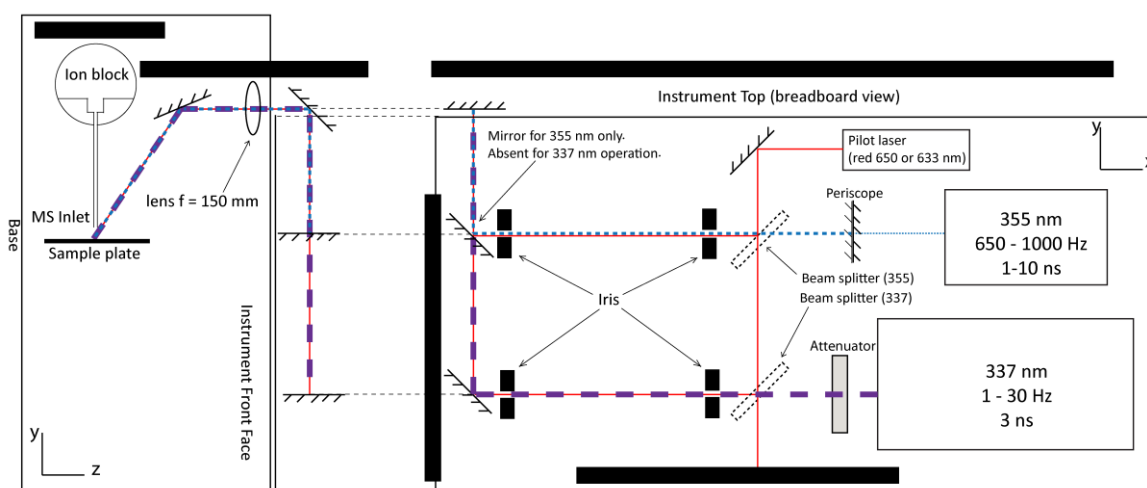

Figure 18: Optical diagram for the AP-MALDI source mounted to the Synapt G2-Si. Components are not to scale. Ghost beams are omitted.

The AP-MALDI ion source was recognised as a nanoESI source by the Synapt, allowing access to the standard nanoESI software controls. This was achieved by ensuring the 'Source ID' voltage was supplied detected as 2.6 V by addition of a 1 kΩ resistor. Other push switch interlocks were defeated to allow operation of source gases and high voltages.

The MALDI sample plate was mounted to an electronically controllable XY-stage assembly (Zaber Technologies Inc., Vancouver, Canada) and a thumbscrew-controlled stage for the Z-axis. A PTFE MALDI sample plate holder with magnets for fixing the metal MALDI sample plate was attached to the XY stage. The Zaber stages were connected to the Synapt control computer by USB for software communication. Three software control schemes were implemented:

1. Zaber Console: The manufacturer's software, providing a graphical user interface (GUI) and scripting environment for stage control. Information such as stage coordinates were available here.

2. Custom-written software with a GUI (developed by former group member Dr Pavel Ryumin) suited to selecting locations on a single sample with mouse clicks. This provided a more convenient way of selecting wells than the Zaber Console for day-to-day tests.

3. Scripts run from within Waters Research Enabled Software (WREnS), enabling control of the stages by the mass spectrometer. This was the control scheme of choice for large sample batches analysed under a fixed set of conditions (see section 2.3.4 and Appendix 2).

The potential difference between the MALDI sample plate and the ion transfer tube was typically set to 3-3.5 kV, the resistance wire around the heated ion transfer was set between 26 and 30 W,provided by a low voltage DC power supply, and the cone voltage at 30-40 V. The ion transfer tube was treated as disposable, since crude samples would slowly contaminate the inlet. Baseline expected performance was accepted on achieving 1e4 counts per second measured on the $[M+2H]^{2+}$ ion of $[Glu^1]$-Fibrinopeptide B from a 100 pmol droplet with LSM **A**. When this was unobtainable the tube was replaced. Counter-flow N_2 gas was provided through the heated ion transfer tube up to a rate of 200 L/h as required (analyte-dependent). The approximate transfer tube temperature, at various N_2 flow rates, has been published previously.[109] A schematic for the ion inlet is shown in Figure 19, with a photograph of the setup shown in Figure 20. A variety of solid and liquid MALDI samples were compatible with this source, including all of those in Table 2.
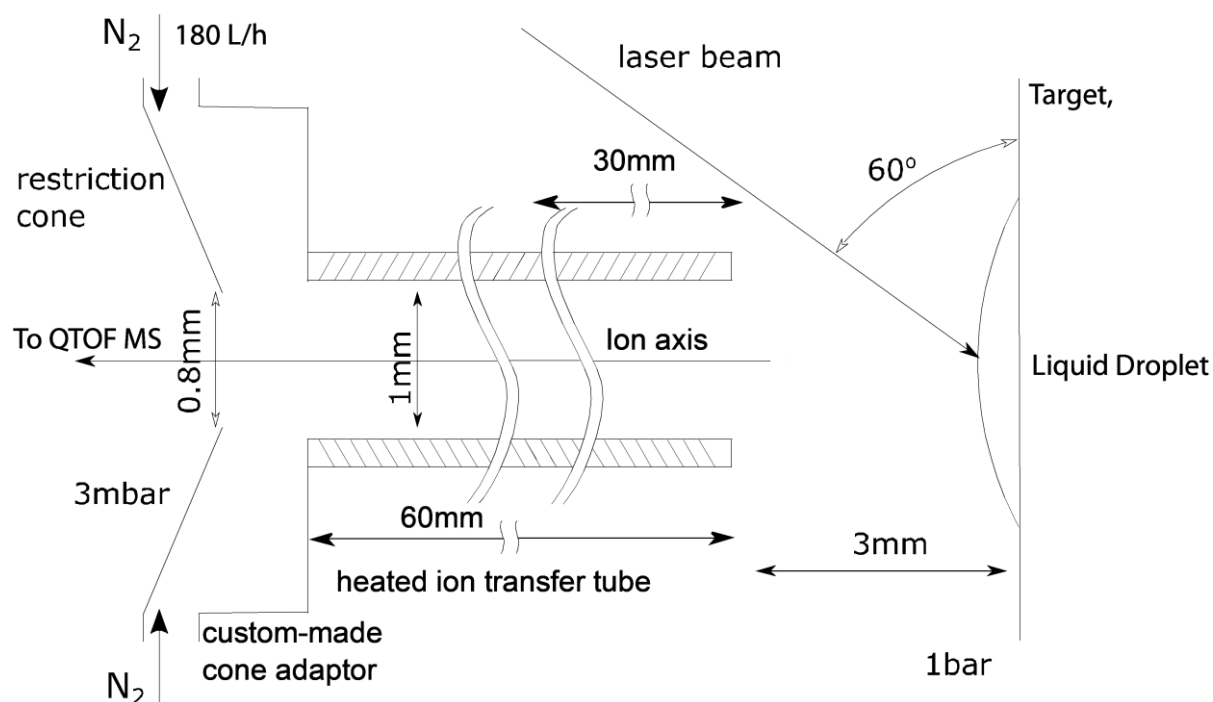
Figure 19: Schematic of the MALDI sample plate and ion inlet on the Synapt G2-Si AP-MALDI source.[109]
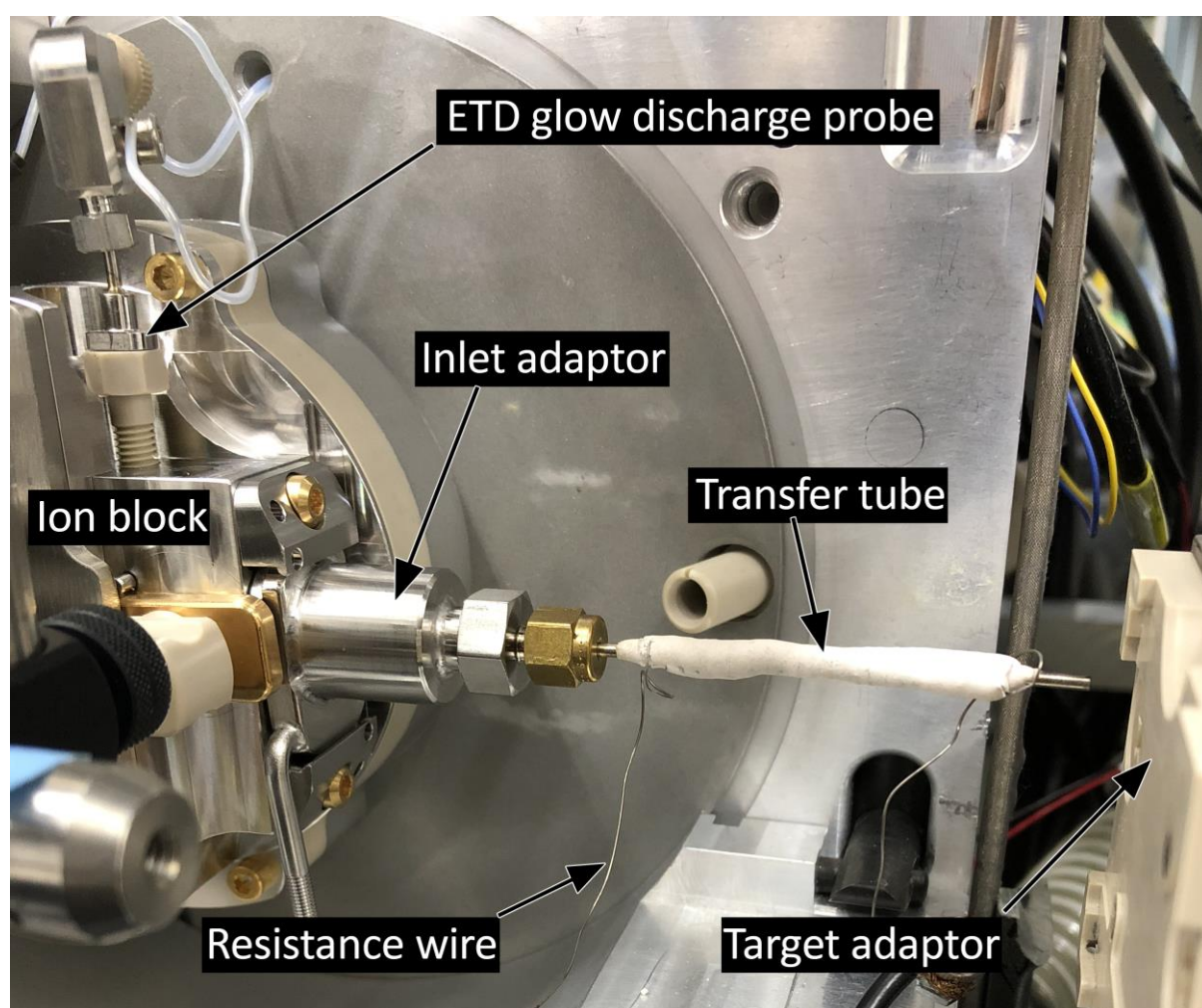
Figure 20: Photograph of the Synapt AP-MALDI source inlet assembly.

Important gases for the ion mobility-TOF MS(/MS) experiments included the drift gas in the travelling wave ion mobility spectrometry (TWIMS) cell, which was N_2 to give a cell pressure of approx. 2.5 mBar, the trap and transfer cell gas for CID, which was argon, to give cell pressures of approx. 0.02 mBar, and the helium cell gas, which was set to approx. 3.8 mBar (*cf.* Figure 5). Specific details of each experiment are discussed in the relevant sections.

For comparative purposes, a QTOF Premier (Waters) was used. It featured the standard intermediate vacuum (<0.1 mBar) MALDI source. The laser was a 337 nm nitrogen VSL-337ND-Si (Spectra-Physics). The TOF was operated with 'V-mode' ion optics. Liquid samples prepared for this mass spectrometer first underwent drying in a rough vacuum chamber to limit outgassing when loaded.

In general, open laser beam experiments were only undertaken, and high voltage interlocks defeated, after consultation with H&S staff at the University, and in accordance with the site-specific H&S regulations. All open beam pulsed lasers were secured with physical or digital keys and interlocked to the laboratory door to ensure immediate shutdown in the event of unauthorised or accidental access.

2.3.2 Lasers and Optics.

This section refers to the components of the diagrams in Figure 14 and Figure 18. Five pulsed UV lasers were used for experiments across the three mass spectrometers. The MNL100 nitrogen laser (with a wavelength of 337 nm, max. pulse rate of 30 Hz, and a pulse width of approximately 3 ns) was used for most of the research reported in this thesis. Nitrogen lasers often have a rectangular beam shape with a corner missing due to the output coupler in the laser cavity (Figure 21). Energy distribution through the beam cross section is relatively even, sometimes described as

a 'top hat'. Generally, the beam quality will decrease with laser age and number of pulses due to leakage from the gas cell, although this is a serviceable part.
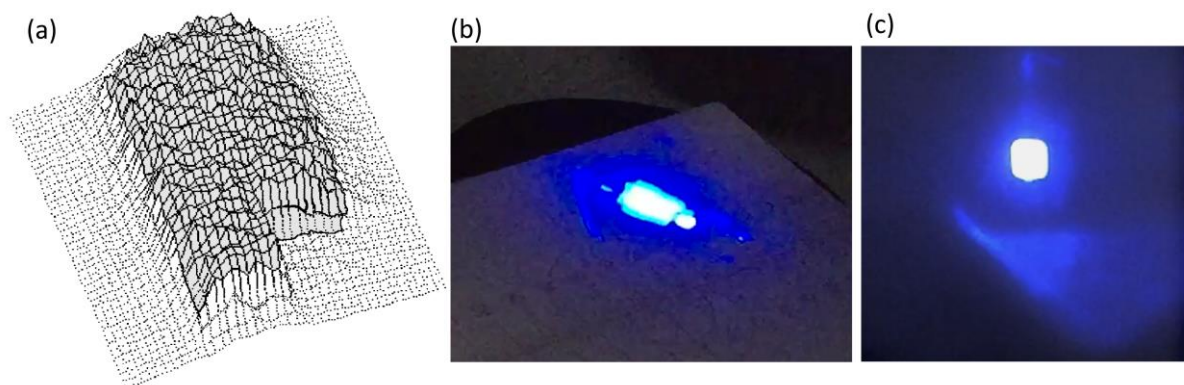
Figure 21: Pulsed N$_2$ 337 nm lasers; (a) schematic from VSL-337-ND-S manual[225] (b) photograph of the fluorescence produced by irradiation of a white card with this laser, and (c) fluorescence from a white card produced with the MNL100 laser.

The nitrogen laser on the Synapt was fitted so that the beam was 10 cm above and parallel to the breadboard surface. This design consideration left space for optical component mounts. The beam was attenuated with a neutral density filter attached to a stepper motor, which was controlled by an Arduino Uno (Arduino, Somerville, USA) microcontroller and auxiliary 12 V DC-power supply. The 3.3 V output of the Arduino Uno was used in place of 2x 1.5 V batteries to power the red pilot laser. A remote switch was added to the pilot laser to allow for easier operation.

The second pulsed laser attached to the Synapt was the Innolight FLARE PQ UV 1k-30, frequency-tripled Nd:YAG (355 nm) laser and introduced much later in the project, following research into high-repetition rate lasers by group member Jeffrey Brown.[132] It was added to the Synapt setup in tandem with the 337 nm laser. An additional mirror was included to swap from the 337 nm to the 355 nm beam. This solid-state laser was vendor-optimised for a maximum pulse rate of 1 kHz, although responded to transistor-transistor logic (TTL) triggers down to 650 Hz. It featured a beam expander immediately after the laser aperture, and an elliptical beam shape with a cone-like energy profile (Figure 22).
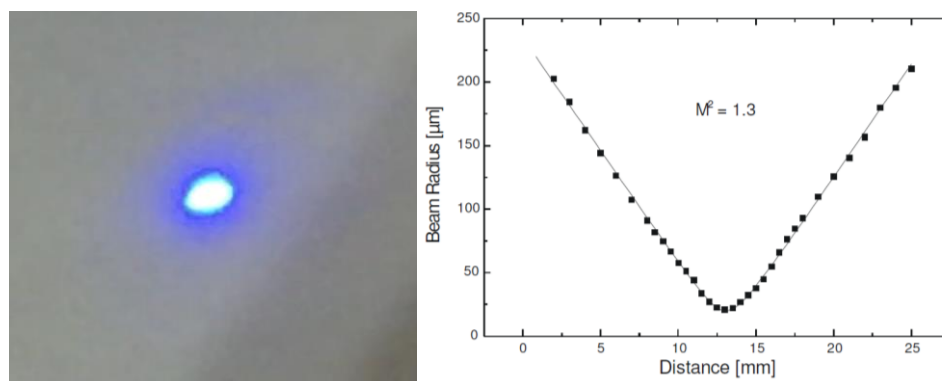
Figure 22 (a) Fluorescence produced by a white card in the path of the 355 nm FLARE PQ laser beam showing the elliptical beam profile. (b) Energy profile for the 1064 nm model of the FLARE PQ laser, reprinted from the manufacturer's documentation.[226] An equivalent for the frequency tripled 355 nm wavelength was not available. Squares indicate measurements, whilst the line is calculated.

While this setup left a laser pulse repetition range of 31-649 Hz that could not be investigated between the two lasers, the high and low ends of the interesting pulse repetition range were both accessible. The Nd:YAG laser head was attached to the breadboard, which functioned as a large heatsink and electrical ground. A periscope device was used to raise the beam to 10 cm above the breadboard, ensuring compatibility with existing optics used with the 337 nm laser. The energy output for the 355 nm wavelength was measured to be approximately 30 µJ per pulse at 1 kHz, considerably lower than the laser label's value of 0.5 mJ per pulse. An attenuator was not fitted as the laser beam's energy was further reduced due to the inherent transmission losses of the optical components. For MALDI experiments, the laser energy per pulse reaching the sample plate was in the range of 20 – 25 µJ, depending on the laser pulse repetition rate.

With both UV lasers operating at the same height, optical components were shared by both beams and only a single pilot laser was required. The optical diagram in Figure 18 depicts how the pilot laser beam was added to the UV lasers' paths with beam splitters (approx. 10% reflective). Irises were used to ensure alignment of the UV and pilot beams was parallel to the breadboard surface. The pilot laser was then directed through the lens onto the MALDI sample plate. The laser beam was checked to be positioned centrally on the optical components for best fine-tuning flexibility.

The QTOF Ultima Global only featured a single laser at a time and was thus simpler in its laser set-up. The laser was operated at a maximum pulse frequency of 30 Hz (337 nm or 355 nm).

The QTOF Premier's laser apparatus was unmodified from the factory default and typically operated with a laser energy set between the arbitrary values of 200 and 500. The maximum pulse frequency used was 20 Hz.

2.3.3 Instrument Calibration

The QTOF Ultima Global was calibrated with a mixture of polyethylene glycol (PEGMIX: average Mw 200, 600 and 2000 g/mol) ionised from liquid samples with matrix **D**. The QTOF Premier was calibrated with the same calibrant but using a standard MALDI dried droplet preparation with CHCA as the chromophore.

For lipid-only analysis on the Synapt in both ion modes, the ion mobility cell and TOF analyser were calibrated with polyalanine. A greater *m/z* range was acquired for protein and peptide-containing samples, so the TOF was calibrated with sodium iodide (2 µg/µL in water/IPA 1:1) up to *m/z* 5000. Interestingly, the latter calibration required a sample prepared without a matrix chromophore (matrix **G**) as otherwise no ions relating to NaI were detected (Figure 23). It is possible that the NaI ions are formed exclusively by ESI processes and the MALDI matrix compound interferes. The NaI solution (0.5 µL) was added to 0.5 µL of a solution of water/MeCN/ethylene glycol (3:7:6; v:v:v) on the sample plate. The laser was set to a pulse repetition rate of 10 Hz and focused on the droplet edge to achieve desorption and ionisation. Typically, data were acquired for three minutes with steady ion generation to produce intense spectra suitable for calibration to an accuracy of approx. 1 ppm.
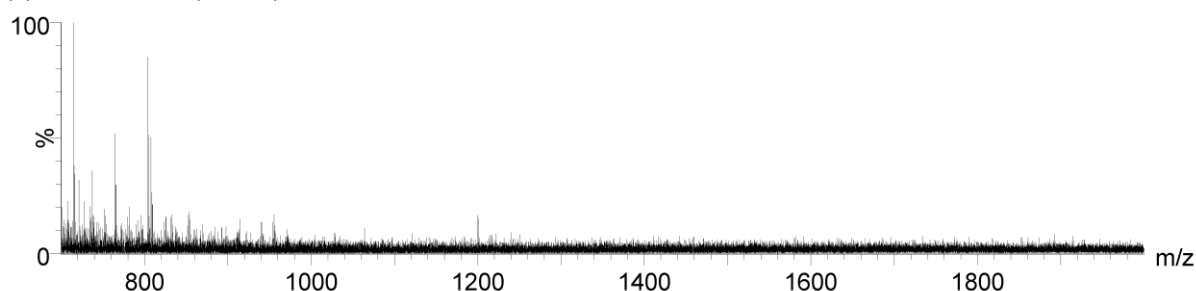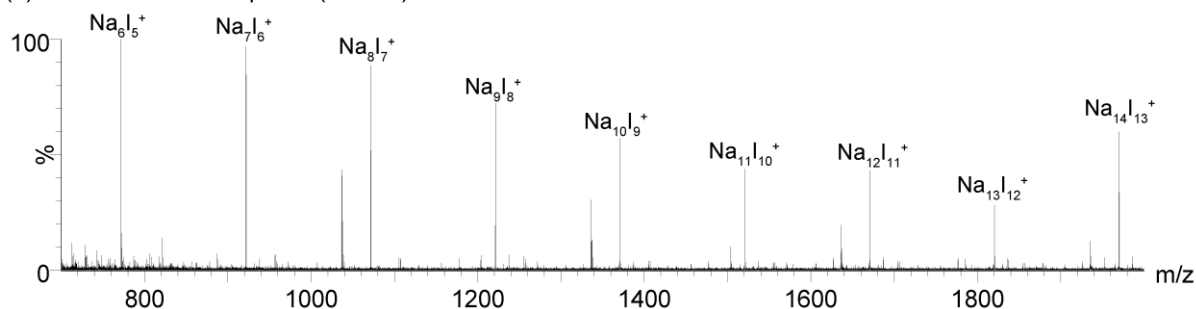
Figure 23: Liquid AP-MALDI-TOF MS spectra for m/z 750-2000 acquired over one minute for samples of sodium iodide with (a) DHB LSM A and (b) LSM G.

2.3.4 Synapt G2-Si setup for automated sample acquisition

The MALDI sample plate was inserted into the PTFE stage adapter. The XY coordinates of the sample well "A2" were recorded from the Zaber Console. Since "A1" was not an accessible well due to the stage travel distance, its value was calculated and added to the WREnS script

"Arbitrary Wells Acquisition" (see Section 6.1). The script referenced "A1" coordinates to calculate values for all other wells of the sample plate. The mass spectrometer was put into "Operate" and the ion source parameters were set. The WREnS script was saved, built and run. A new acquisition was started through the MS Tune window in MassLynx. On starting the acquisition, a new data file was created in the MassLynx project "data" folder, which triggered the WREnS script to initiate control of the XY-stage. The laser was started manually once the stage positioned the first sample well in front of the mass spectrometer inlet and was run until the acquisition completed.

2.4 Software and data processing

MassLynx 4.1 was used to acquire data from the Waters mass spectrometers and process mass spectra. DriftScope versions 2.4, 2.7 and 2.8 were used for post-acquisition ion mobility filtering and CCS calibration. Abstract Model Builder [Beta] (AMX, version 1.0.1563.0), prototype software developed by Waters, was used for most multivariate model building and validation from TOF MS spectra. The LIPID MAPS database and Waters Metabolic Profiling CCS Library were accessed from within Progenesis QI (v2.2, Nonlinear Dynamics, Newcastle Upon Tyne, UK) for lipid identification.[227] Progenesis Bridge (v.1.0.29, Waters) was required to process profile spectra prior to importing into Progenesis QI. The EZInfo plugin (v3.0, Umetrics, Umeå, Sweden) was also accessed from Progenesis QI for additional multivariate data analysis tools. Data were plotted with MATLAB R2017a and Microsoft Excel as necessary for visualisation or raw outputs (e.g. 3D plots). Microsoft Visual Studio 2015 was used for developing attenuator and stage communication software. MarathonControl (v.1.60) was used to control the MNL100 laser. UniDec (University of Oxford, Oxford, UK) was used to deconvolute multiply charged protein and peptide mass spectra.[228] Microsoft Excel and Mmass[229] were used for Kendrick mass defect (KMD) calculation. Where used in this thesis, the KMD was calculated with Equation 9:

$$KMD = \left(IUPAC\ mass \times \frac{14.00000}{14.01565} \right) - nominal\ mass \qquad \text{Equation 9: Kendrick Mass Defect}$$

2.5 Determination of structural information from metal-cationised phospholipids

Individual methanolic solutions of phosphatidylcholines (PCs) (2 pmol/μL, 200 pmol/μL for TAP MS) and an ethanolic solution of sphingomyelin (SM) (d18:1/12:0) (2 pmol/μL) were prepared. Multivalent metal salt solutions ($MgCl_2$, $CaCl_2$, $SrCl_2$, $BaCl_2$, $MnCl_2$, $CoCl_2$, $ZnCl_2$, $FeCl_2$, $AlCl_3$, $ScCl_3$, $CrCl_3$ and $LaCl_3$; 1 pmol/μL – 1 μmol/μL) were prepared in water. A hexane/isopropanol (3:2 v:v) extraction of soya milk was prepared as in Section 2.6. Sample droplets were added to a stainless-

steel MALDI sample plate by addition of LSM **A** or **G** (500 nL), salt solution (500 nL) and analyte solution (500 nL) (in that order) to a well. For CID, argon was the collision gas. Nitrogen was the drift gas for ion mobility measurements. TAP experiments were performed according to the workflow in Figure 24. It was not necessary to adjust the TriWave settings manually for these experiments.

```
                                 ─────────── TriWave ───────────
                          ┌──                                 ──┐
  Source      Quadrupole  │ Trap          IMS          Transfer │   TOF

  All ions ─→  Selection of [L+Ba]²⁺ ─→ Fragmentation ─→ Separation of 1st ─→ Fragmentation ─→ Detection of
                                        CE = 22 V       generation products   CE = 40 V        1st and 2nd generation
                                                        by drift time.        Drift time       products with aligned
                                                                              maintained.      drift time.
```
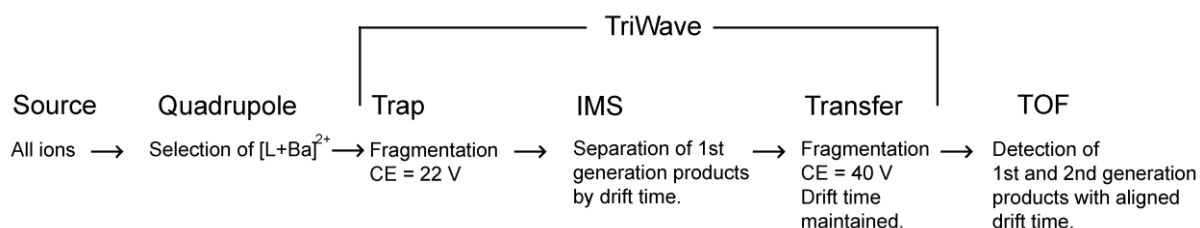
Figure 24: Schematic for the TAP MS workflow

The reagent 1,4-dicyanobenzene was used for ETD experiments with nitrogen as the make-up gas (25 mL/min). A specific ETD ion block was fitted to the Synapt's source to allow introduction of the reagent through a glow discharge needle. The trap gas flow was manually set to 14 mL/min, and the fill time set to 0.1 s, for *m/z* 128 (negative reagent ions). The trap DC wave was set to 15 V for trapping and lowered to between 0.2 V and 1.5 V for ion-ion reactions. Supplementary collisional activation was not used. Initial ETD setup was performed with an ESI source and a solution of substance P (50 fmol/µL in water/methanol 1:1) to confirm fragmentation efficiency (at least 2%) on the $[M+3H]^{3+}$ ion. All AP-MALDI experiments in this section were conducted with the heated inlet counter-flow gas set to 180 L/h and approximately 200 °C.

2.6   Phospholipid sample preparation ("HIP" extraction)

N-hexane/IPA (HIP, 450 µL, 3:2, v:v) was added to 50 µL milk. The mixture was vortexed for approximately 5 seconds then centrifuged at 13,000 rpm for two minutes at room temperature. A volume of 750 nL of the clear, colourless supernatant was added to 750 nL of LSM, which was previously spotted on a stainless steel MALDI sample plate. This ensured a domed droplet morphology, filling the entire well, which was seemingly advantageous for automation as it meant the centre of the well would reliably produce strong signal. Homogenous sample morphology also reduced variability in the laser incidence angle between samples, a factor which is suspected to

influence overall TIC. The samples were each analysed for ten seconds by automated AP-MALDI MS to ensure strong signal intensity for best compatibility with software tools used later.
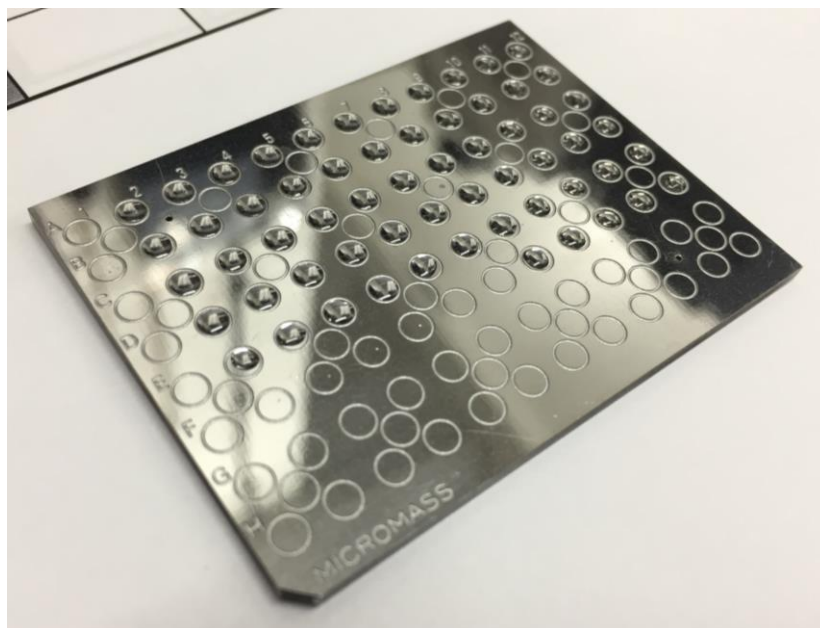
*Figure 25: Liquid MALDI droplets of DHB/glycerol LSM (**B**) and HIP lipid extract.*

2.7   Phospholipids, peptides and proteins analyte extraction.

Aqueous trichloroacetic acid (TCA) solution (250 µL, 5% in H_2O, m:v) was added to 50 µL of milk. The mixture was vortexed, followed by centrifugation for 2 minutes at 13,000 rpm and room temperature. The supernatant was removed by a Pasteur pipette and discarded. The precipitate was dissolved in sodium hydroxide solution (200 µL, 0.5 mM, water/isopropanol 1:1). A few seconds of sonication was typically required for complete resuspension. The final analyte solution was cloudy and white, not dissimilar to milk in appearance. Samples were spotted as in section 2.3 and analysed for one minute each by automated AP-MALDI. The extra analysis time was required to build signal for the protein ion signals specifically, since they were generally much weaker than the phospholipid signals.

2.8   Antibiotic extraction and analysis.

A solution of the antibiotics amoxicillin, ampicillin, cloxacillin, methicillin, penicillin G and cefalexin (30 µL) was spiked into water or bovine skimmed milk (50 µL) for a final concentration of 62.5 pg/µL of each antibiotic. A β-lactamase solution (1 µL, 1.5-3.0 units/µL) prepared in Trizma solution at pH 7.4 with BSA filler protein (approx. 1 mg/mL) was added to half of the sample batch to mimic the presence of β-lactamase-excreting bacteria and hydrolyse the antibiotics. All samples were incubated at 30 °C for two hours without further modification to the solution's pH.

Aqueous TCA solution (170 µL, 5% m:v) was added to each sample, which were then vortexed and centrifuged at 13,000 rpm for two minutes. The supernatant was the final analyte solution. An AP-MALDI parallel reaction monitoring (PRM) ion mobility-TOF MS/MS method was developed for confidently detecting the antibiotics and their hydrolysis products by their precursor and CID product ions. Since most of the antibiotics were supplied as sodium salts and no further purification was undertaken, the sodium adduct ions were generally preferential precursors. PRM was used to monitor up to three product ions of each antibiotic, to improve specificity and sensitivity.[230] In contrast to a triple quadrupole MS, all transitions were monitored on the same scan. The PRM scan list is shown in Table 3. Data for ions not specified in the list were stripped from the MS/MS functions i.e. only signals for monitored precursor and product ions were retained. The quadrupole was set to transmit a narrow window of approximately m/z 1 for the precursor since the analytes were low mass, and thus mixed with much of the MALDI background signal; low mass (LM) resolution was set to 12.0 and high mass (HM) resolution to 15.0 on the instrument tune page.

Table 3: PRM scan list for the antibiotic detection method.

PRM scan list file	Product Ions (m/z)			Transfer collision potential ramp (V)		
Set m/z	1	2	3	Start	End	Ion
350.1	333.1	305.1	192.2	10	32	[Ampicillin+H]$^+$
372.0	196.2	182.2	-	10	32	[Ampicillin+Na]$^+$
366.0	349.1	208.1	114.3	10	32	[Amoxicillin+H]$^+$
458.0	284.1	299.1	-	10	32	[Cloxacillin+Na]$^+$
403.0	244.1	227.2	285.1	10	32	[Methicillin+Na]$^+$
335.1	289.1	176.2	-	10	32	[Penicillin G+H]$^+$
357.1	313.1	198.2	-	10	32	[Penicillin G+Na]$^+$
348.1	303.1	218.1	-	10	32	[Cefalexin+H]$^+$
406.0	362.0	284.0	-	10	32	[AmoxH$_2$O+Na]$^+$
390.0	268.0	346.0	-	10	32	[AmpH$_2$O+Na]$^+$
476.0	432.0	354.0	-	10	32	[CloxH$_2$O+Na]$^+$
421.0	377.0	299.0	-	10	32	[MetH$_2$O+Na]$^+$
375.0	253.0	331.0	-	10	32	[PenGH$_2$O+Na]$^+$
388.0	266.0	344.0	-	10	32	[CefaH$_2$O+Na]$^+$

3 Results and Discussion

3.1 Determination of structural information from metal-cationised phospholipids

The results of an investigation into the generation and CID of [L+Ba]$^{2+}$ ions of phosphatidylcholines and sphingomyelins were reported in the journal *Analytical and Bioanalytical Chemistry* under the title '*Collision-induced dissociation of doubly-charged barium-cationized lipids generated from liquid samples by atmospheric pressure matrix-assisted laser desorption/ionization provides structurally diagnostic product ions*'.[112] This section includes some additional results, as well as the results published in the aforementioned article.

3.1.1 Generation of metal-cationised doubly-charged phospholipids

Lipids are typically detected as singly charged ions in MS. This may be in the protonated or deprotonated forms, or as adducts formed with ions such as Na$^+$, NH$_4^+$ or Cl$^-$ depending on the chosen ion mode and specific lipid species. More unusually, doubly charged lipid-metal adducts in the form [L+2M(I)]$^{2+}$ and [L+M(II)]$^{2+}$ may be produced.[48, 55, 112] For characterising lipid structure, multiply charged lipids can be beneficial for detecting structurally informative fragment ions by MS/MS. It was discovered that the Synapt's AP-MALDI ion source can generate phospholipid adduct ions with doubly charged metal cations, and further investigation into MS/MS capabilities was conducted. Figure 26 shows the signal intensity for analyte ions detected from DHB-based (LSM **A**) liquid AP-MALDI samples of PC (16:0/18:0) (1 pmol) containing alkaline earth metal chloride salts (500 pmol). 5 droplets were analysed for one minute per group. The ion species [L+M]$^{2+}$ was notable, since it contained only a single lipid molecule, but was doubly charged. Barium chloride-doped samples produced the most intense [L+M]$^{2+}$ ions, even when considering divalent transition metals, so was prioritised for further investigation. This is likely related to the ionic radii of the metal ions and how the lipid incorporates the ion into the gas phase structure. Shown in Table 4, the radius of the alkali earth metal divalent ions increases down the group.[231] This corresponds to the trend seen in Figure 26 where [L+M]$^{2+}$ ion intensity increases with effective ionic radius. Trivalent metals ions such as La^{3+} were not found to produce multiply charged lipid adduct ions.

Table 4: The effective ionic radii of divalent metal cations used for CID MS/MS.[231]

Ion	Charge	Effective ionic radius (pm)
Mg	2+	72
Ca	2+	100
Sr	2+	118
Ba	2+	135
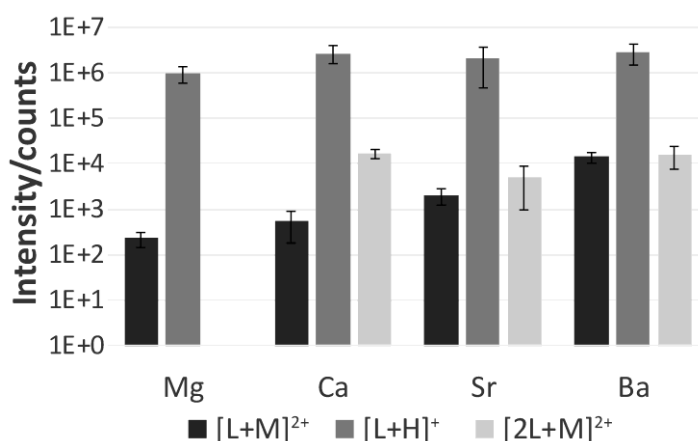
Figure 26: Signal intensity for ions from samples of PC (16:0/18:0) (1 pmol) doped with alkaline earth metals (500 pmol) for 50 s of analysis time. Error bars indicate the 95% confidence interval for 5 individual droplets prepared from the same stock solutions.

Figure 27 depicts the results of optimisation of the amount of MCl_2 (M=Ba, Sr) in liquid MALDI samples of PC (16:0/18:0) (1 pmol) with respect to obtaining the most intense lipid ion signals. An amount between 0.5 – 5 nmol of metal salt per droplet was determined to be optimal after assessment with 5 sample droplets per group. Both metal ions investigated produced similar trends in the production of the three ions, but absolute signal intensity was lower for Sr.. Interestingly, addition of too much MCl_2 was detrimental to overall analyte ion intensity. This could be due to ion signal suppression caused by the abundance of ions such as Cl^-.

$[L+Ba]^{2+}$ was detected for PCs (16:0/18:0), (16:0/18:1) and (16:0/18:2) and SM (d18:1/12:0) (Figure 28 a-c and d) but could not be found for PE (16:0/16:0). It was suggested that although the PE amine group would be protonated under the matrix conditions, the proton itself was mobile and could leave should a barium ion coordinate to the phosphate group. This could not occur for PC and SM since the head group positive charge was fixed at the trimethylamine group. The formation of $[PE+H]^+$ appeared to be favoured over $[L-H+Ba]^+$, however. The presence of isobaric PC and PE could be determined, in theory, with this observation. At this point, this has not been shown since additional PE standards were not purchased.

(a)

(b)
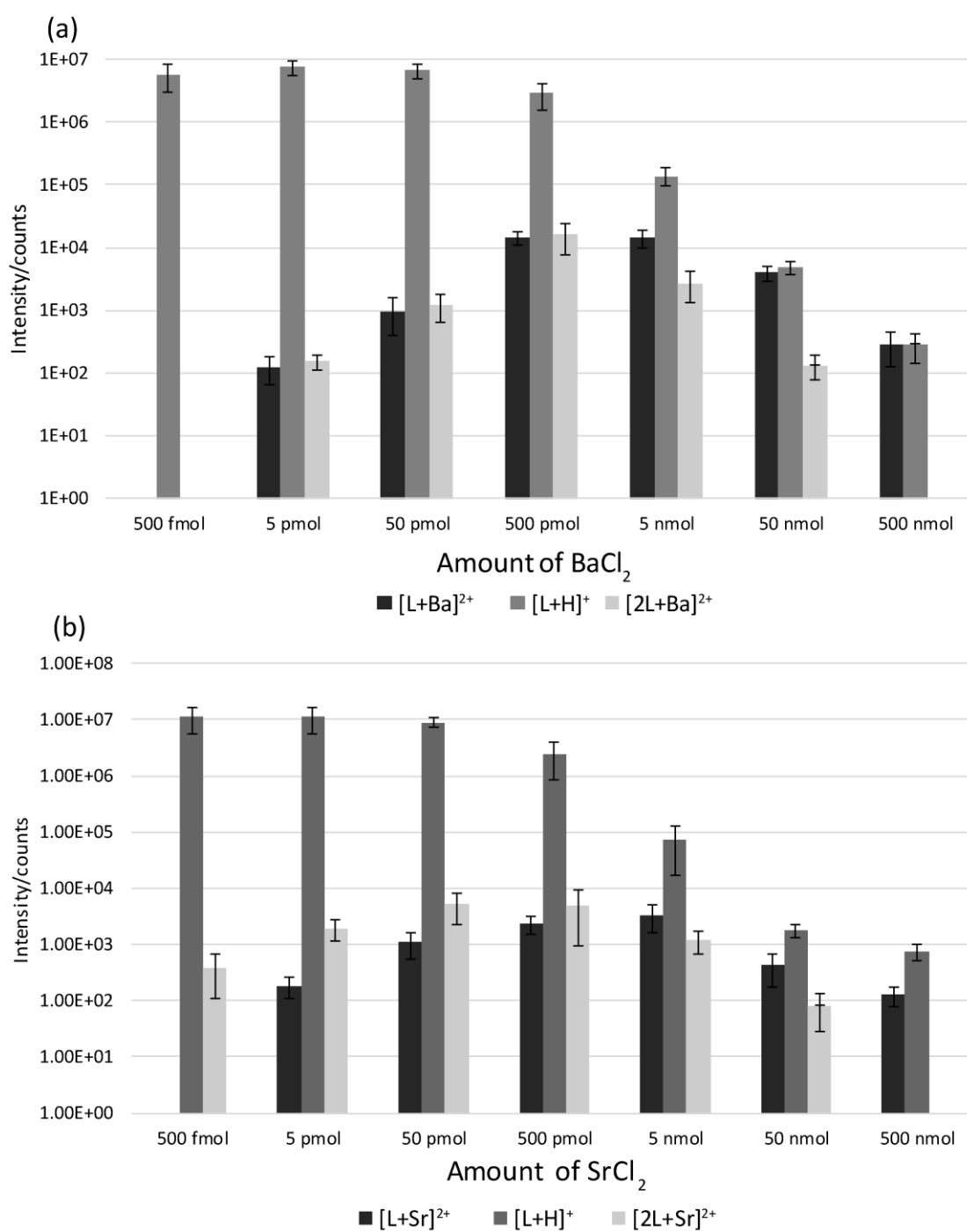
Figure 27: Ion signal intensity of three ion species from an AP-MALDI sample droplet containing 1 pmol PC (16:0/18:0) and (a) Ba^{2+} or (b) Sr^{2+} ions (500 fmol-500 nmol). Error bars indicate the 95% confidence interval for 5 sample droplets from the same stock solutions.
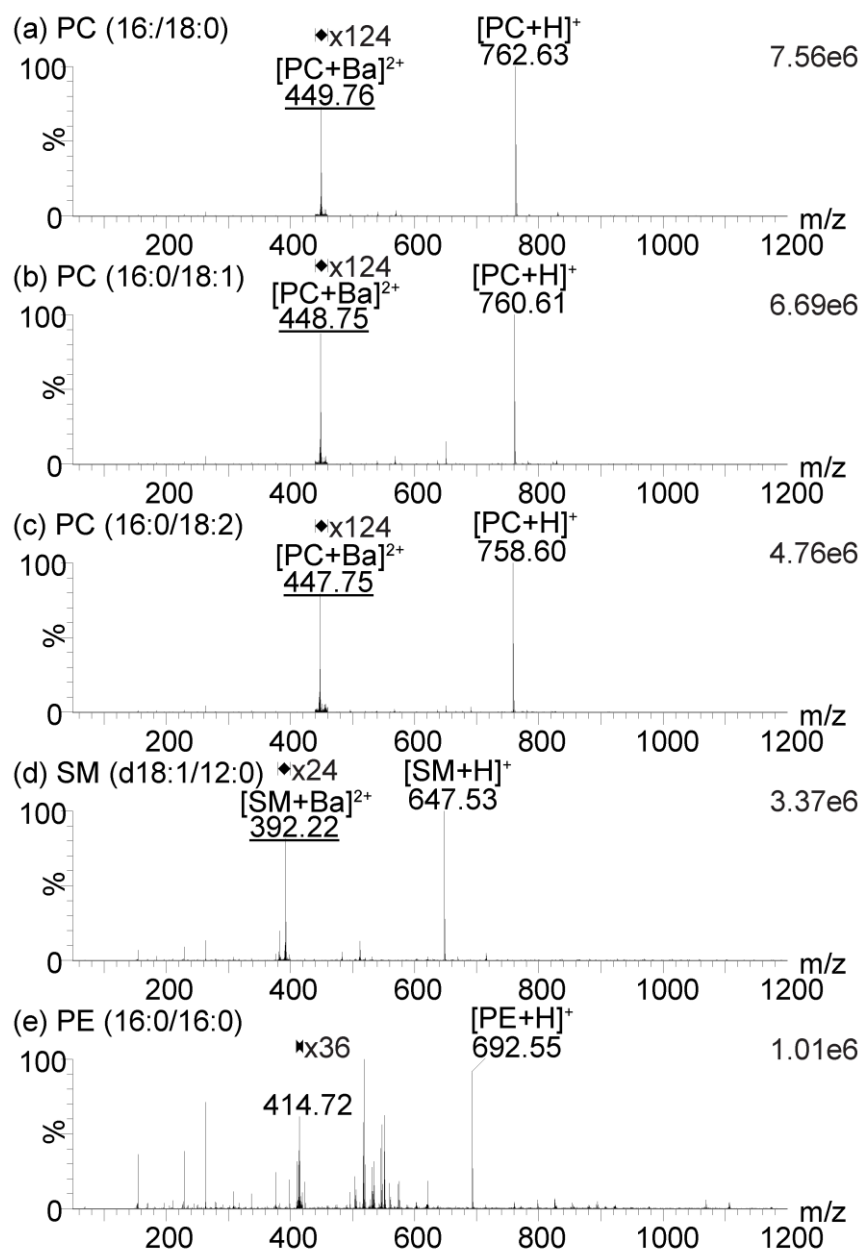
Figure 28: Liquid AP-MALDI-TOF MS spectra exhibiting [L+Ba]$^{2+}$ (underlined labels) and [L+H]$^+$ of phosphatidylcholines (a) PC (16:0/18:0), (b) PC (16:0/18:1), (c) PC (16:0/18:2) and sphingomyelin (d) SM (d18:1/12:0). CID of the peak m/z 414.72 in the spectrum for (e) did not confirm that this peak represents [PE+Ba]$^{2+}$.

3.1.2   CID

In Figure 29, CID product ions of PC and SM were classified into two major groups: intact fatty acids with the formula [*sn*x+Ba-H]$^+$ and protonated fragments resulting from the loss of a fatty acid [L+H-*sn*x]$^+$ (with x = 1 or 2). The former are highly abundant, a distinction from the CID product ions of singly charged lipid-metal adducts where the fatty acids are detected as part of [L+H-*sn*x]$^+$ and inferred from neutral loss. Further MS/MS analysis on individual [*sn*x+Ba-H]$^+$ is possible. Detection of [*sn*x+Ba-H]$^+$ suggests a putative alternative attachment of the divalent metal cation closer to the acyl chains or a relocation of Ba^{2+} during CID. The ratios of [*sn*1+Ba-H]$^+$ /

[*sn*2+Ba-H]$^+$ within the same spectra suggest a mechanistic preference for product ion formation dependent on the *sn*1/*sn*2 position of the fatty acids. The dominant product ions were recorded for the fatty acids at the *sn*2 site, which is in agreement with previous research that indicates the *sn*2 bond as being more labile in CID of [L+H]$^+$ and [L+Li]$^+$.[232] CID of SM (d18:1/12:0) revealed a low abundance peak for the *sn*2 chain (*m/z* 336.1); other peaks were all identified as being headgroup-related, although with some additional peaks to those usually detected with CID of protonated sphingomyelin ions.

A comparison of CID MS/MS product ions of [L+M]$^{2+}$ is shown in the spectra depicted in Figure 30. Mn is included as an example of transition metals performing similarly to the Group 2 elements. The trend for producing metal-cationised fatty acid CID product ions is evident for all.
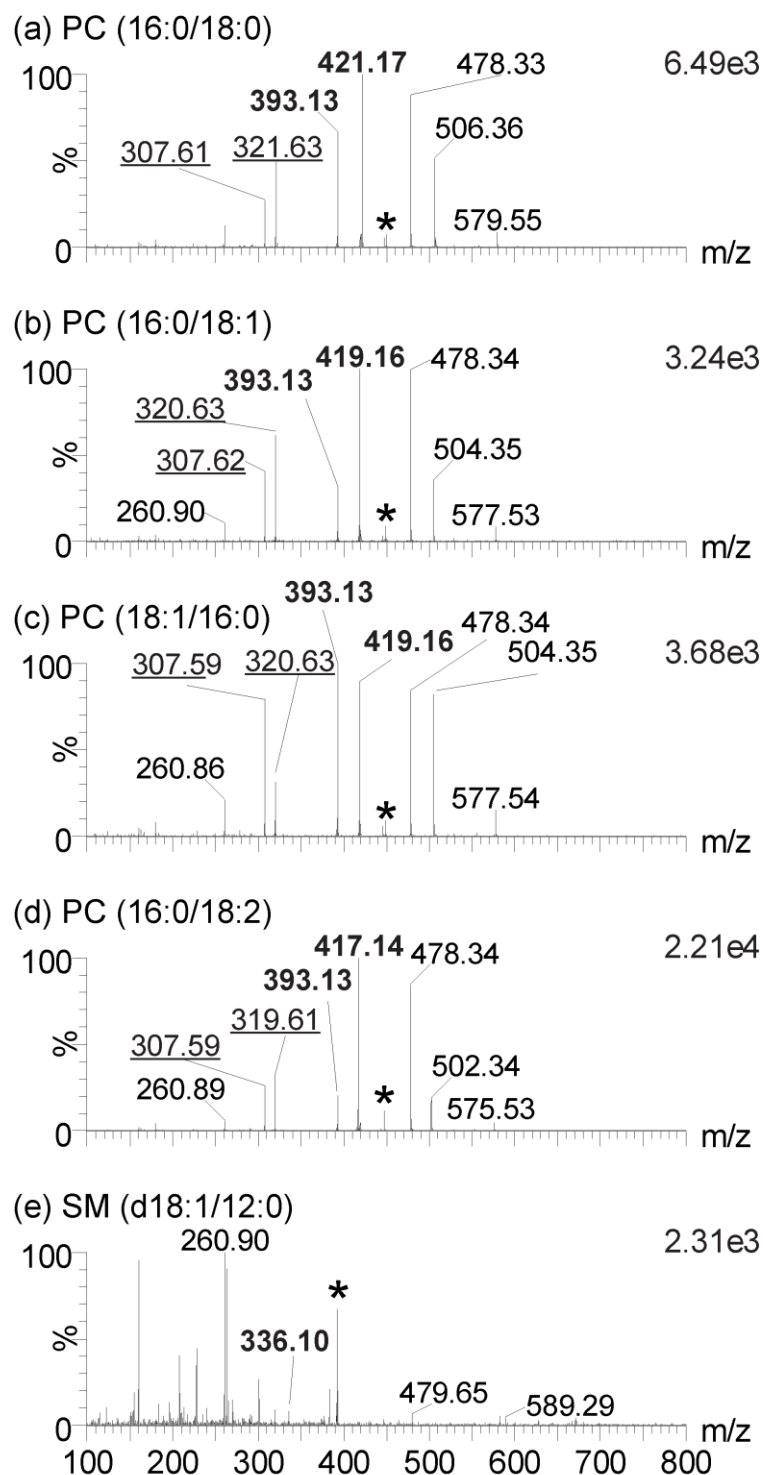
*Figure 29: Liquid AP-MALDI-CID MS/MS spectra of the precursor ions [L+Ba]$^{2+}$, indicated by *.. (a) PC (16:0/18:0), (b) PC (16:0/18:1), (c) PC (18:1/16:0), (d) PC (16:0/18:2), and (e) SM (d18:1/12:0). For the PCs, the dominant product ions are [sn1+Ba–H]$^+$ and [sn2+Ba–H]$^+$ (**bold** label), and [L+H–sn1]$^+$ and [L+H–sn2]$^+$ with other products including [L+H-sn1-N(CH$_3$)$_3$]$^+$ and [L+H-sn2-N(CH$_3$)$_3$]$^+$ (regular label), and [L+Ba–H-sn1]$^{2+}$ and [L+Ba–H–sn2]$^{2+}$ (underlined label). For the SM, many head group-related product ion peaks and a minor peak at m/z 336.10, presumably [amide+Ba–H]$^+$ from the sn2 position, were detected.*
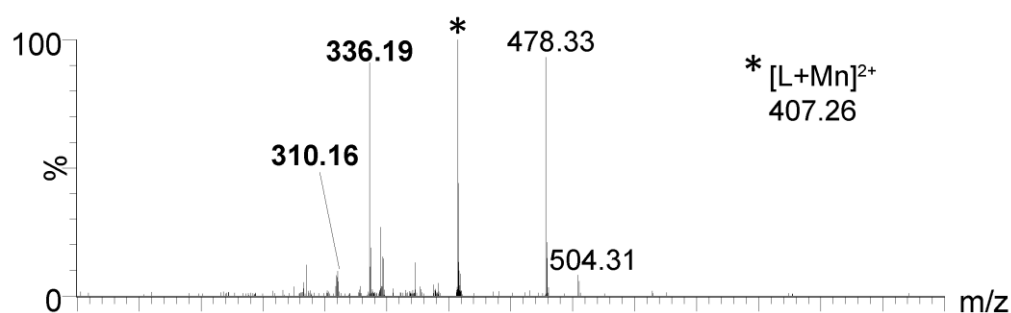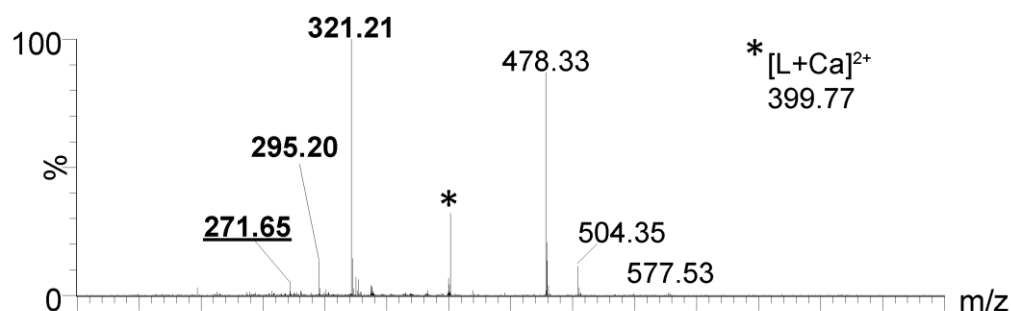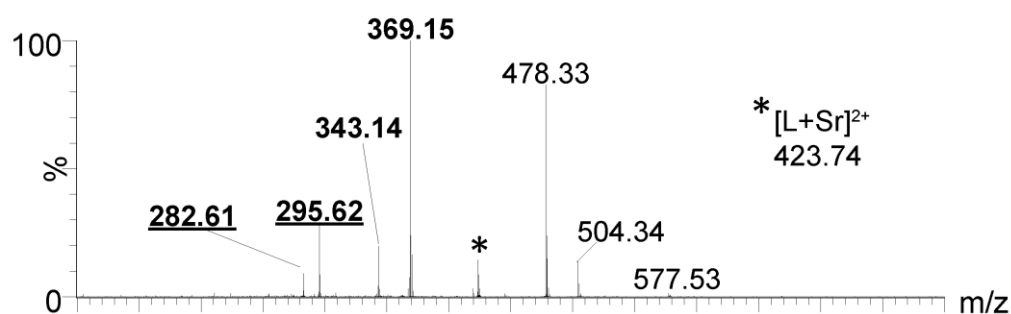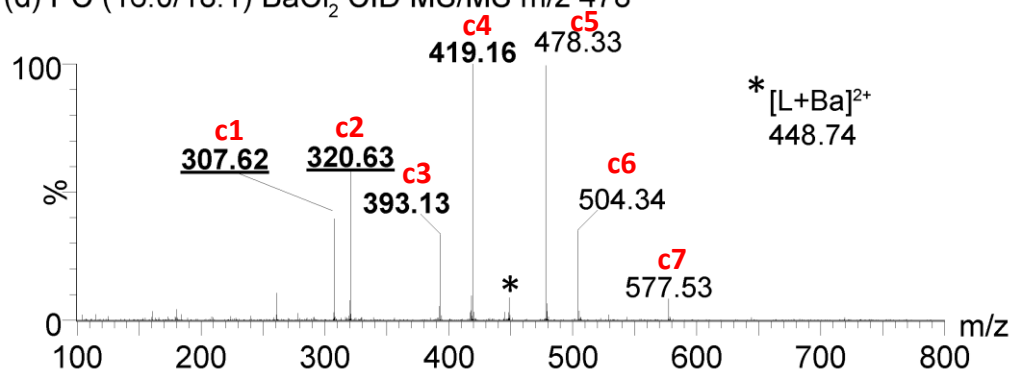
*Figure 30: Liquid AP-MALDI-CID MS/MS spectra of [L+M]²⁺ where M is (a) Mn (b) Ca (c) Sr and (d) Ba. Peak labels in **bold** describe ions that contain the metal and <u>underlined</u> labels indicate doubly charged product ions. Metal-cationised fatty acids are **bold** only. * indicates the precursor ion. The formation of fragments described with 'cn' labels are shown in **Error! Reference source not found.**.*
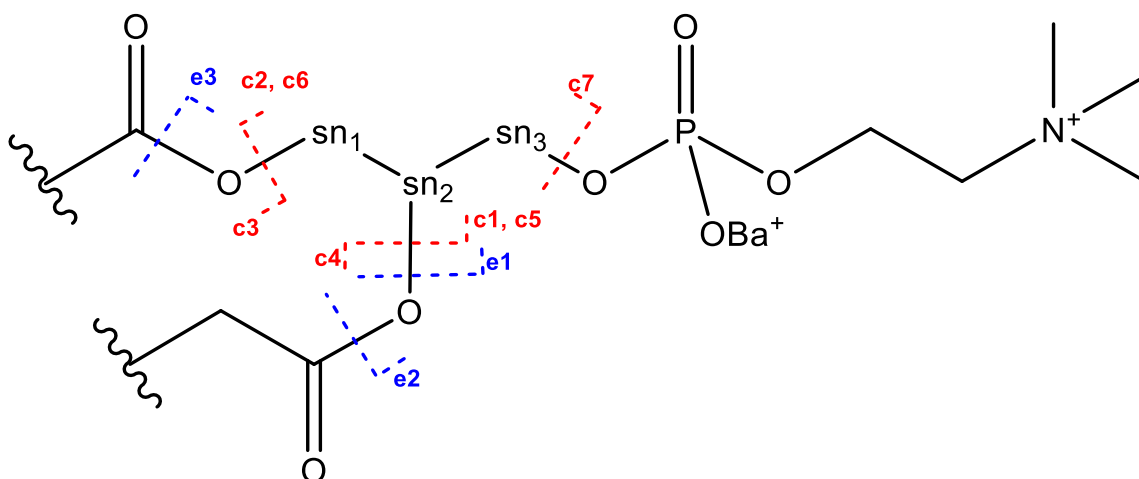
Figure 31: Bonds determined to be fragmented by CID (red) and ETD (blue) of [PC+Ba]$^{2+}$ from experimental data. The 'sn' positions, CID and ETD product ions (discussed later) are indicated.

Pseudo-MS^3 performed on the [snx+Ba-H]$^+$ fragment ion by in-source decay (ISD) and CID produced product ion spectra that indicated the position of double bonds (Figure 32). Selection of these ions from a drift plot in a TAP-MS experiment is shown in Figure 33. Charge-remote fragmentation of the acyl chain resulted in peaks every *m/z* 14 (CH$_2$), except where a double bond was present. The double bond resulted in a diagnostic peak separation of *m/z* 40 (H$_2$C-CH=CH). Formation of these fragments is proposed to occur by the mechanisms in Figure 34. Abundant ion signals of presumably Ba$^{+\bullet}$ (*m/z* 137.9) and BaOH$^+$ (*m/z* 154.9) were also detected. In the future,

enhancement of the [*sn*x+Ba-H]⁺ CID product ion signal might be possible by reducing the formation of these two ions, i.e. retaining Ba on the acyl chain.
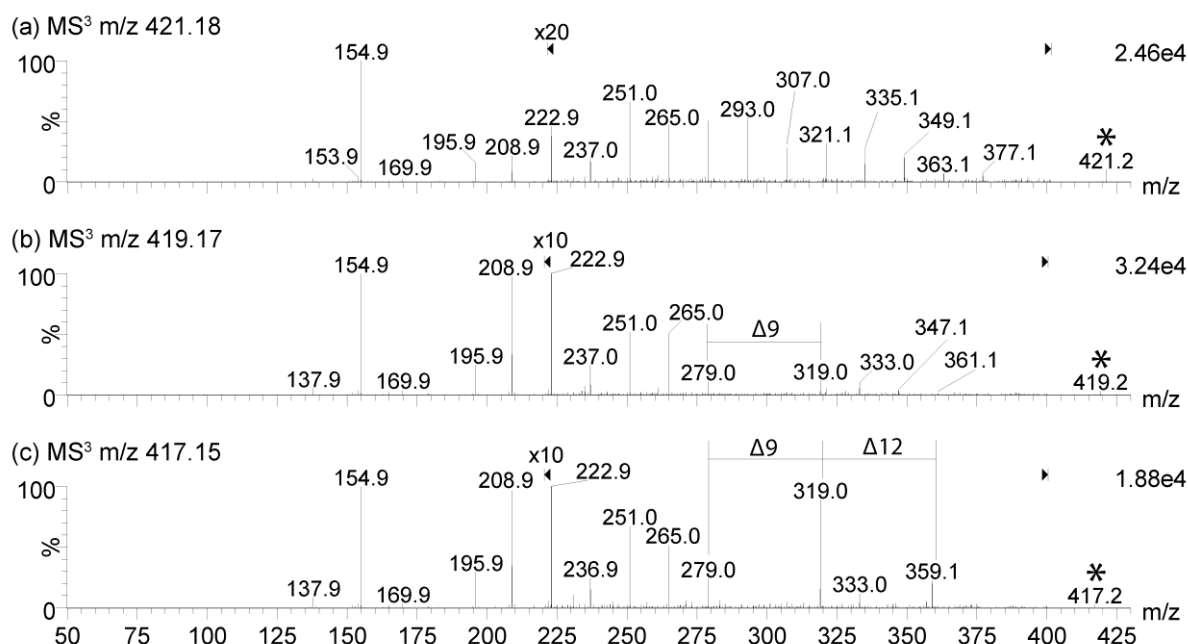
Figure 32: Liquid AP-MALDI-CID MS/MS (pseudo-MS³) spectra of ISD product ions of the precursor [L+Ba]²⁺ (*) for (a) [PC (16:0/18:0) +Ba]²⁺ ➔ [C18:0+Ba–H]⁺ ➔ products, (b) [PC (16:0/18:1) +Ba]²⁺ ➔ [C18:1+Ba–H]⁺ ➔ products, and (c) [PC (16:0/18:2)+Ba]²⁺ ➔ [C18:2+Ba-H]⁺ ➔ products.
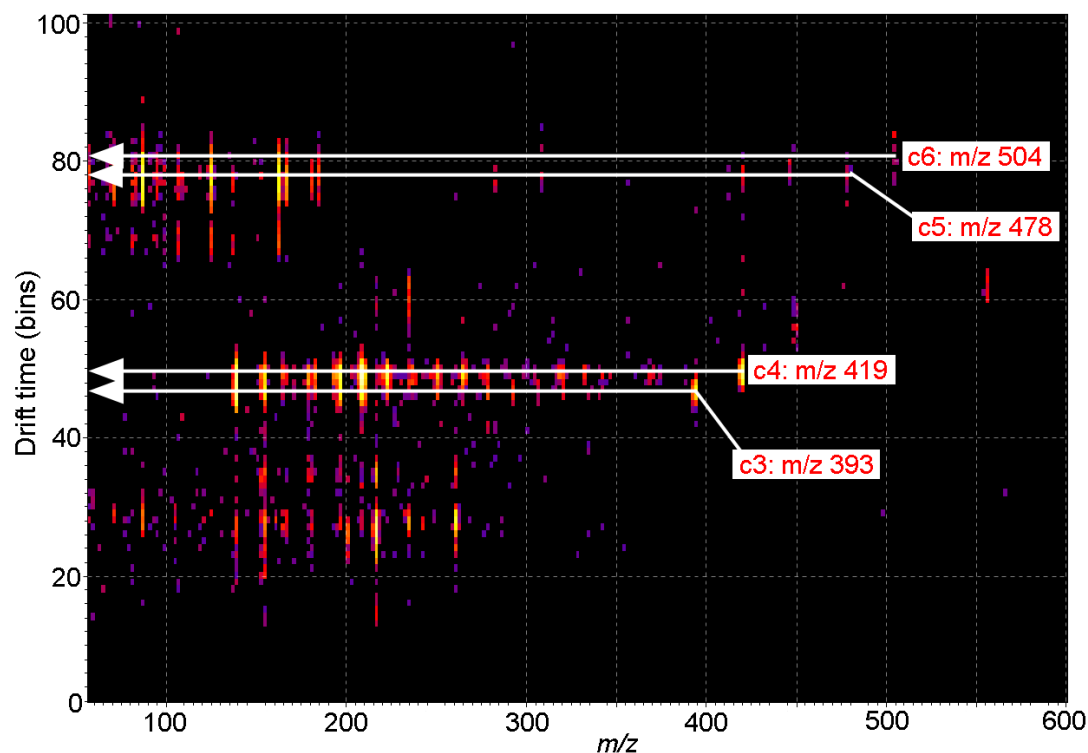
Figure 33: Drift plot from the TAP analysis of [PC (16:0/18:1)+Ba]²⁺ showing time-aligned MS³ fragments of MS² ions depicted in Figure 30d.
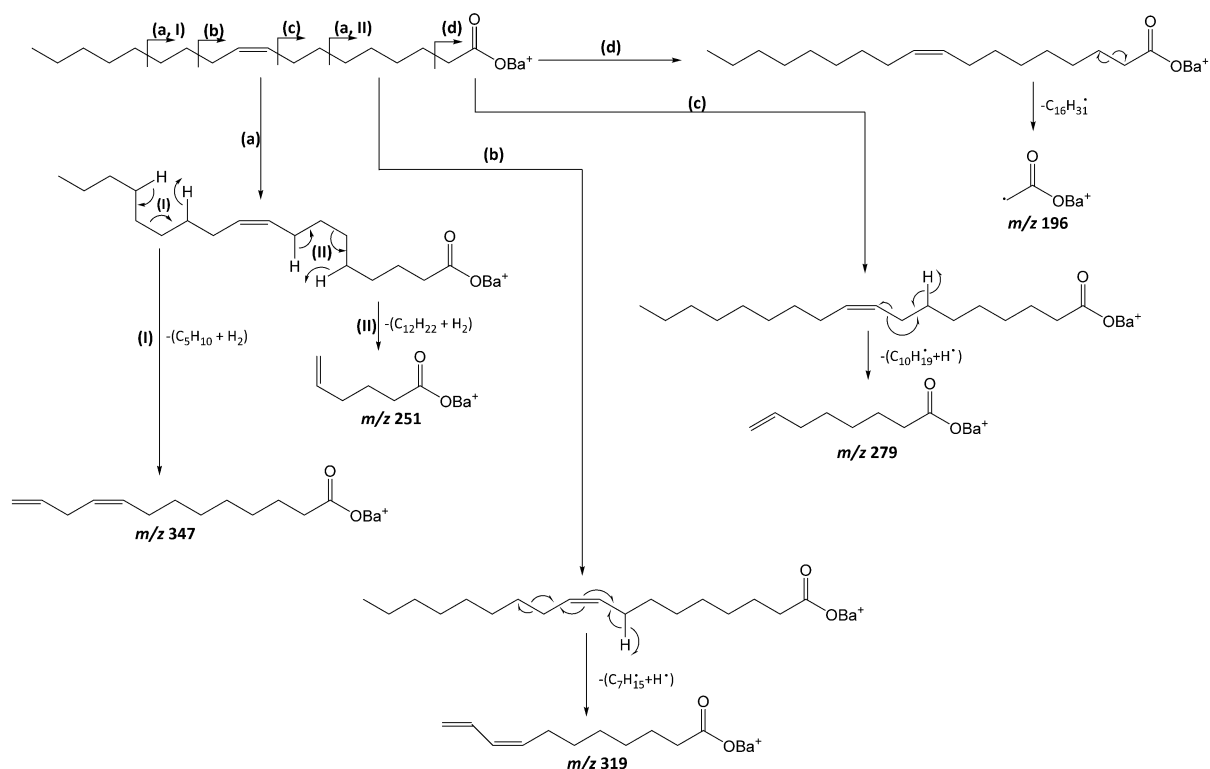
Figure 34: Fragmentation mechanisms proposed for the formation of fragments observed from [PC (16:0/18:1)+Ba]$^{2+}$

The spectra in Figure 35b and c show the recorded product ion spectra of a lipid found in soya milk for the first and second stages of CID, respectively. From these CID spectra and the MS precursor ion spectrum (Figure 35a), it was determined that the peak at *m/z* 459.8 can be assigned to PCs, of which the majority contain isomers of the fatty acid C18:2 (*m/z* 417.1, [C18:2+Ba-H]$^+$). This is further supported by the product ion peak at *m/z* 502.34 which corresponds to the loss of [C18:2+Ba-H]$^+$. At both *sn* positions the acyl chains must be C18:2 since *m/z* 782.6 is the corresponding [L+H]$^+$ ion. The TAP product ion peaks of [C18:2+Ba-H]$^+$ shown in Figure 35c reveal that the majority of the C18:2 isomers are C18:2(Δ9, Δ12). There is some evidence to suggest C18:2(Δ10, Δ12) is also present (*m/z* 293.01), another isomer of linoleic acid. Linoleic acid is the major fatty acid in soya milk and as such this was not unexpected.[233]
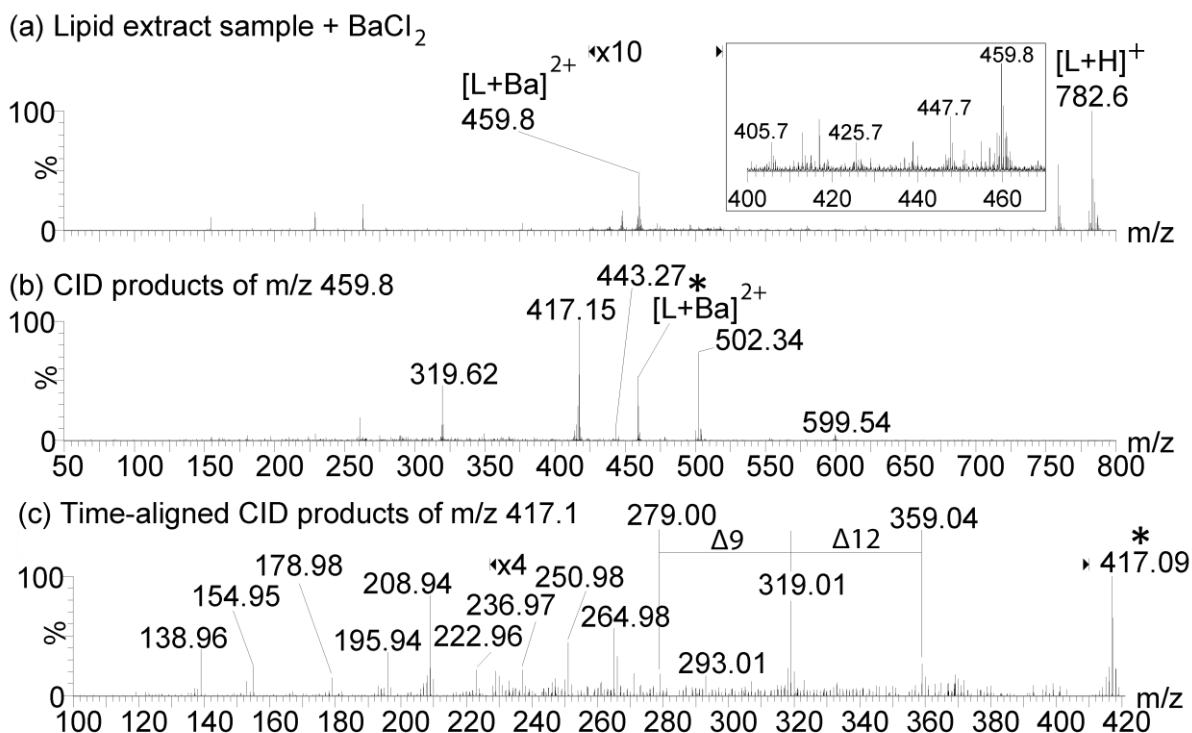
Figure 35: Liquid AP-MALDI TAP pseudo-MS³ analysis of lipids from soya milk extract. The profile spectrum (a) featured several doubly-charged peaks (labelled in inset). CID fragments of m/z 459.8²⁺(b) suggest both acyl chains were the same length. TAP MS³ of m/z 417.1 suggests the presence of double bonds at the △9 and △12 positions. Precursor ions are indicated by *.

3.1.3 ETD

The following work is of interest because not only is the literature on ETD of lipid ions limited, but ETD performed on MALDI-generated ions is not common either. Barium-cationised PC ions [L+Ba]$^{2+}$ produced by liquid AP-MALDI were successfully fragmented by ETD (Figure 36). The product ion spectra were comparable to the spectra obtained from additional experiments performed with an ESI source (data not shown). The charge-reduced product ion [L+Ba+e]$^{+•}$ and ETD reagent adduct ion [L+Ba+1,3-dicyanobenzene]$^{+•}$ were detected, showing the ion-ion reaction was taking place. Structurally important ETD fragments denoted **e1** – **e3**, are indicative of the fatty acid side chain identities. For each of the lipids in Figure 36, the **e1** and **e2** ions, with *m/z* 617.2 and *m/z* 632.2 respectively, correspond to the loss of the C18:n fatty acid. The **e3** ions correspond to the alternative; loss of the C16:0 side chain. Interestingly, the mechanism for **e2** and **e3** results in different bonds being cleaved compared to CID (**Error! Reference source not found.**). Fragments **e2** and **e3** are the result of ester bond cleavage, as previously reported for [L+2Na]$^{2+}$ ions.[48] Unlike CID, fragments corresponding to barium-cationised fatty acids [*sn*x+Ba-OH]$^{+}$ were not detected, suggesting the barium ion does not relocate during the ETD process. A product ion indicative of the headgroup was not identified.

Figure 36: Liquid AP-MALDI-ETD-QTOF MS/MS spectra of barium-cationised (a) PC (16:0/18:0), (b) PC (16:0/18:1) and (c) PC (16:0/18:2).

3.2 High-throughput milk speciation

3.2.1 Initial considerations and requirements

An aim of ambient ionisation techniques is to significantly reduce, or eliminate entirely, the requirement for sample preparation. REIMS provides a good example where this is successful for direct analysis of biological samples when equipped with an iKnife electrosurgical tool; to the end user it is simply a point-and-click sampling device with much of the complexity of mass spectrometry obfuscated.[138, 234-236] However, the electrosurgical device is optimised for tissue analysis. Liquids analysis has proven more difficult and development of laser and sonic probes to adapt for these is still ongoing.[237] Conversely, our in-house developed AP-MALDI source was designed for native compatibility with liquid samples. However, preliminary liquid AP-MALDI MS investigations revealed that crude milk would require analyte extraction to produce information-rich mass spectra. Attempts to analyse crude milk directly, mixed with LSMs, produced sparsely populated spectra. Detrimentally, crude milk added to LSM also rapidly became viscous enough to behave as if it were solid. This is because the low pH of the LSM, typically around pH 2 for 25 mg/mL DHB, causes the milk to curdle. Unfolded proteins form clumps and increase the viscosity of the droplet. Figure 37 shows a typical spectrum for crude bovine skimmed milk (3-minute summed spectrum at 30 Hz laser repetition rate) when analysed with LSM **B**. Predominantly featured were sodium and potassium adduct ions of lactose (m/z 365.12, m/z 381.09), and evidence for sodiated and potassiated polysaccharides (m/z 707.23, m/z 723.21). This was recently evident in spectra of water/methanol diluted samples analysed by direct infusion ESI in the literature.[190] Other ions in the range m/z 650-900 were potentially attributable to phospholipids but were of low relative intensity considering the acquisition duration. The following sections demonstrate how simple sample preparation led to improved spectra obtained from milk samples enabling rapid analysis and classification.
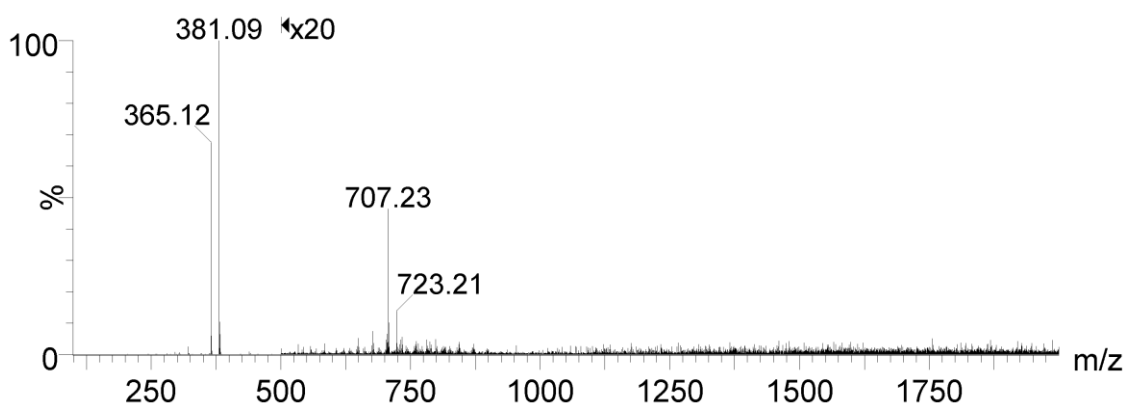
Figure 37: Positive ion mode AP-MALDI-TOF MS spectrum of crude skimmed milk from direct addition to the DHB-based LSM **B**. Sodium and potassium adducts of lactose dominate the spectrum at m/z 365.12 and m/z 381.09.

The initial development of lipid profiling by liquid AP-MALDI MS was carried out on the QTOF Ultima. Despite attempts to achieve ionisation with a variety of LSMs, only the LSM **D** was able to be used. Nevertheless, lipid profiles were obtained after HIP extraction of cow milk in the first analysis. Analysis with LSM **D** was also trialled on the QTOF Premier to provide a comparison to a commercially available, intermediate vacuum MALDI source (Figure 38).
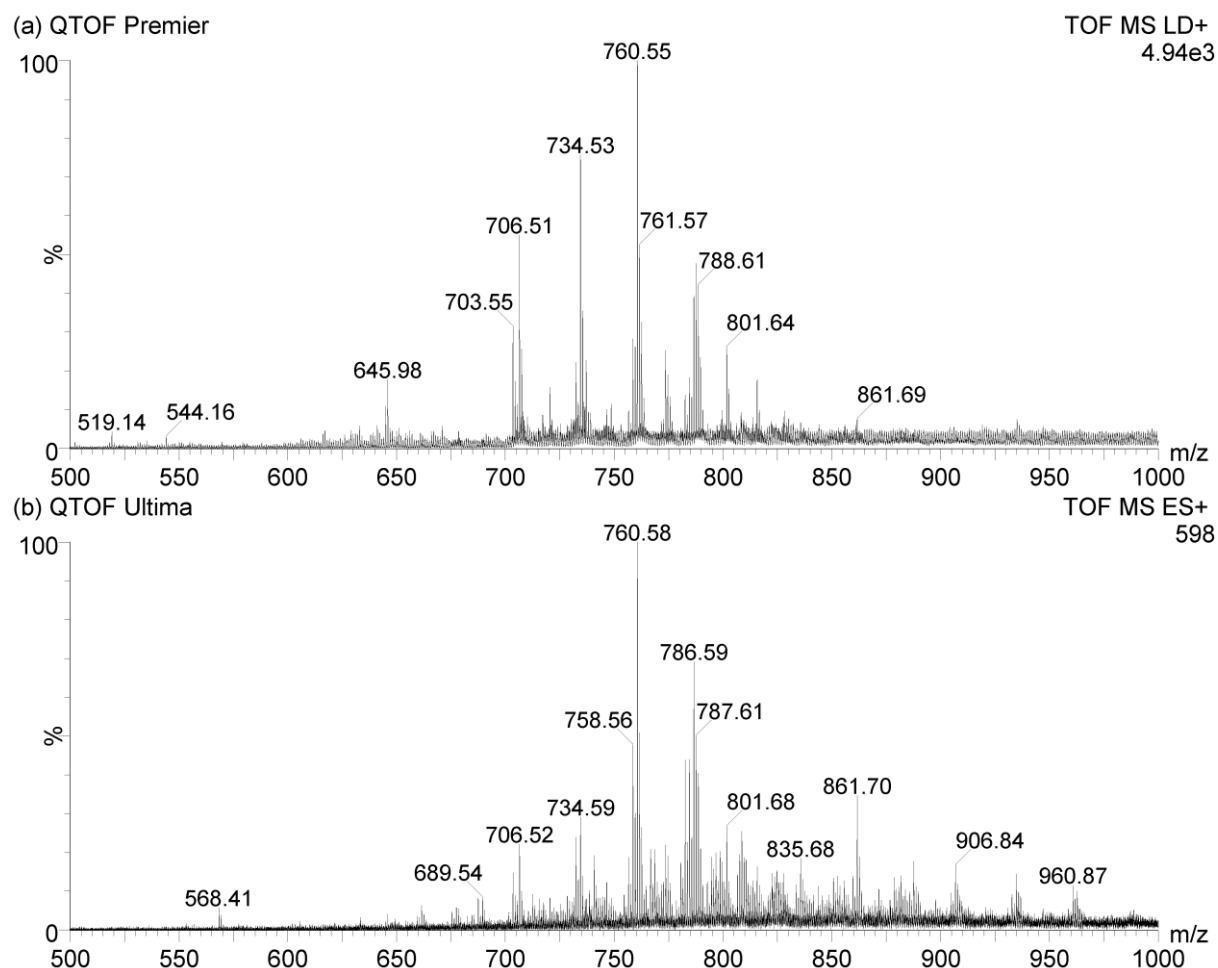
Figure 38: Cow milk lipid extract spectra acquired on the (a) MALDI-QTOF Premier and (b) AP-MALDI-QTOF Ultima MS.

Once the principle for lipid extraction and MS analysis was demonstrated, the decision was made to further develop the method with the Synapt's AP-MALDI source. The inherent sensitivity of the Synapt instrument itself, and the flexibility of the ion source would allow analysis on the order of seconds rather than minutes, as was necessary with the QTOF Ultima. The potential for automated analysis through the integration of the electronic sample stages and WREnS was also attractive.

Important for high-throughput analysis is limited carryover between samples. The Synapt AP-MALDI source was assessed with a liquid sample of leucine-enkephalin (1 pmol, LSM **A**). The source was operated with the heated ion transfer tube and counter-flow gas set to the typical settings for peptide analysis i.e. 180 L/h counter-flow gas, 200 °C. The TOF analyser was set to a

scan time of 100 ms, with an interscan delay of 100 ms. The 337 nm laser was set to a pulse

repetition rate of 1 Hz. One peak per 5 scans was observed on the extracted ion chromatogram

(Figure 39). Examination revealed that the analyte signal did not carry over to the next scan on

this timescale. This was important, as high-throughput could be achieved without overlapping

sample signals, since the sample stage took at least 0.65 s to move between samples. It was,

however, beneficial to run analysis with a longer scan time (1 s) to improve per-scan peak

intensity and compatibility for data analysis.
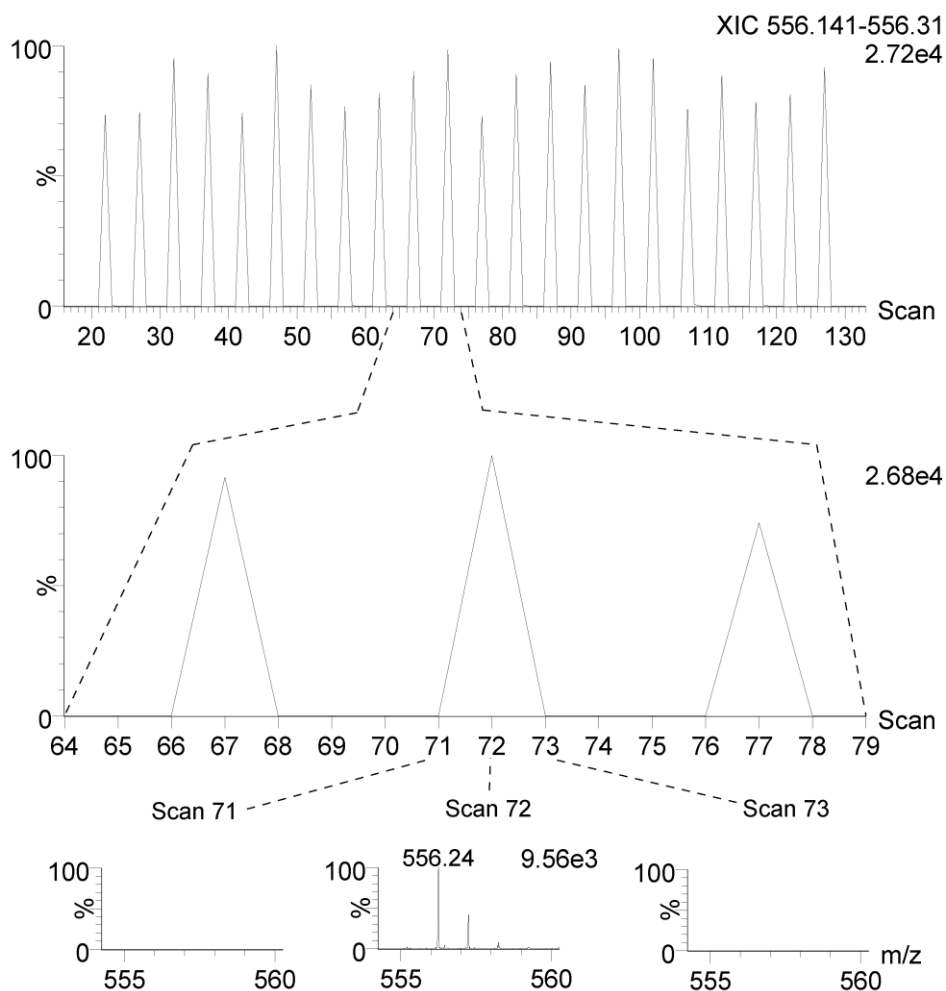
Figure 39: Chromatogram demonstrating the absence of carryover between scans for the liquid sample analysis by AP-MALDI. The leu-enk signal (m/z 556.24) is present only for the scan on which the laser was pulsed. The scan time was 100 ms, with an interscan time of 100 ms. The laser pulse rate was 1 Hz.

3.2.2   Sample classification using phospholipid models

The supernatant of the HIP extraction was analysed in both the positive and negative ion mobility-

TOF modes with the Synapt's AP-MALDI ion source. Mobility mode was found to generally

improve signal intensity for analyte ions. The proposed explanation is that the pulsed ion beam

from the AP-MALDI source was gated in the ion mobility device and hence better synchronised

with the TOF pusher cycles. Ion mobility data was also beneficial to record, enabling charge-state filtering, background subtraction and collision cross section calculation.

Each phospholipid extract sample spot was exposed to approximately 300 laser pulses over a period of 10 seconds. Visible consumption, as viewed through the source camera, of the sample was not evident after this short analysis period, allowing multiple analyses of the sample if required. Advantageously, the DHB LSM **A** provided sufficiently intense spectra in both ion modes, thus a single sub-1.5 µL sample, considering the evaporation of volatile components, provided positive and negative mode data. The source parameters were optimised for each polarity; ion transfer tube power was kept in the region of 26-30 W to promote desolvation and the MALDI sample plate voltage at 3-3.5 kV. The counter-flow gas was set to 180 L/h and 200 L/h for positive and negative ion modes, respectively, tuned on the intensity of the peak at *m/z* 760.6 in an animal milk sample.
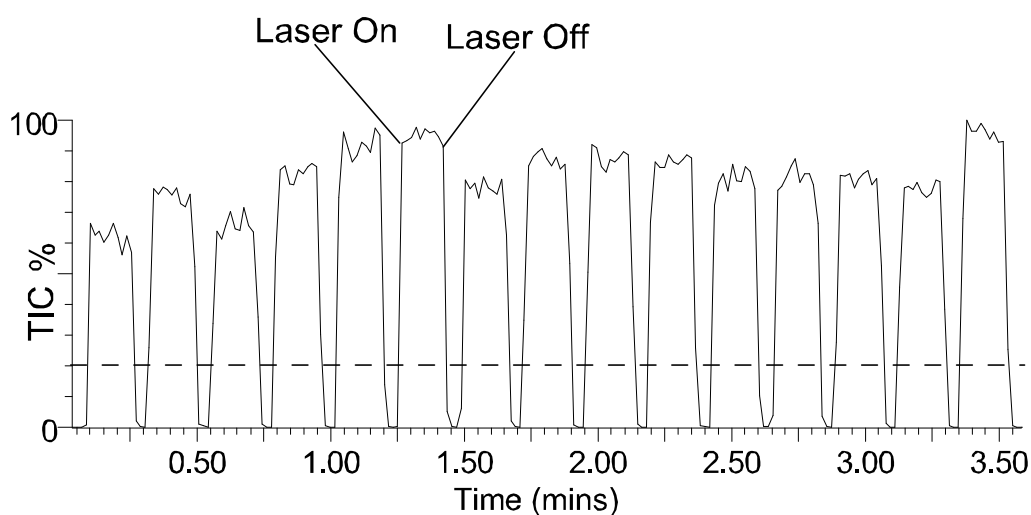
Figure 40: The distinctive total ion chromatogram produced from automatic sampling using the AP-MALDI source with a laser pulse repetition rate of 30 Hz, ten seconds allowed per sample and an inter-sample time of three seconds. The dashed line indicates %TIC used for automatic sample detection in software, based on the rise and fall of the signal.

Distinctive, 'square wave' chromatograms were produced by the automated analysis (Figure 40). It is interesting to note Figure 41, which shows an example chromatogram and mass spectrum for lipid extracts (HIP) of bovine milk analysed for one second each. This is approximately the fastest achievable analysis rate with this ion source and the standard Waters MALDI 96-well sample plate, limited by the movement speed of the XY-stage, and still produces good quality spectra.
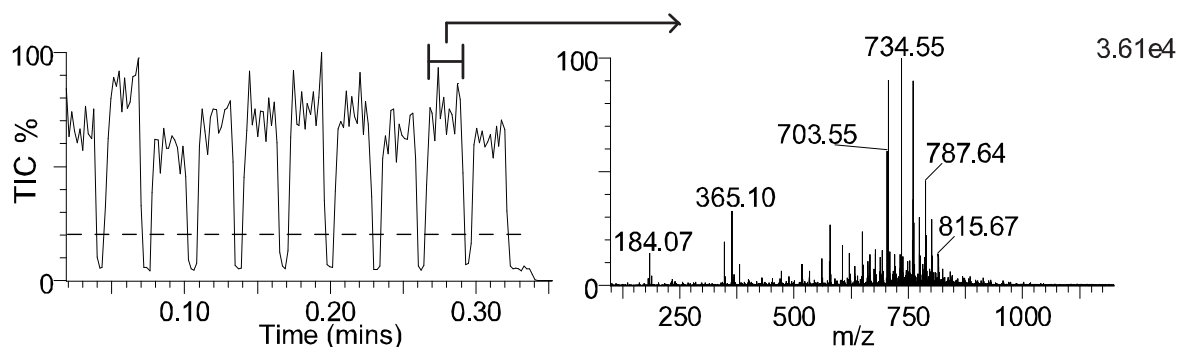
Figure 41: (Left) AP-MALDI MS chromatogram of ten lipid extracts using HIP extraction of bovine milk analysed for one second each with a 10 Hz TOF scan rate. (Right) The mass profile of the sample indicated on the chromatogram.

Richly populated positive ion mode mass spectra were produced, particularly in the region from *m/z* 600 to *m/z* 1000, appearing like those reported in the literature for standard MALDI MS of bovine, caprine and ovine milk.[178] Ions of two phospholipid classes PC and SM were abundant. TAGs were not readily detected presumably because they are typically suppressed by phospholipids without additional sample preparation.[238] As an aside, preliminary experimental data suggested that substituting water for 10 mM ammonium phosphate solution as the matrix aqueous component can result in the detection of TAG [M+NH$_4$]$^+$ alongside phospholipids, however this was not realised until late into this project.

45 samples of each of bovine milk (whole and skimmed), whole caprine milk, raw and pasteurised ovine milk and unsweetened soya milk were analysed. Individual aliquots were taken from the same stock bottle and extracted with the HIP solvent mixture. One liquid MALDI droplet was spotted for each of the 45 samples per milk type for a total of 270 spectra. Droplets that escaped the sample well during analysis and produced low intensity spectra were ignored for subsequent multivariate analysis, resulting in 266 positive ion mode spectra and 268 negative ion mode spectra. Likewise, for protein analysis the TCA-based extraction was performed but on half the number of aliquots (30 for each of bovine skimmed milk, caprine whole and ovine raw. 15 for each of bovine whole, ovine pasteurised and soya), this consideration being made due to the 6-fold increase in time required to analyse these extracts.

Example positive ion mode mass spectra for samples from the six milk classes, each analysed for ten seconds, are shown in Figure 42. The soya spectrum was distinctly different from the rest: some ions with *m/z* greater than 800 have greater relative abundance than those ions in the animal milk. The ions at *m/z* 782.57 and 820.51 are examples that are essentially absent in the animal milks. These signals are attributable to PC(36:4) and PS(39:7) by their *m/z*. The animal milks feature ions such as m/z 815.69, attributable to SM(32:1), which the soya milk lacks. Relative intensity variation in ions common to the animal milk spectra can be seen, for example, the ratio of ions at *m/z* 703.57 [SM(34:1)+H]$^+$ and *m/z* 706.53 [PC(30:0)+H]$^+$ as described

previously in the literature.[178] This was evident for all technical replicates from the same milk stock. Bovine milk featured higher intensity of the PC ion than either caprine or ovine milk. Ovine milk featured the highest relative abundance for the SM ion. This trend matches with that described in the literature, suggesting it is a characteristic of each type generally.

Negative ion mode mass spectra (see Figure 43) featured peaks for different classes of phospholipids than the positive ion mode. PS, PI and PG ions were identified by *m/z* and collision cross section. Collision cross sections were only able to be used as available in databases or published research articles. As with the positive ion mode, visual distinction between soya and animal milk is possible and the animal milks were more visually similar. Overall, TIC was approximately 1 order of magnitude lower in negative mode, with a similar decrease in base peak ion intensity.
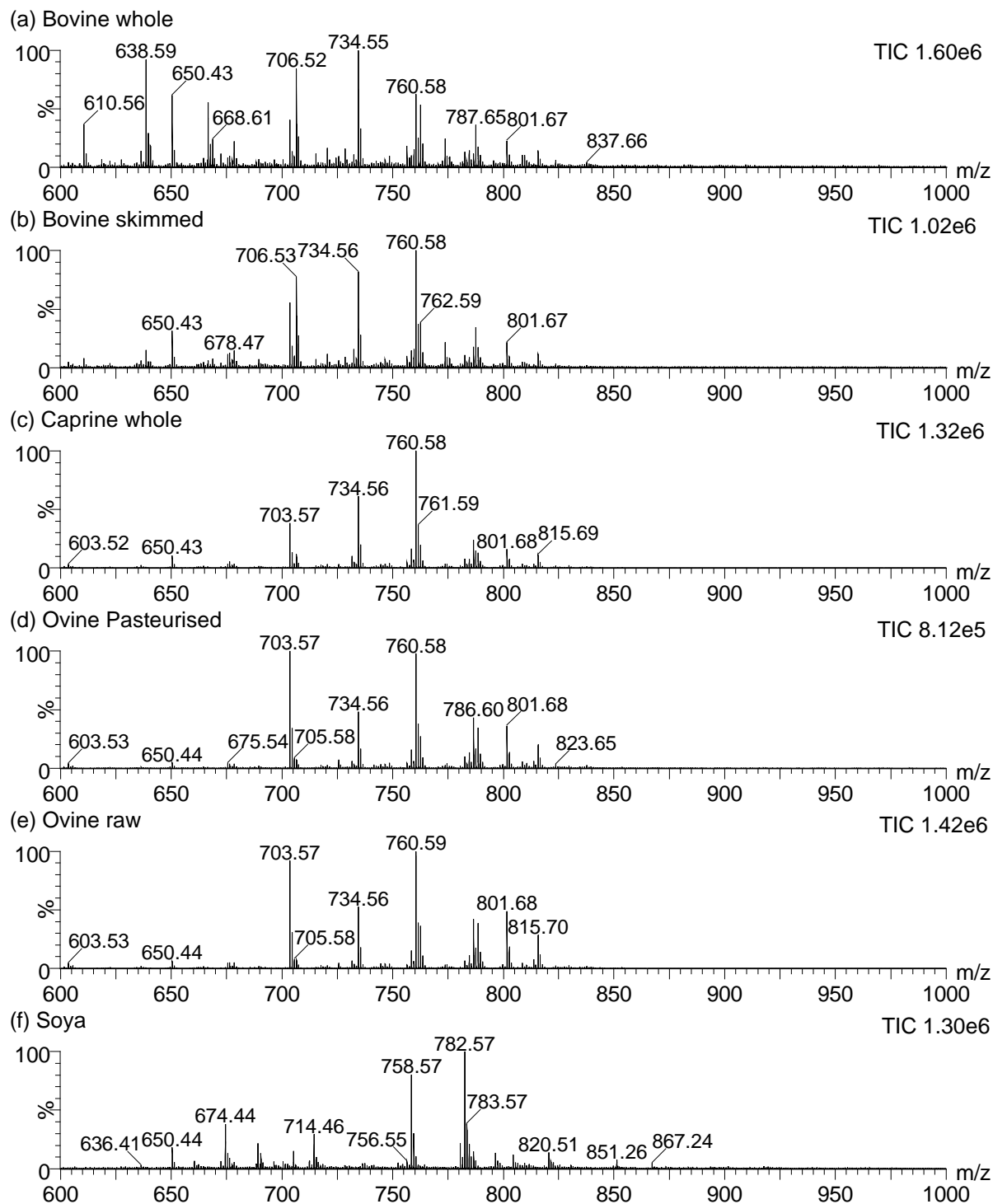
Figure 42: Example positive ion mode AP-MALDI TOF mass spectra for the six milk classes investigated (a) bovine whole milk, (b) bovine skimmed milk, (c) caprine whole milk, (d) ovine pasteurised milk, (e) ovine raw milk and (f) soya milk. The reported TIC is for a one second scan.
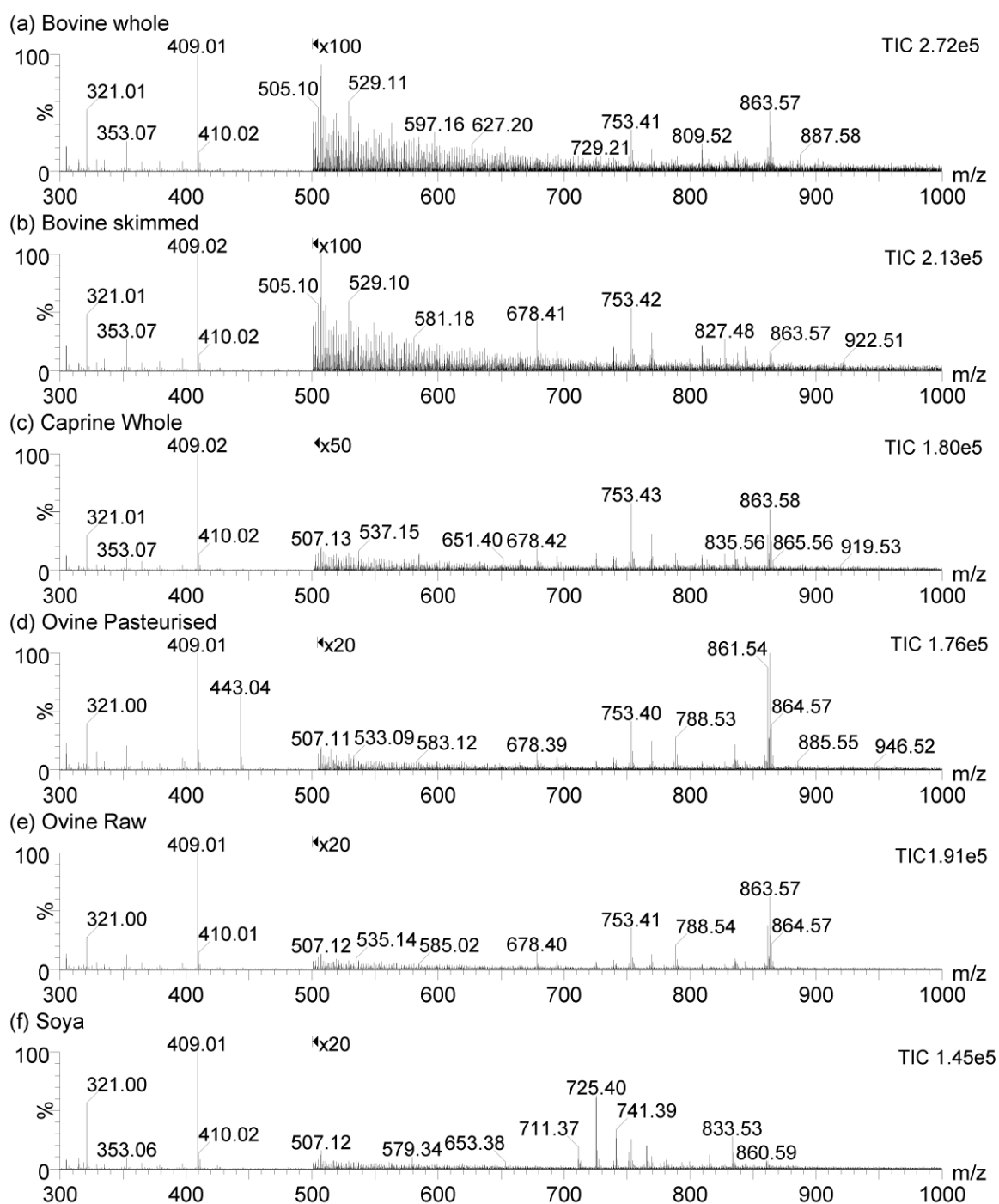
Figure 43: Examples of negative ion mode mass spectra for the six milk classes investigated (a) bovine whole milk, (b) bovine skimmed milk, (c) caprine whole milk, (d) ovine pasteurised milk, (e) ovine raw milk and (f) soya milk. The reported TIC is for a one second scan. The greater m/z range compared to the positive ion mode was used because of fewer abundant matrix-related ions and thus the potential for using lower mass metabolites for classification, e.g. m/z 321.00, 409.01.

For automatic classification of samples according to their mass spectra, multivariate statistical techniques were used within commercial software packages 'AMX' and 'Progenesis QI' with the 'EZInfo' plugin. Both programs feature mass spectrum normalisation, background subtraction and

lockmassing functions for enhancing mass spectra prior to multivariate analysis. The combination of these functions can improve the comparability of spectra with varying absolute ion signal intensities. Each uses the normalised signal intensities to then compare the variance between spectra. AMX sums signal intensity using a binning method (Figure 44) to reduce computational load for best compatibility with online classification. The bin size can be specified but it is not practical to use a bin size smaller than the peak width, and smaller bins increase model computation time significantly. Progenesis QI detects normalised peak intensity for each individually resolved peak.
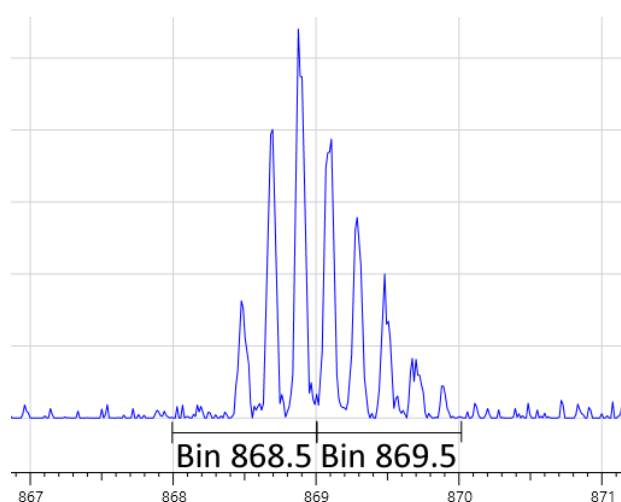
Figure 44: The spectral binning method for a bin size of m/z 1 as used in AMX to reduce computation load. The normalised signal intensity is calculated for each bin.

Multivariate statistical techniques such as principal component analysis (PCA), invented in 1901 by Karl Pearson, have found utility in high throughput MS analysis as they allow the variance between many mass spectra to be assessed.[239, 240] A computer model can be trained to characterise samples from mass spectrum features. PCA is an unsupervised method, i.e. the correct sample classification is unknown to the model. Thus, it is a way to determine if variable features between sample spectra actually exist.[241] Linear discriminant analysis (LDA) and partial least squares discriminant analysis (PLS-DA) are related techniques suited for automatic classification but differ from PCA in that they actively build the model with the sample classes predefined.[242] It is important that PCA is performed prior to discriminant analyses, as this helps prevent overfitting of the data, a problem that can cause a model to fail validation tests. A variety of multivariate analysis methods can be found in the research articles in (Table 1) and are commonplace in modern research. To consider the subtle, but significant, differences between lipid profiles, PCA models were constructed. The parameters for these models are shown in Table 5. Following separation of samples by class in a 10-dimensional PCA calculation, LDA was

performed with three dimensions to build a model suitable for classifying unknown samples. The positive and negative ion mode LDA projections are shown in Figure 46. Classification by each species is most visually distinct for the positive ion mode.

Table 5: PCA/LDA model parameters for lipid analysis by liquid AP-MALDI MS profiling

Model Property	Value
PCA dimensions	10
LDA dimensions	3
Intensity limit (arb. unit)	0
m/z range (Ion mode)	600 – 1000 (+) 300-1000 (-)
Binning (*m/z*)	1
Binning mode	Advanced
Lockmass (*m/z*)	706.5381, 758.5699 (Soya only)
Spectrum Interpretation	One per burn
Lockmass	Yes
Background subtraction	Yes
Normalisation	Yes

Results of cross validation of all LDA models discussed in this section by a "leave 20% out" method, described graphically in Figure 45, are summarised in Table 8. While the positive ion mode achieved 100% classification accuracy accounting for the four organisms, the model suffered when samples from two groups in a single species were included. Visual overlap of the caprine and bovine milk types in Figure 46 suggests little difference in the negative ion mode lipid profiles of these samples. The negative ion mode model returned a classification accuracy of 92.16% for both 4- and 6-class models. The 6-class models split the bovine and ovine milks into their two subtypes. While the clustering in the negative ion mode model is less tight, perhaps because of lower and more varied analyte ion intensity, there is less overlap between closely related samples. The clustering could likely be improved through longer acquisitions improving signal to noise ratio and overall signal intensity. It should be noted that the visual projection of the data for these models is not precisely an indicator of the model's accuracy.
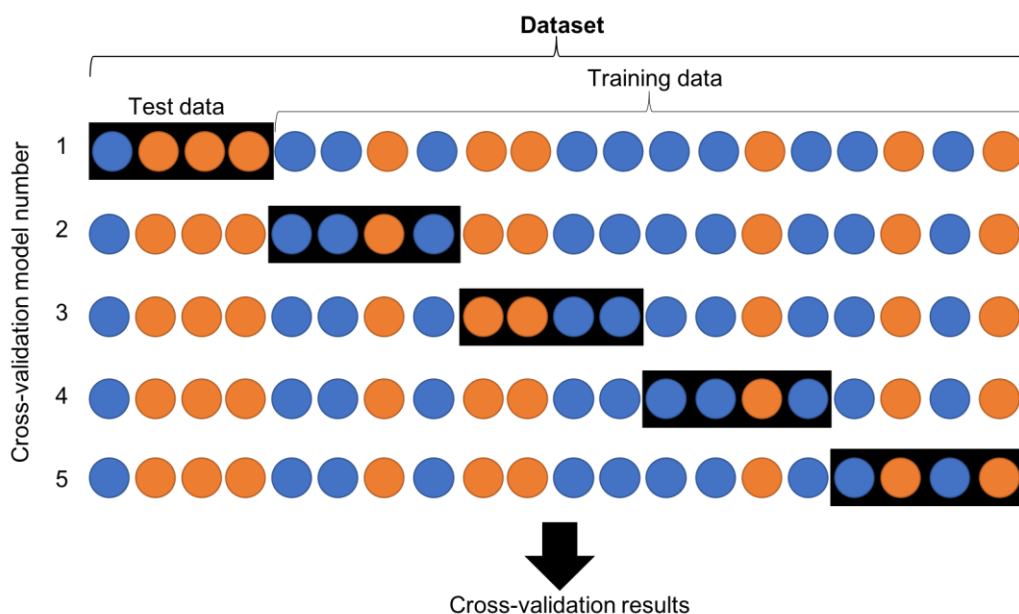
Figure 45: Graphical representation of the 'leave 20% out' cross-validation. 20% of the data is used as the test set and the remaining 80% is used to build the model. The 20% is then tested against this. The process is repeated until all data has been in a test set.
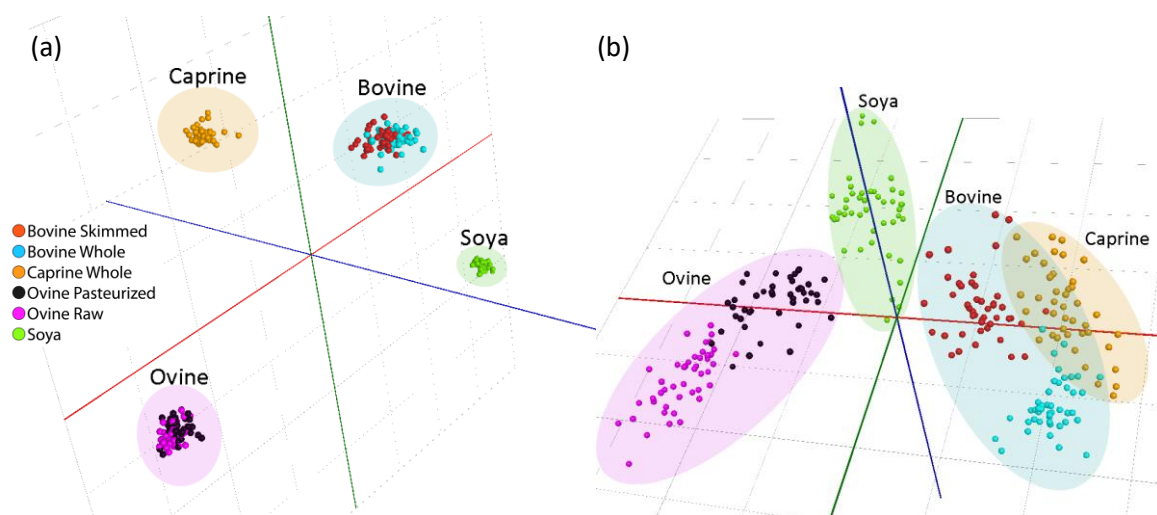
Figure 46: 3-dimensional LDA plots of six milk classes constructed from (a) positive and (b) negative ion mode MS profiles of lipid extracts; bovine whole (blue), bovine skimmed (red), caprine whole (orange), ovine raw (purple), ovine pasteurised (black) and soya (green).

3.2.3   'AMX Recognition' with lipid profiles.

The 'AMX Recognition' software was developed by Waters for real-time sample classification with ion sources like REIMS/iKnife, and AMX models. However, it is essentially compatible with any ion source, providing the data is collected in the correct fashion, i.e. with distinct increases in TIC against the background when a sample is analysed. This is necessary so that the software can detect sample mass spectra without the need for manual definition by a user. When a sample is analysed the TIC increases. If this passes through the set threshold the software detects the following mass spectra as from a sample and will attempt to classify them against an LDA model

previously trained with AMX. This makes for a streamlined screening process with real-time feedback if required.

The compatibility with data generated by the AP-MALDI source is much the same as with the core AMX model builder. To demonstrate this an example screenshot of the AMX Recognition application window is shown in Figure 47. The classification of the sample being analysed is shown in the window as the sample is analysed, or as data is read by the software in a postprocessing mode. The classification is made by comparison to a model built with training data from samples of known classification. A record of the assignment of each mass spectrum, plus a probability score, allows data to be reviewed and stored for future reference.

The 4-class PCA/LDA milk speciation model was used to classify a mixed test set of 15 HIP-extracted samples, technical replicates of the milk stocks used to build the model. The classification model was selected, the data file for the test set was added to AMX Recognition in the post processing mode and the analysis initiated. As was to be expected from the cross-validation results above, the 15 samples were classified correctly. The AMX Recognition settings are shown in Table 6.
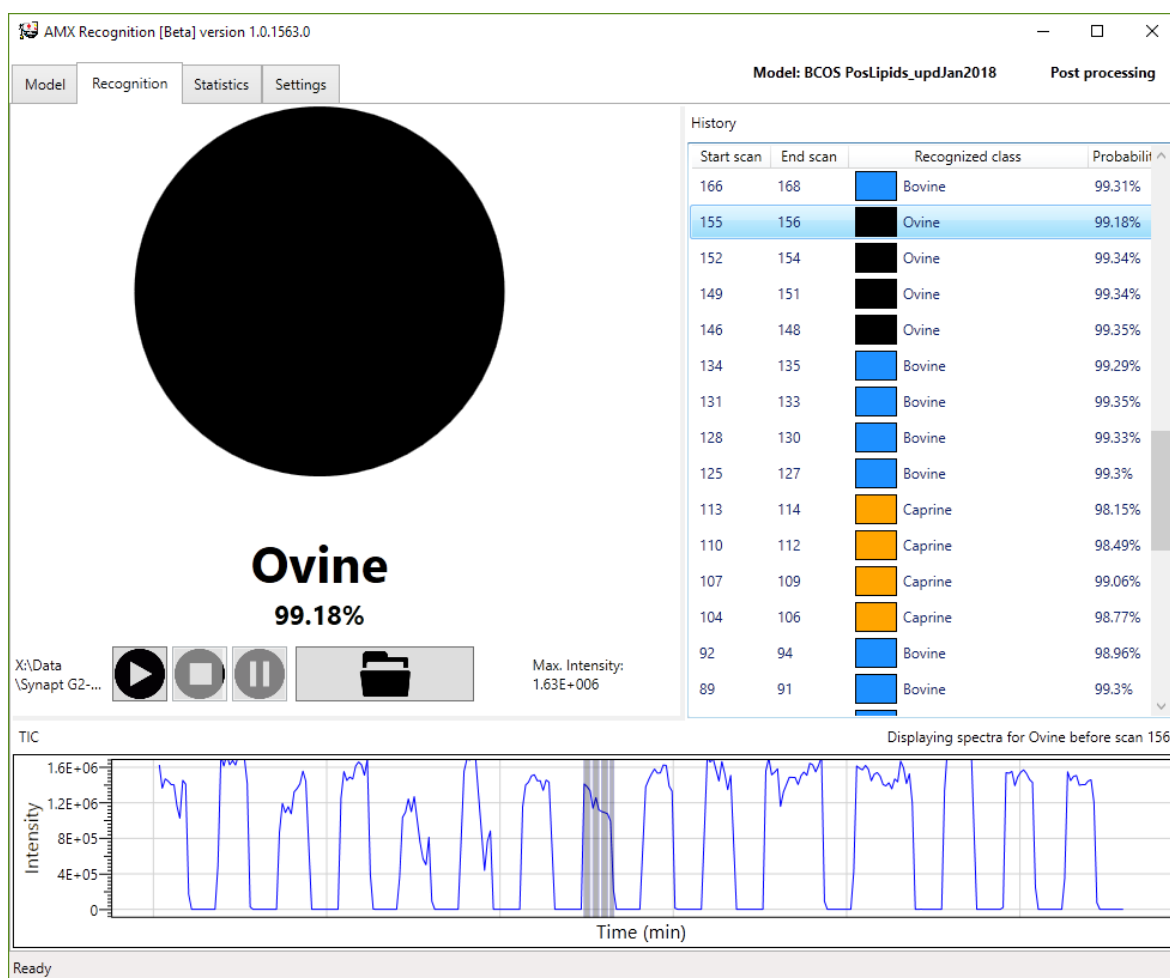
Figure 47: The application window of AMX Recognition. The left pane shows the current sample classification, the right pane shows all analysed spectra with their identification probability and the bottom pane shows the chromatogram.

Table 6: Parameters for the 4-class positive ion mode classification model in AMX Recognition.

Parameter	Value
Apply Lockmass correction	Yes
Apply background subtraction	Yes
Threshold	10000 counts
Lockmass	*m/z 758.5699 (soya) m/z 760.5851 (all others)*
Outlier type	Standard deviation
Standard deviation multiplier	5
Intensity limit	4E5 counts
Scans per spectrum	2
Good spectrum timeout	1 s
Probability threshold	Global (0%)

3.2.4   Modelling with ions of charge >1$^+$.

The custom AP-MALDI ion source was designed for promoting the generation of multiply charged ions.[108, 109] Recent publications have demonstrated the ability of this ion source for protein analysis in both ion modes, although generally negative ion mode signal is poorer in comparison to the positive ion mode.[222, 223] The use of highly charged positive ions for classification models was investigated as a way of making more specific models from underlying organism biology. Enhanced understanding of the optimum ion source conditions, along with modifications to milk sample preparation, were required for the detection of ions of charge state greater than 1 in positive ion mode experiments. No cone gas was used when acquiring the spectra, as it was found to promote sodium and potassium adducts for protein ions (see Figure 48). This could be due to the increased residence time within the heated inlet while the gas flow is on, allowing more time for sodium and potassium ions from the inlet side walls to become attached to the protein ions. In addition, the charge state envelope appeared to shift to favour higher charge states without cone gas.
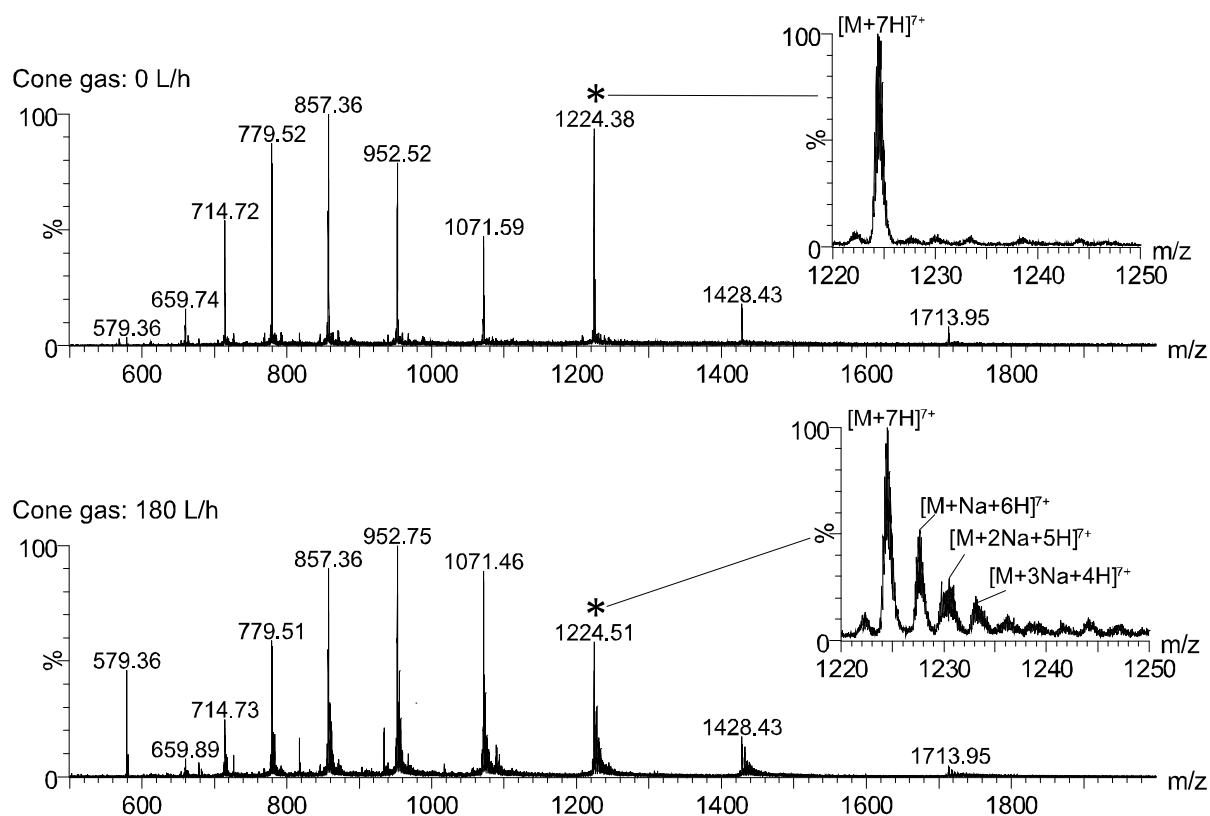
Figure 48: Liquid AP-MALDI mass spectra of ubiquitin (10 pmol) produced with (top) and without (bottom) cone gas. Sodiation was promoted when cone gas is used, as evidenced by the zoomed 7+ ion region.

Positive ion mode mass spectra included ions of the abundant proteins in the animal milk samples, however negative ion mode spectra were absent of useable information from multiply charged ions (Figure 49).
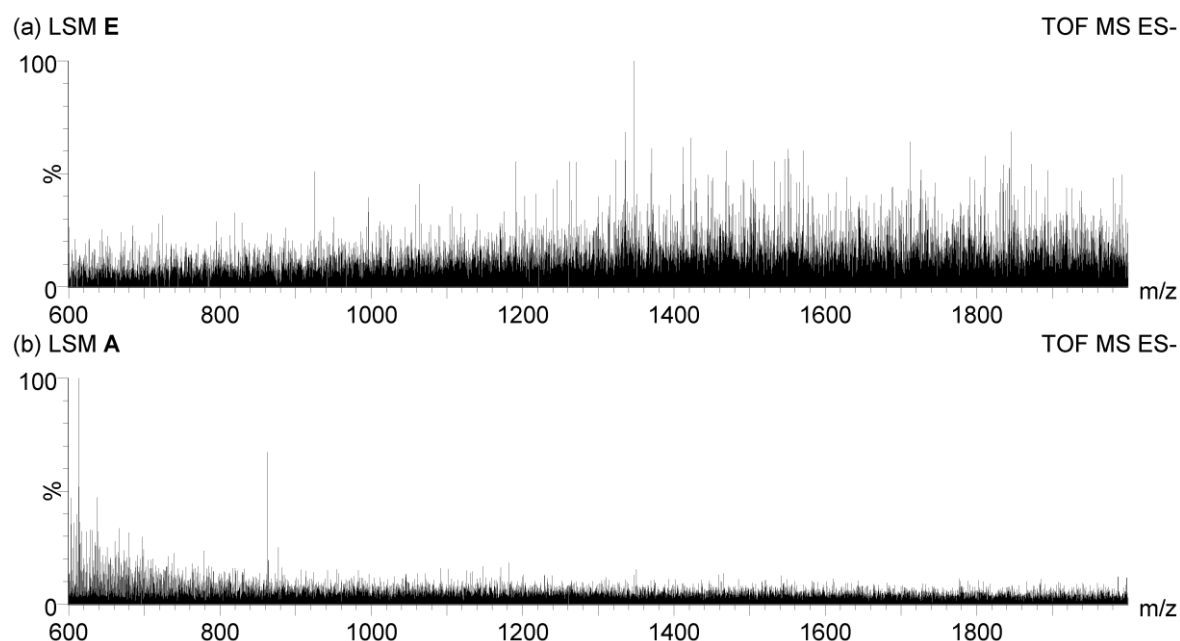
*Figure 49: Ion mobility-filtered AP-MALDI-TOF mass spectra (5 min acquisition) show a distinct lack of multiply charged ions detected from milk samples in the negative ion mode for the two liquid matrices **E** and **A**.*

Figure 50 shows representative mass spectra for the six milk classes, analysed in the positive ion mode after extraction by the protein-favourable TCA-based extraction.

Each sample was analysed for 60 s versus the 10 s used for the lipid-only samples to compensate for a relatively low TIC. Despite low analyte ion signal, the specificity of proteins to each species provided the opportunity to enhance the classification ability of the model. For example, bovine α-s$_1$-casein is 239 Da greater than caprine α-s$_1$-casein meaning the two produce distinguishable ions at high charge states. This is due to some differences in the amino acid sequences, as displayed in Figure 51.
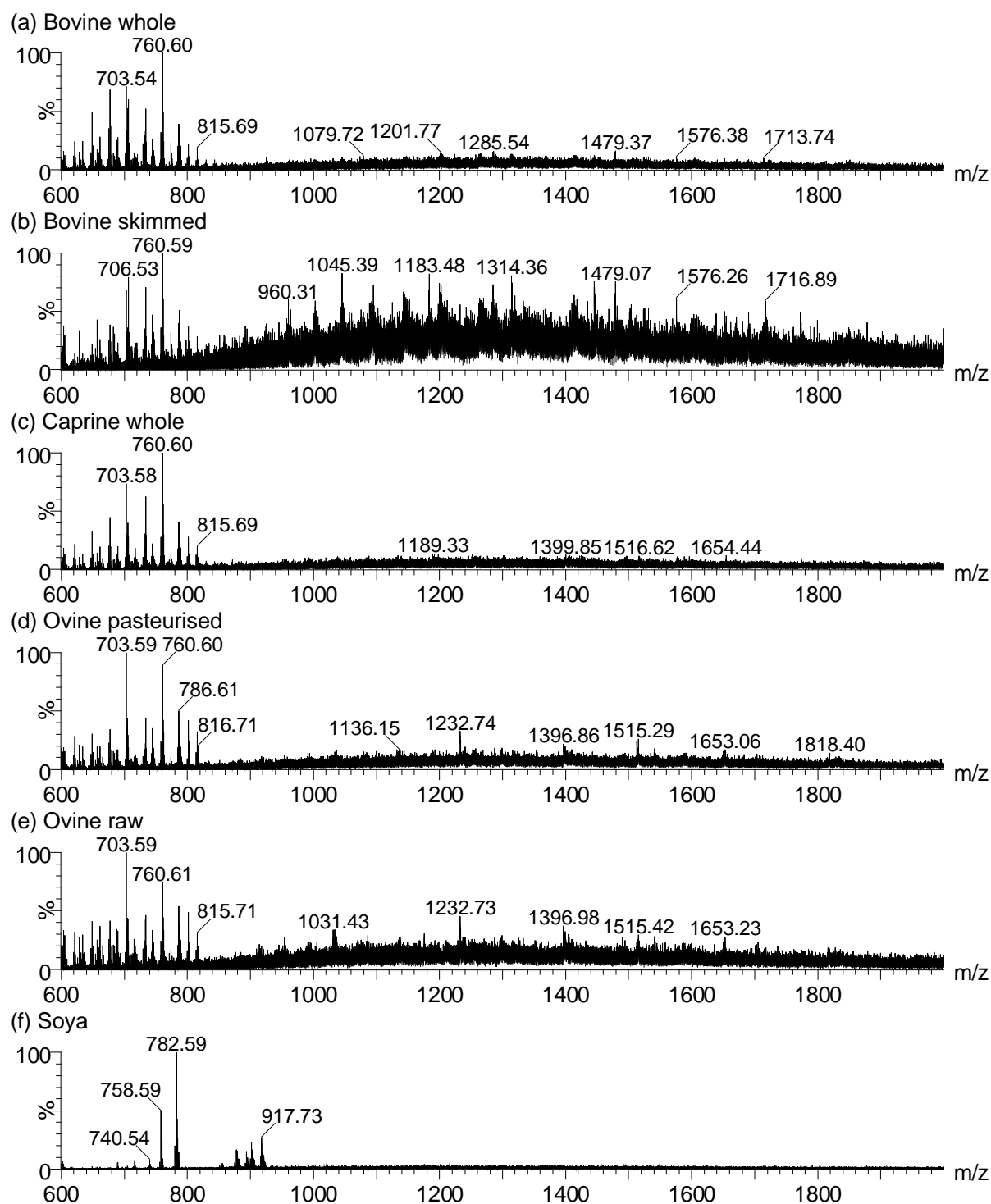
Figure 50: Positive ion mode mass spectra of the six milk types acquired over one minute. Ions of high-charge state for the abundant proteins in the animal milk (e.g. caseins) were detected above m/z 900.

```
α-s1-casein (bovine) Query   1   MKLLILTCLVAVALARPKHPIKHQGLPQEVLNENLLRFFVAPFPEVFGKEKVNELSKDIG   60
Sequence differences            MKLLILTCLVAVALARPKHPI H+GL   EV NENLLRF VAPFPEVF KE +NELSKDIG
α-s1-casein (caprine) Sbjct   1   MKLLILTCLVAVALARPKHPINHRGLSPEVPNENLLRFVVAPFPEVFRKENINELSKDIG   60

                     Query  61   SESTEDQAMEDIKQMEAESISSSEEIVPNSVEQKHIQKEDVPSERYLGYLEQLLRLKKYK  120
                                 SESTEDQAMED KQM+A S SSSEEIVPNS EQK+IQKEDVPSERYLGYLEQLLRLKKY
                     Sbjct  61   SESTEDQAMEDAKQMKAGSSSSSEEIVPNSAEQKYIQKEDVPSERYLGYLEQLLRLKKYN  120

                     Query 121   VPQLEIVPNSAEERLHSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWY  180
                                 VPQLEIVP SAEE+LHSMKEG  A QK+PMI VNQELAYFYP+LFRQFYQLDAYPSGAWY
                     Sbjct 121   VPQLEIVPKSAEEQLHSMKEGNPAHQKQPMIAVNQELAYFYPQLFRQFYQLDAYPSGAWY  180

                     Query 181   YVPLGTQYTDAPSFSDIPNPIGSENSEKTTMPLW   214
                                 Y+PLGTQYTDAPSFSDIPNPIGSENS KTTMPLW
                     Sbjct 181   YLPLGTQYTDAPSFSDIPNPIGSENSGKTTMPLW   214
```

Figure 51: the sequences of α-s1-casein for cow and goat aligned to show the differences in the amino acid sequences. This results in a mass difference between the two proteins and thus their multiply charged ions exhibit different m/z values.

A multivariate model featuring lipid, peptide and protein signals, for speciation of the six milk classes was built with 10 PCA components and 3 LDA components. The model used mass profiles from 30 technical replicates of each of bovine skimmed milk, caprine whole milk and ovine raw milk and 15 for each of bovine whole, ovine pasteurised and soya milk). This mass profiles were generated from AP-MALDI analyses lasting 1 minute per sample and included ions relating to the most abundant proteins in the milk typically with >m/z 1000 . The model parameters are shown in Table 7.

Table 7: Properties for PCA/LDA model for milk speciation incorporating high charge state ions.

Model Property	Value
PCA dimensions	10
LDA dimensions	3
Intensity limit	0
m/z range	600 – 2000
Binning (m/z)	1
Binning mode	Advanced
Lockmass (m/z)	706.5381, 758.5699 (Soya only)
Spectrum Interpretation	One per burn
Lockmass	Yes
Background subtraction	Yes
Normalisation	Yes

The projection of the LDA components is shown in Figure 52, where there is greater visual distinction between classes, compared to the equivalent lipid-only model (Figure 46a). Inspection of the score and loading plots in Figure 53 reveals highly charged ions that contribute to the classification power of the model. For example, the bins m/z 1515.5 and 1232.5 contributed to ovine milk classification and bins including m/z 1314.5 and 1445.5 contributed to bovine milk classification. The results of the '20% out' cross validation are shown in Table 8. There was a

notable improvement in classification accuracy over the positive ion mode lipid-only model (77.44% increased to 92.59%).

Table 8: '20% out' cross validation results for 5 PCA/LDA models.

Model	Number of spectra	Number correct	Classification accuracy (%)
Positive 4-class	266	266	100
Negative 4-class	268	246	92.16
Positive 6-class	266	206	77.44
Negative 6-class	268	247	92.16
Positive Proteins	135	125	92.59

N.B. '4-class' refers to models with the two bovine/ovine milk types (i.e. skimmed and whole, pasteurised and raw) combined. '6-class' refers to models where the two bovine/ovine milk types were treated as individual classes.

Table 9: Comparison of per-class classification accuracy for the positive ion mode models.

Class	Accuracy for Lipids-only model (%)	Accuracy for Protein-peptide-lipid model (%)	Accuracy increase (%)
Bovine skimmed	76	87	11
Bovine whole	56	93	37
Caprine	100	100	0
Ovine pasteurised	64	80	16
Ovine raw	71	93	22
Soya	100	100	0

The reason for this increase in overall accuracy was improved accuracy in classification of individual bovine and ovine milk types (Table 9). Bovine skimmed milk spectra were separated from bovine whole milk by highly charged ions corresponding to intact proteins such as β-casein; the bovine whole milk was better characterised by lipids in the range *m/z* 600 – 760. The pasteurised ovine milk spectra featured greater abundance of highly charged ions in bins *m/z* 1232.5, 1396.5 and 1513.5 than the raw milk.
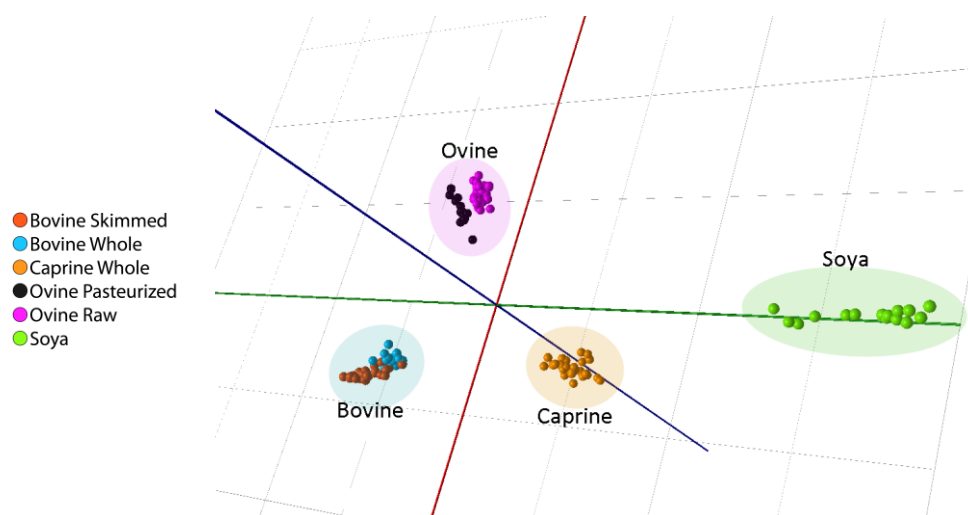
Figure 52: 3-dimensional LDA plot of six milk classes for the positive ion mode protein-peptide-lipid model; bovine whole (blue), bovine skimmed (red), caprine whole (orange), ovine raw (purple), ovine pasteurised (black) and soya (green). The shaded regions indicate the species.
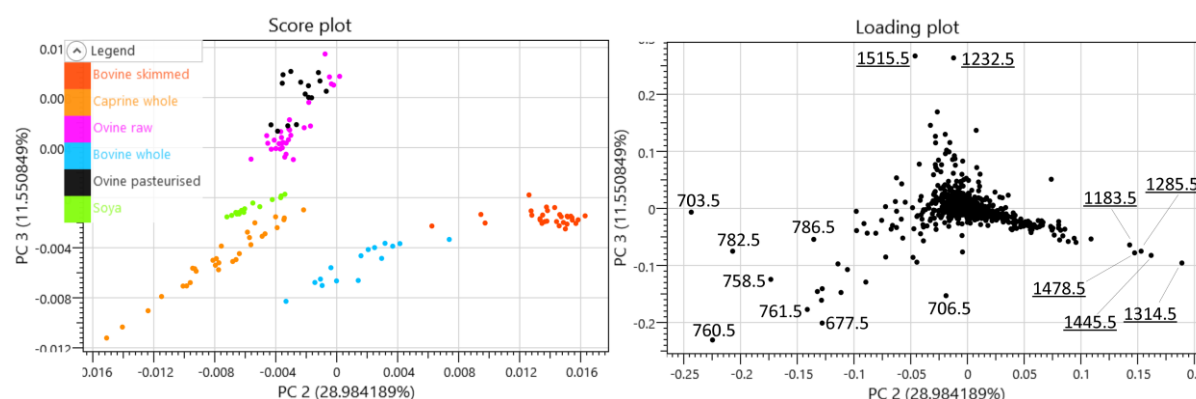
Figure 53: Score plot (left) and corresponding loading plot (right) from the six-class high-charge ion model. Some protein ions (underlined labels) strongly influenced the classifications.

3.2.5 Lipid Identification

Compound identification is not strictly necessary for classification models to be useful and, from a high-throughput perspective, adds complexity to experiments or data processing. However, knowledge of a model's influential compounds can feedback into understanding the underlying biology.

In this section, identification of compounds in TOF MS profiles is primarily based on accurate mass data. In addition, ion mobility data recorded in the N_2 gas-filled Synapt G2-Si ion mobility cell enabled an orthogonal dimension of separation. Additional dimensions of data can be used to filter out background signal and provide greater confidence in compound identification. The combination of *m/z* ratio, collision cross section (CCS), and Kendrick mass defect (KMD) has been demonstrated in the literature for identifying molecules in complex mixtures.[243, 244] The TWIMS cell was calibrated with polyalanine, allowing the CCS of the ions to be calculated. Practically, CCS

is a system-independent value, which means the values are comparable throughout the literature, whereas the raw TWIMS drift time measurements are not. In the positive ion mode, PC and PE [M+H]+ may be isobaric, along with similar *m/z* isotopic peak overlap with sphingomyelins (Figure 54).
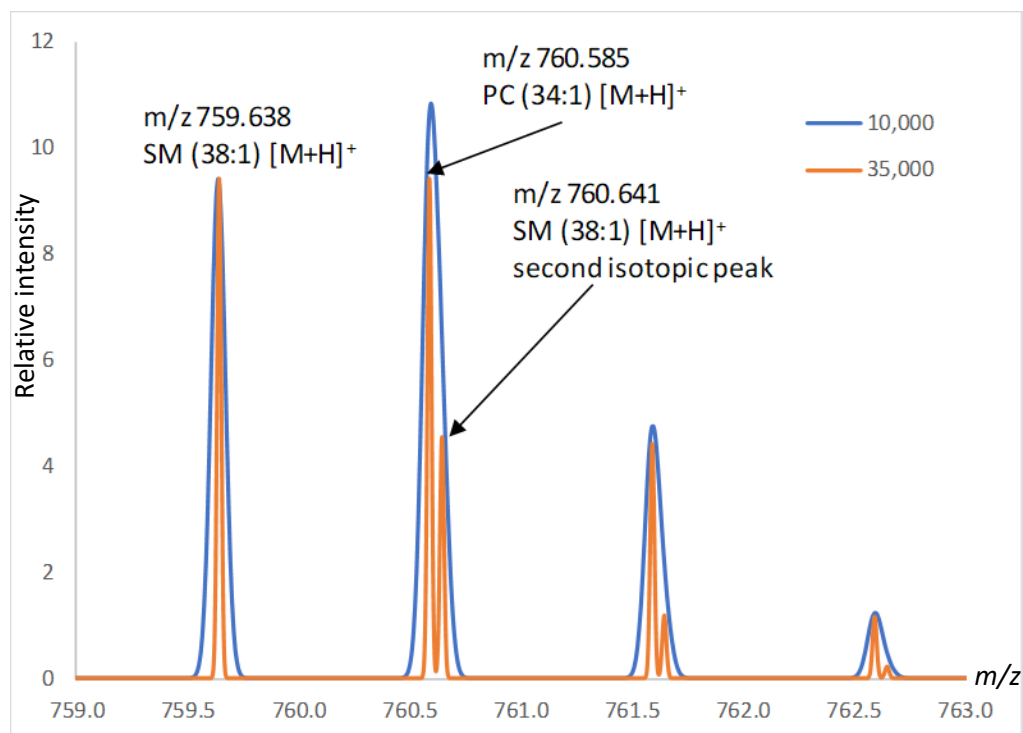
Figure 54: Example of the peak overlap of two lipids, SM (38:1) and PC (34:1). At a resolution of 10,000 (blue) the second isotopic peak of SM (38:1) [M+H]+ is unresolved from the [M+H]+ peak of PC (34:1). At higher resolutions this issue is alleviated.

However, each of these is discernible through IMS separation as they differ in structure, and thus CCS. In addition, matrix-related peaks and ubiquitous background ions like polyethylene glycol can be differentiated. Finally, the KMD was calculated to determine the degree of unsaturation. This provided a third value for identifying lipids as ions possessing differing degrees of unsaturation are separable by their KMD. This is a particularly useful feature for lipids, where there are many double bonds possible along the acyl chains.

Three-dimensional plots for lipid ions detected in positive and negative ion mode from caprine milk are shown in Figure 55. CCS for different lipid classes are powerful in distinguishing ions of similar *m/z*, for example isotopic peaks of [SM(38:1)+H]+ (*m/z* 759.637, CCS 305 Å^2) and [PC(34:2)+H]+ (*m/z* 758.569, CCS 293 Å^2). Since these ions have one and two double bonds respectively, they are also separated on the KMD axis. Phospholipid accurate mass and collisional cross section data were searched against the LipidMAPS structural database and Waters Metabolic Profiling library through Progenesis QI.[227] Additional CCS data were also sourced from the literature since this is a relatively new area of research and standardised data are only just

becoming widely available. A table containing identifications for phospholipids detected from goat milk by *m/z* and CCS can be found in Appendix 2 (Section 6.2).
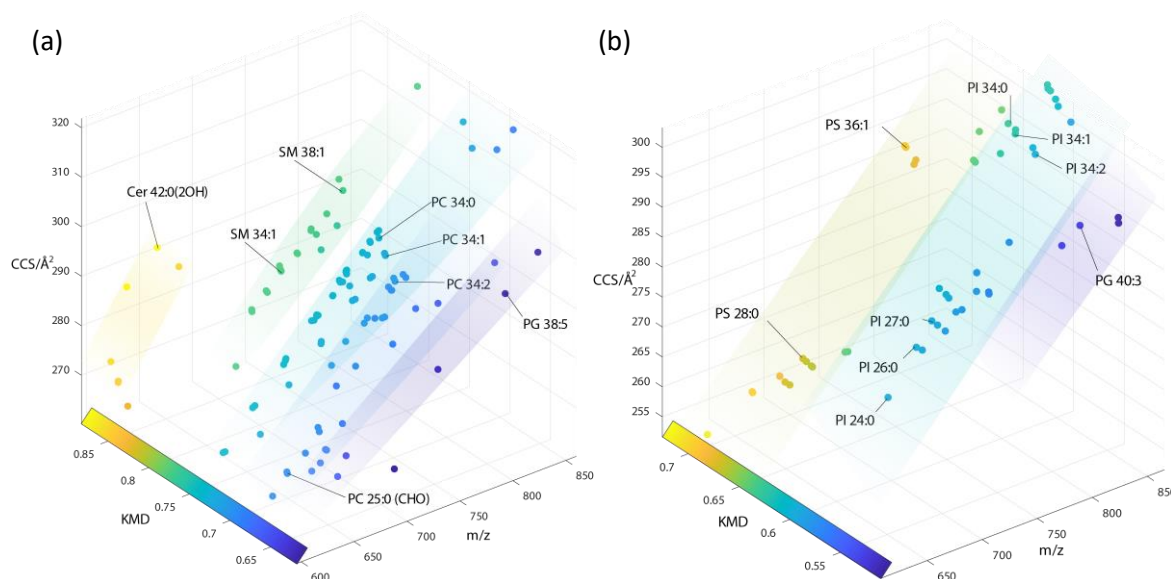
Figure 55: Three-dimensional plots of m/z, CCS and KMD for caprine milk lipid ions. Examples of ions are labelled for the positive ion mode (a) and negative ion mode (b).

As an alternative to identifications that can be made with MS/CCS/KMD, MS/MS experiments can be used for structural analysis, as was demonstrated in Section 3.1. Fragmentation products of phospholipids are generally able to establish the lipid headgroup identity, but information about the side chain structure is lacking. This was a driving force behind the experiments described in Section 3.1. Additionally, a mass profile can include hundreds or thousands of individual compounds without chromatographic separation. This makes MS/MS of phospholipids challenging due to overlap and convoluted signals.

Identification of higher charge state ions is discussed in detail in Section 3.3.2.

3.2.6   Dual ion mode classification models

Since the positive and negative ion mode lipid mass spectra featured exclusive classes of lipids, the possibility of models of greater classification power from the combination of this data was investigated. Multiple challenges arose with this; foremost, all commercial TOF analysers typically operate in only one polarity at a time, the Synapt G2-Si included, although there are some references to rapid polarity switching technology in the literature.[245] Thus, data acquired from the positive ion and negative ion modes had to be combined in an offline process. The "Combine All Files" function of MassLynx was not capable of combining data from the two different ion modes to create summed spectra. At the time of writing, AMX does not feature tools for assigning multiple spectra to one sample, for combining multiple data files or support for importing an

externally-produced data matrix. This final point was crucial for the eventual test workflow; thus, two Progenesis QI projects were used to pre-process the lipid spectra from each ion mode with background subtraction, normalisation and peak picking and EZInfo was used for multivariate analysis. EZInfo allowed the export of data matrices for each polarity, which were combined in Microsoft Excel. The combined matrix was imported back into EZInfo. A similar classification method to LDA, PLS-DA, was available in EZInfo. PLS-DA is another supervised machine learning technique. Like LDA, the discriminating factors between sample groupings are assessed and subsequently used for classification of unknowns.[246] PLS-DA models were then built with the lipid spectra for each individual ion mode, followed by the combined data matrix.

A combined polarity model was constructed from lipid profiles of the 6 milk classes. Data were first acquired in the positive ion mode (TWIMS cell and TOF analyser calibrated with 10 µg/mL polyalanine + LSM **A**) from sample droplets co-spotted with LSM **A**. These same samples were reanalysed without the need to re-spot after switching the TOF analyser to the negative ion mode. Polarity switching was performed immediately after positive ion mode analysis, followed by a period of approximately 30 minutes to allow the electronics to settle. Negative ion mode calibration of the TWIMS cell and TOF analyser was performed with polyalanine (10 µg/mL) and LSM **A** after settling. Samples were analysed for ten seconds in each ion mode, with the laser pulse repetition rate of 30 Hz and approx. 20 µJ of laser energy per pulse. The PLS-DA model projection is shown in Figure 56. The results of the cross validation of this combined polarity model are shown in Table 10. This initial investigation suggests that there is the potential to increase classification power through the combination of positive and negative ion mode spectra. Total correct classifications for the positive and negative mode models were 94.4% and 84.6% respectively, whilst for combined data this was 98.4%. Whilst this is not suitable for online classification with current Synapt-generated data, offline classification could still benefit. Mass spectrometer types able to perform real-time polarity switching on the second timescale could also take advantage of this. Additionally, the technology for dual polarity TOF analysers has been demonstrated; it may be that commercial versions are available in the future.[247, 248] Electronics that allow faster switching may yet be implemented on a future QTOF platform, although the cost is currently prohibitive.

Table 10: PLS-DA model cross-validation results for individual and combined ion mode analysis of milk species.

Model	MS Polarity	Correct classification rate %
	+	94.4
	−	84.6
PLS-DA 6-class	+ and − (offline combination of MS data from consecutive runs)	98.4

N.B. '6-class' refers to both bovine milk types, both ovine milk types, caprine and soya milk.
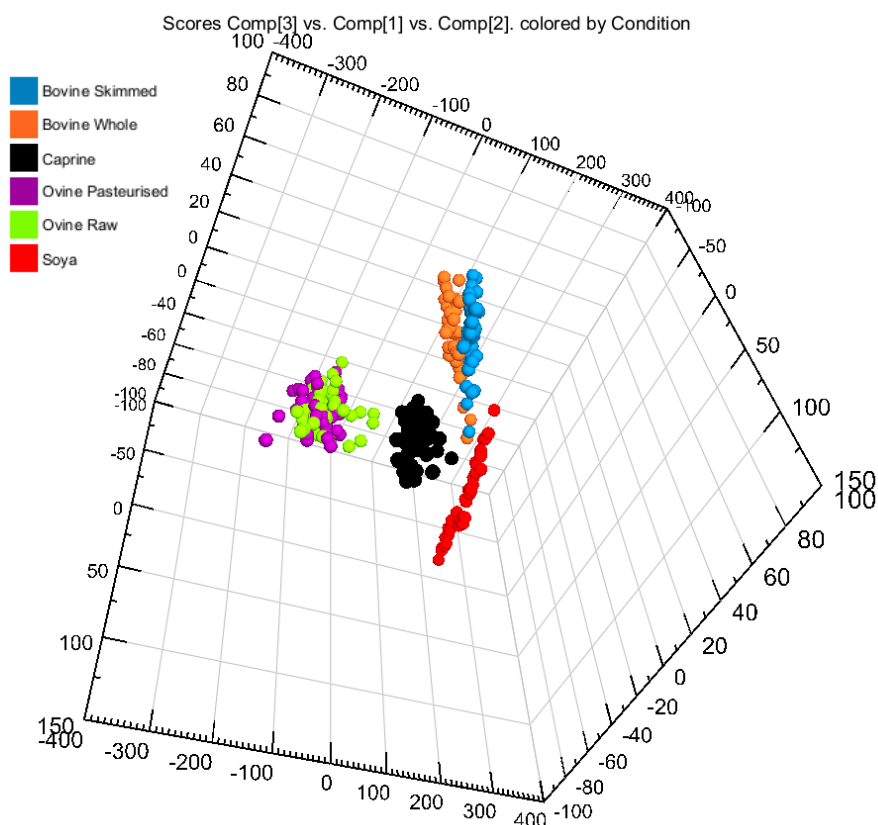
Figure 56: PLS-DA plot of lipid profiles for 6 milk classes with positive and negative ion mode data used for construction.

3.2.7   Milk adulteration level estimation

As an extension of the former sections, the potential for food product adulteration detection using a similar procedure as described above was assessed. Calvano *et al.* have previously described a MALDI-TOF MS method showing that the ratio of PC (30:0) [M+H]+ (*m/z* 706.6) and SM (16:0) [M+H]+ (*m/z* 703.6) can be used to determine whether milk is of bovine, caprine or ovine origin.[178] Furthermore, they showed that from mixtures of two milk types the ratio of the two ions could be used as a predictive value for the %composition. The method was deemed to function as an adulteration screening test, to as low as a 10% adulteration level.

A MALDI-TOF MS method for detecting adulteration of edible oils used PCA to distinguish oil types.[176] This method also used the ratio of two lipid ions to estimate olive oil adulteration levels. Both lipids were TAGS; the ratio of m/z 881.75, which contains palmitic acid, to m/z 907.76 increased with greater olive oil content. The limit of detection was estimated to be approximately 2% olive oil in canola oil. Both studies relied on comparing only two ions for adulteration estimates. This would make them easy to fool if they were standardised tests, since the adulteration could be masked by addition of the molecule in deficit to change the ratio. In the following section, use of the dimensions of a multivariate predictive model is demonstrated for adulteration level estimation.

Positive ion mode spectra from HIP lipid extracts were acquired for 10 seconds per spectrum, at 30 Hz laser pulse repetition rate, for caprine-bovine milk mixtures in 7 classes; 0%, 5%, 22%, 40%, 60%, 77% and 100% bovine milk, 11 samples per group class. The percentage composition of the mixtures was determined by mass, rather than volume due to the potential for inconsistencies in pipetting because of the milk consistency.

First, for comparison to the literature methods discussed above, the ratios of some abundant ion pairs in the same AP-MALDI-TOF MS profiles were evaluated. The regression plots are shown in Figure 57. Specifically, the R^2 for the ion pair ratio 706.5/703.6 (Figure 57a) was 0.9891, comparable to the value reported by Calvano *et al.* (0.9954).

To demonstrate the use of multivariate models for composition determination, an orthogonal (O)PLS-DA model was built in EZInfo. OPLS-DA allows for better visualisation than PLS-DA, and reduces model complexity by removing variation in descriptor variables that is orthogonal to the property variables.[249] In this application, the descriptors are the sample classifications, and the property variable is peak intensity. This initial model contained only 100% bovine whole and 100% caprine whole milk spectra with the m/z range 600-1000. This made an orthogonal model with single-dimension 't[1]' class discrimination based on the differences between bovine and caprine milk, and intraclass discrimination 'to[1]'. Spectra from the caprine-bovine milk mixtures were processed and overlaid on these two dimensions, where separation on t[1] occurred per the mixture composition. Figure 58 shows the individual samples per class plotted in colour. Black points show the mean for each class with error bars indicating the 95% confidence interval. The trendline R^2 of 0.9996 indicates a linear relationship between %mass bovine milk and mean interclass separation, t[1]. This type of plot would be suitable for use as a calibration curve for binary mixture content estimation. The model is likely to be more robust to individual molecule tampering since it is not dependent on only two ions.

The R^2 value of the OPLS-DA model suggests it is beneficial to assess many spectral features for this application for best linearity. The R^2 for the ion pair ratio 706.5/703.6 reported by Calvano *et al.* (0.9954) is slightly worse than the 0.9996 achieved with the OPLS-DA model. It is also true for all additional ion pair ratios investigated here (Figure 57). It is likely that the reason for this is that the variability in relative intensity when monitoring just two ions is greater than when all ions within a range are considered, thus leading to a less linear fit. With both the ion pair and OPLS-DA approaches, it is advantageous to produce the calibration plot based on the mean of multiple measurements since variation is evident on an individual sample basis.
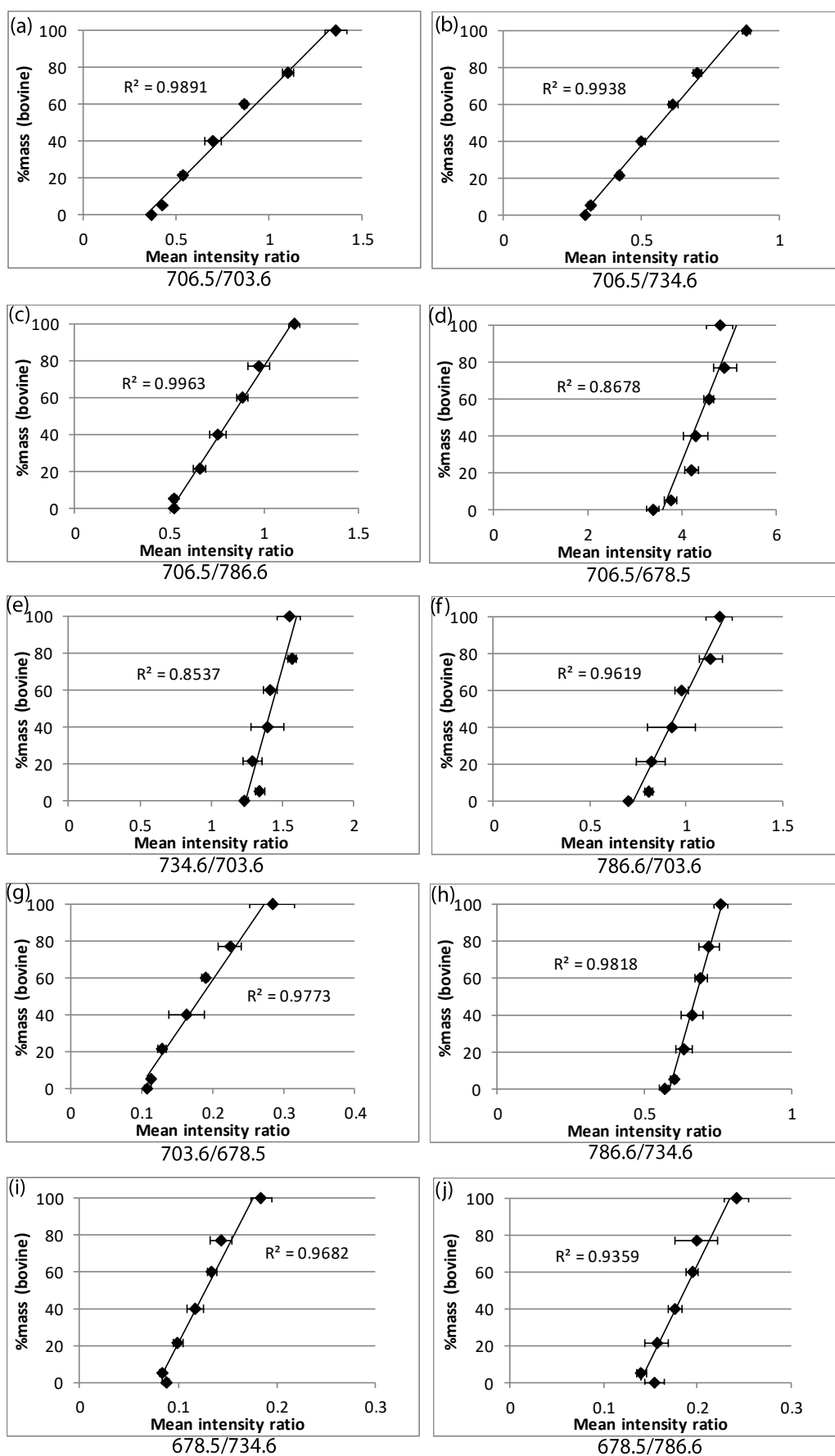
Figure 57: Regression plots for ion pairs in the same bovine/caprine milk mixtures used for Figure 58. The coefficient of determination, R^2, is displayed for each plot. The error bars show the 95% confidence interval.
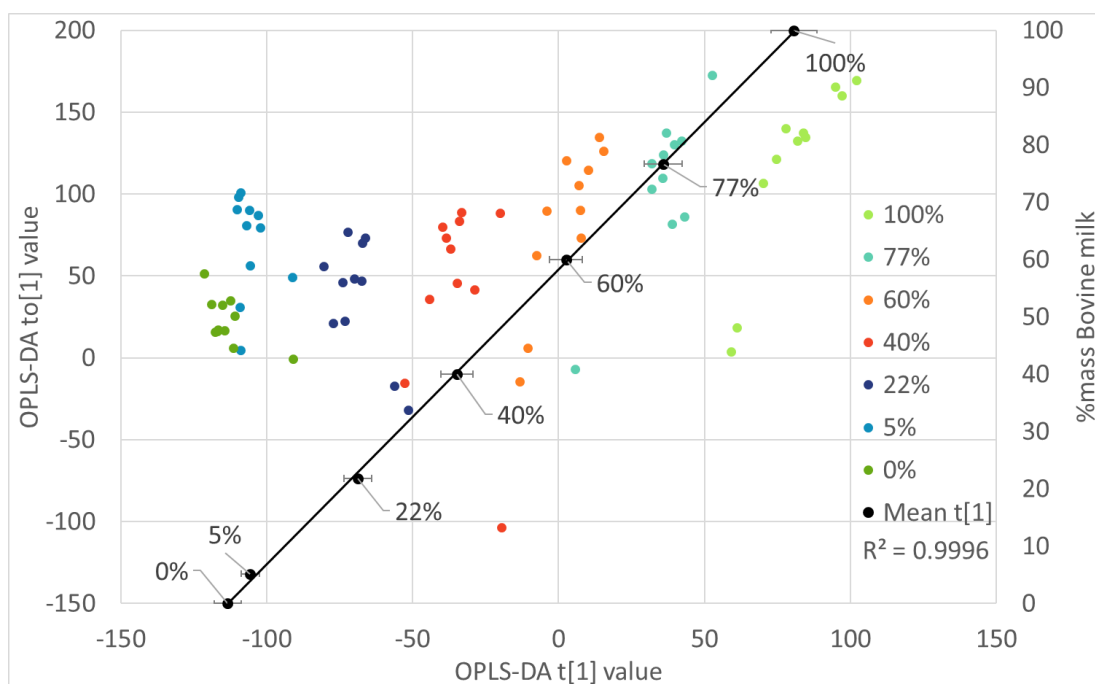
Figure 58: OPLS-DA plot for mixtures of bovine and caprine milk (% mass bovine milk). Black points mark the mean of each class and error bars indicate the 95% confidence interval. The R² value indicates a linear relationship between sample composition and t[1].

Composition prediction modelling with PCA/LDA is also possible using AMX. As of AMX version 1.0.1563.0 the data matrix does not export the raw LDA dimension values, so plotting of a calibration curve is not possible. The model was built using the parameters shown in Table 11 and the LDA model is shown in Figure 59.

Table 11: Parameters for constructing the bovine/caprine mixtures PCA/LDA model.

Model Property	Value
PCA dimensions	10
LDA dimensions	2
Intensity limit (arb.unit)	0
m/z range (ion mode)	600 – 1000
Binning (*m/z*)	1
Binning mode	Advanced
Lockmass (*m/z*)	760.5851
Spectrum Interpretation	One per burn
Lockmass	Yes
Background subtraction	Yes
Normalisation	Yes

AMX '20% out' cross validation resulted in a correct classification rate of 95.95%. The failed classifications were always classified as a member of an adjacent group (e.g. 'bovine 5%' identified as 'caprine 100%), not unreasonable considering their closeness in the predictive space.
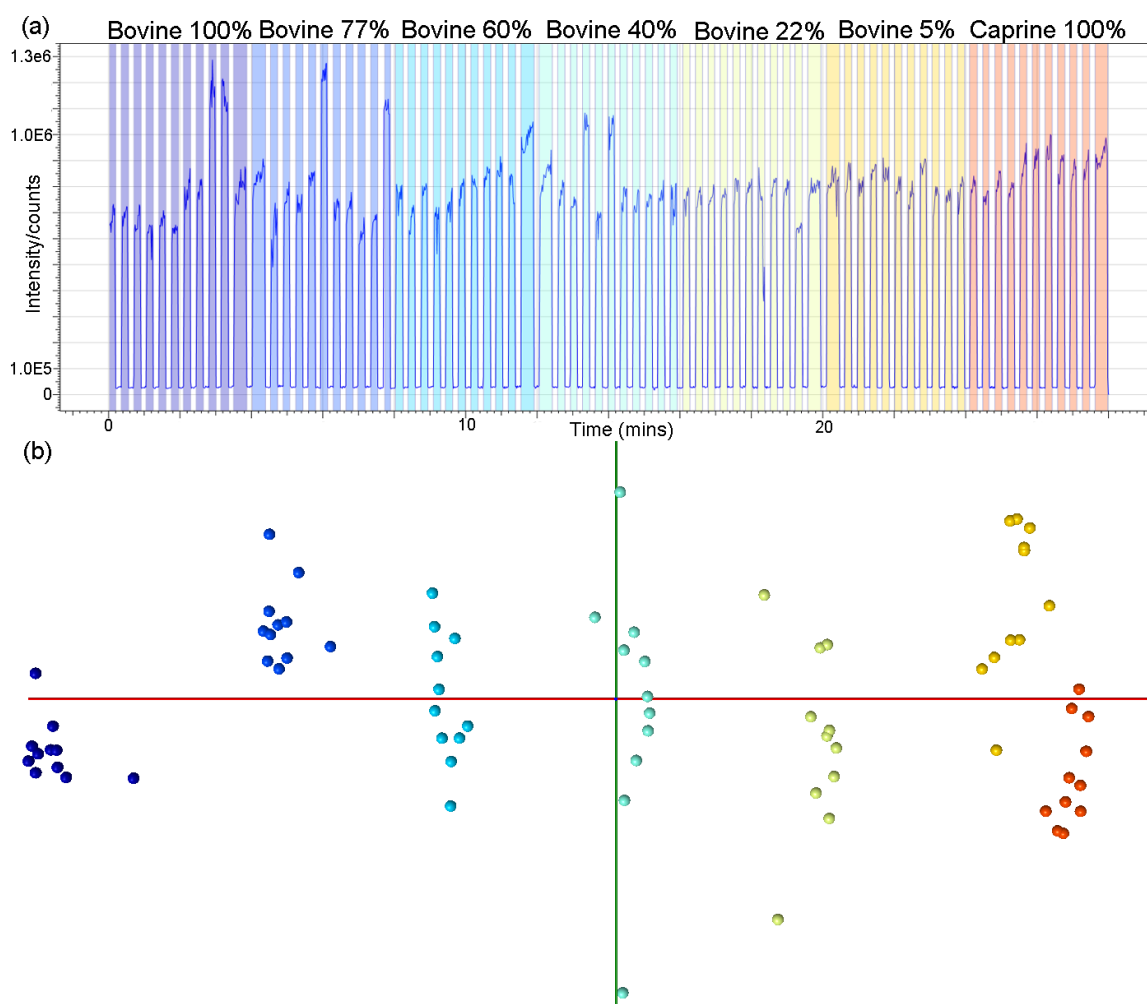
Figure 59: (a) Chromatogram as seen in AMX, where samples were automatically detected, and their classification defined (b) 2D LDA plot from mass profiles of mixtures of bovine and caprine milk (%mass). The colour of the sample indicates their classification as specified in (a). The samples are separated along the x-axis, which is the component that enabled discrimination between classes.

AMX enabled the use of 'AMX Recognition [Beta] version 1.0.1563.0' for real-time and offline classification of samples. A disadvantage of this over the plotting of a calibration curve is that samples can only be classified into discrete groups. The classification scale is only as fine as the user decides by producing calibration samples for those discrete classes. Nevertheless, the speed of post-processing and the ability to run the software in real-time with sample acquisition makes for a rapid screening process. To demonstrate this concept, the PCA/LDA model described above was exported to AMX Recognition. The same data was loaded into the software for assessment, with the settings as shown in Table 12.

Parameter	Value
Apply Lockmass correction	Yes
Apply background subtraction	Yes
Threshold	10000 counts
Lockmass	*m/z* 760.5851
Outlier type	Standard deviation
Standard deviation multiplier	5
Intensity limit	$6E^5$ counts
Scans per spectrum	10
Good spectrum timeout	1 s
Probability threshold	Global (0%)

Figure 60 shows the processing window. The classification of MS scans is shown in the right pane, with the selected or current scans in the left pane. The bottom pane shows the chromatogram. Lockmass errors were reported when the laser was not sampling i.e. there was no *m/z* 760.5851 peak to recalibrate to and indicates the inter-sample period. This does not apply where the 'good spectrum' timeout has been set optimally and 'bad' scans are ignored. The classification results of the model when assessed using AMX are shown in Figure 61.
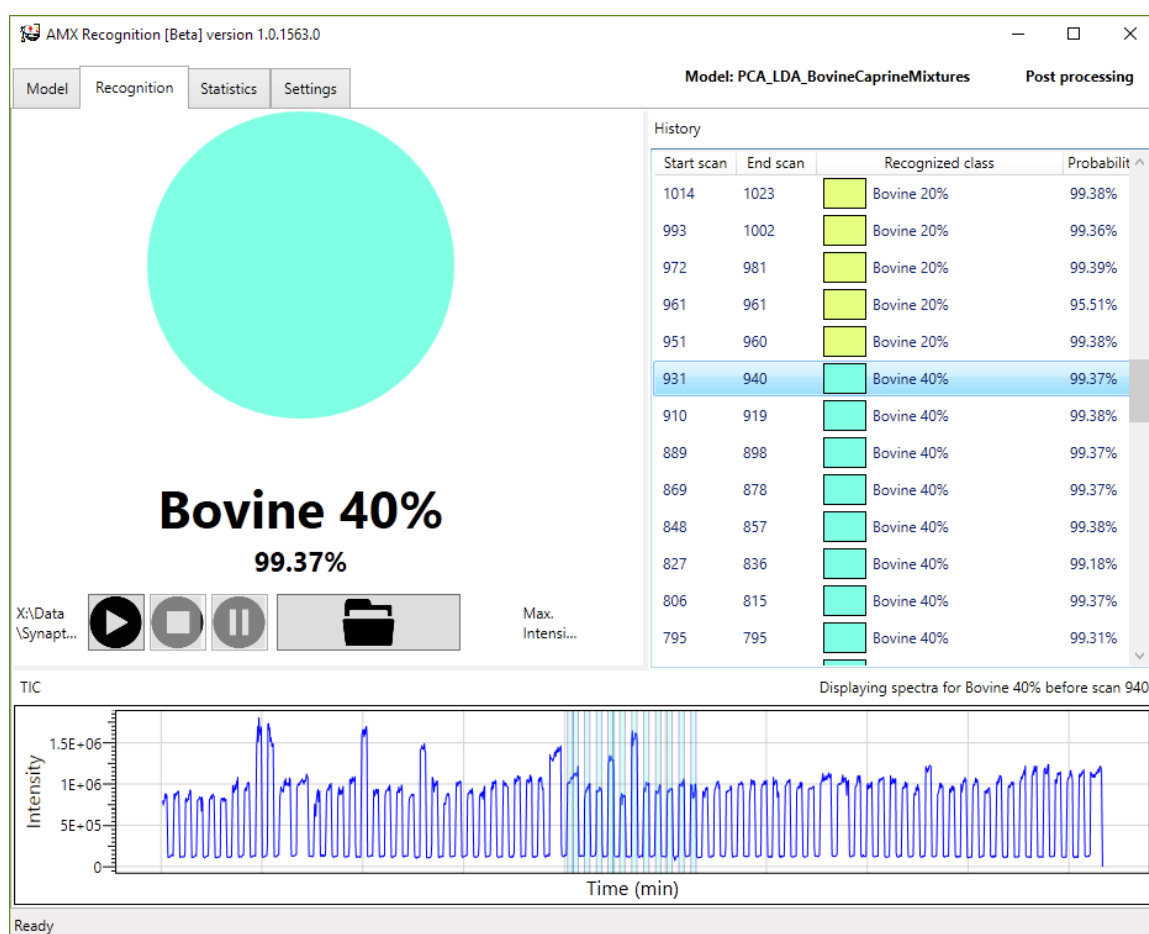
Figure 60: Demonstration of AMX Recognition in 'Post-Processing' mode. The left pane shows the current or selected MS scan, the right pane shows the scan list and the bottom pane contains the chromatogram. Ten scans were combined per reported classification and the 'intensity limit' set to 6e5 counts to correctly select samples.

It must be mentioned that AMX Recognition will require some usability upgrades to be fully compatible with the data produced by AP-MALDI. It is not easy to set the software to correctly detect samples and ignore background, particularly for samples with many scans of low signal intensity. This will be discussed later in relation to the contents of Section 3.3. In this case it was less of an issue because the signal intensity was high compared to the background. However, using too few scans would result in inaccuracies as shown in Figure 62.
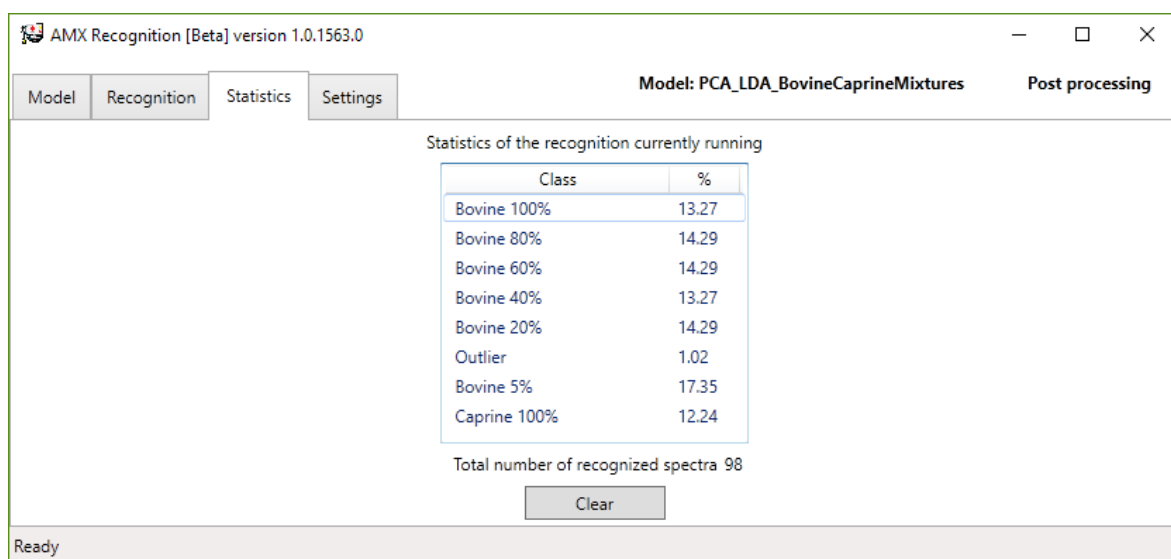
Figure 61: Statistics for the reclassification of the samples used to build the model. 100% class accuracy is theoretically achieved at 14.29%.
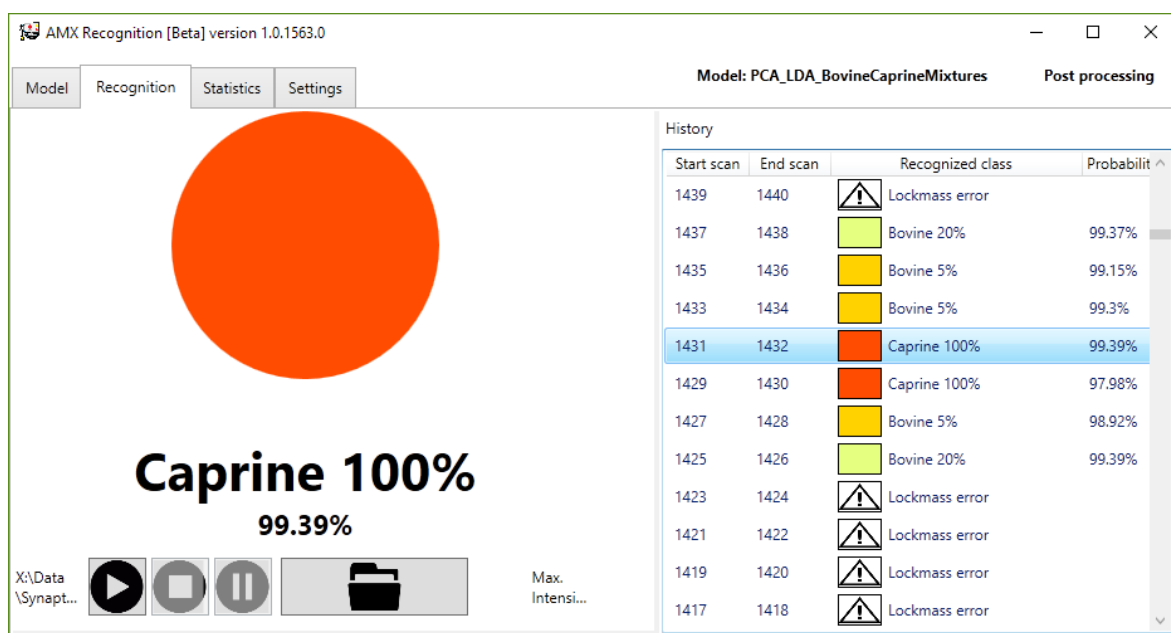
Figure 62: Example of classification inaccuracy in AMX, linked to sampling too few MS scans (2/spectrum); the correct classification for the sample in scans 1425-1439 was caprine 100%, but the high sampling rate did not allow a suitably intense spectrum to be produced.

3.3 Detection of mastitis in dairy cows by liquid AP-MALDI MS of sampled milk

The development for this method began with an auspicious visit to the Centre for Dairy Research (CEDAR) at the University of Reading. A test set of milk samples was acquired from healthy and mastitis-affected cows. Additionally, some samples with a high somatic cell count, but not diagnosed as having mastitis, were included. In total, 10 healthy cows, 9 with mastitis and 10 with high SCC were included in the initial sampling. After HIP extraction, liquid AP-MALDI MS profiling and PCA in Progenesis QI, a loose clustering of samples according to the above groups was demonstrated (Figure 63). From this point, the method was developed using aqueous extractions to include more biologically specific molecules (proteins, peptides), as is explored in the following sections.

Figure 63: The initial principal component analysis performed on HIP extracts using positive ion mode liquid AP-MALDI mass spectra. Blanks were the solvents used for lipid extraction (HIP).

3.3.1 Classification

The method was further developed for the Synapt's AP-MALDI ion source, utilising the heated inlet to generate multiply charged peptide and protein ions, in addition to the lipids used in the original experiment. The samples were prepared by the method in Section 2.7. AP-MALDI-TOF MS profiles of 135 individual samples from two groups 'Healthy' and 'Mastitis' were compared (example spectra in Figure 64). For healthy samples, lipids and highly charged milk protein ions were detected. Mastitis samples featured additional ions, with charges between 2^+ and 7^+ indicative of peptides and larger protein fragments. In addition, the protein peaks detected in healthy samples were reduced in intensity or absent in the mastitis samples. The peaks at *m/z* 868.86 and *m/z* 724.23 were obvious differentiating ions in the example spectra. However, using the multivariate statistical techniques discussed in Section 3.2 allowed multiple spectral features to be evaluated. An ion mobility-filtered spectrum that shows multiply charged ions from a mastitis sample is shown in Figure 65, revealing the richness of the information available from

peptide and protein ions.

Figure 64: AP-MALDI-QTOF MS spectra representative of the two classes, (a)'Healthy' and (b)'Mastitis'. Ions of m/z less than 600 were not used for classification to avoid matrix-related ions. Underlined labelling indicates an ion charge >1.

Figure 65: Ion mobility-filtered AP-MALDI-QTOF MS spectrum focused on the range m/z 500 -1000 for a mastitis sample featuring many multiply charged ions. The singly charged ion signals obscured many of the peaks without the filtering applied. Charge states were determined through a combination of isotopic pattern assessment and deconvolution.

Cone gas was not used for the mastitis model data acquisition as it was detrimental to the ion signal intensity of higher charged ion species and led to increased sodiation (as demonstrated in Figure 48). The signal intensity per second was relatively low compared to the HIP extracts, so an acquisition time of one minute was required allowing for a more intense summed spectrum. This especially benefitted the protein ion signals. Further improvements to sample preparation and the ion source could reduce the acquisition time, increasing the throughput of the method. Initial

experiments with the 1 kHz Nd:YAG laser have suggested multiply charged ion signal intensity from these complex samples could not be enhanced just by increasing laser pulse rate, regardless of the cone gas setting. This was in contrast to results of liquid AP-MALDI experiments that have shown an approximately linear increase in signal intensity of multiply charged peptide ions at pulse rates up to 500-1000 Hz with a Nd:YLF laser (λ = 349 nm; approx. 8 ns pulse duration).[132] Where that research showed the potential for a 100-fold increase in analysis speed (from 10 Hz to 1 kHz), the same benefit was not found in a brief investigation with the 1 kHz laser in our set-up (λ=355 nm). While a benefit was seen with the doubly and triply protonated angiotensin I ions, it was not observed to the same degree with peptide ions in the crude milk extracts (Figure 66), which may be due to the effect of the biological matrix. It is likely that the high pulse repetition may effectively 'drill' through the liquid droplet if replacement of material by healing is not fast enough. This would lead to a lower than expected increase in absolute signal intensity. It may be necessary to use laser pulse energy nearer to the threshold energy when pulsing at higher frequency to avoid this.
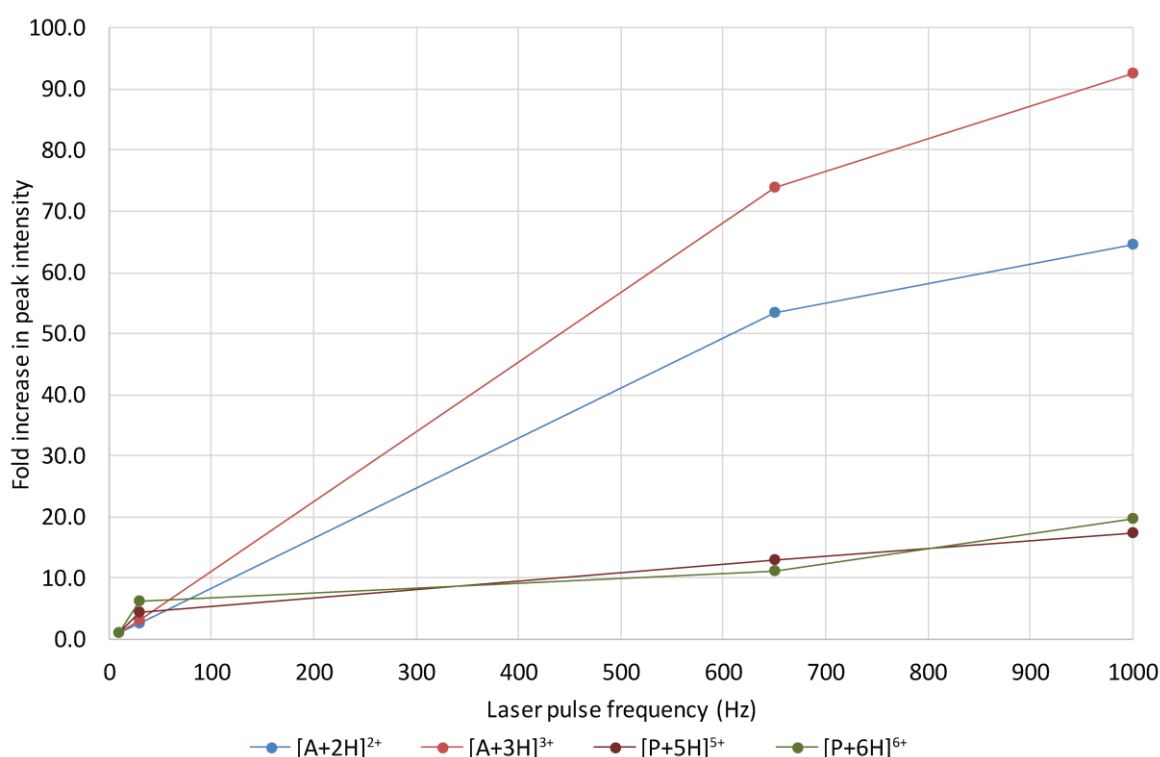
Figure 66: Fold increase in signal intensity at 10 Hz (all signals = 1), 30 Hz (λ=337 nm laser) to 650 Hz and 1 kHz (λ=355 nm laser) pulse rates for ions of the peptide standard angiotensin I (A) and the 4340 Da peptide (P) found in mastitis samples.

Figure 67 shows mass spectra of a mastitis sample analysed at 1 kHz. Compared to Figure 64, the relative ion signal intensity for multiply charged ions is significantly reduced with respect to lipid signals. However, there are multiple factors to consider here. First, two different lasers were used each featuring a different UV wavelength and pulse width, likely influencing how much material was ablated from the droplet. The ion suppression effects of a complex biological matrix like milk is also likely to impact signal intensity. This could be investigated by spiking peptide standards into the milk and assessing signal intensity versus that of a standard of the same concentration in high-purity solvents. As such, further investigation of different lasers and sample preparation modifications may provide advantages at a higher repetition rate.
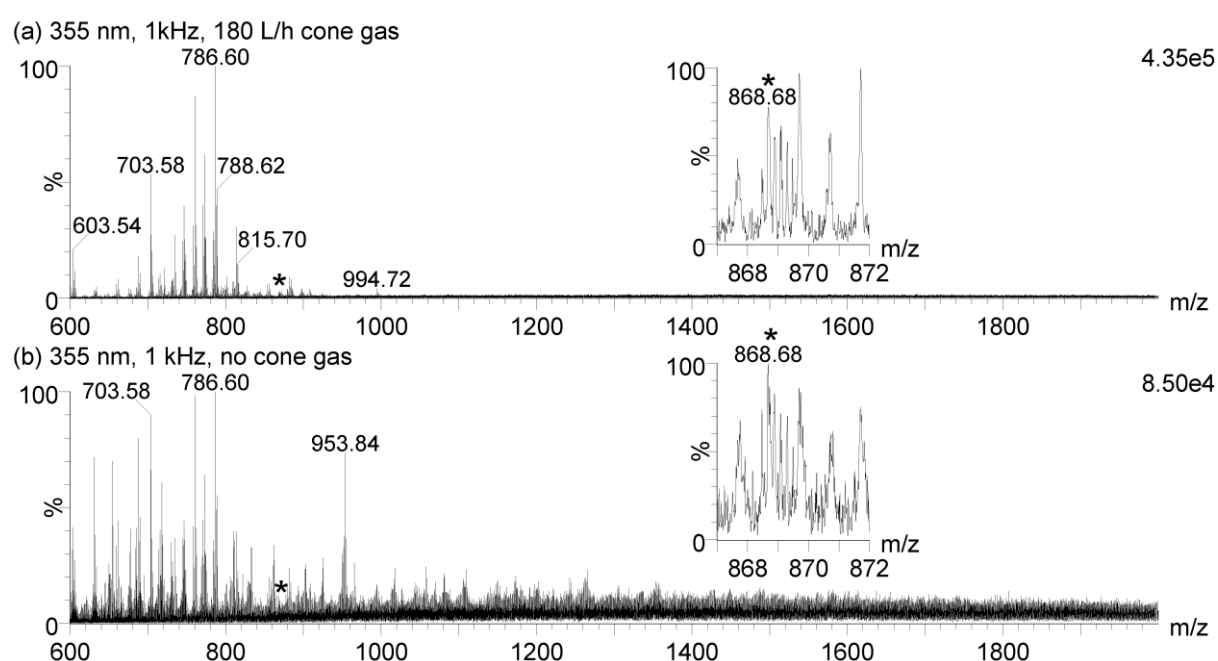
Figure 67: Liquid AP-MALDI mass spectra acquired with a 1 kHz Nd:YAG laser with a pulse energy of approximately 22 μJ (a) with cone gas and (b) without. Multiply charged peaks had a lower relative intensity compared with those acquired with a N_2 laser at 30 Hz (Figure 64). The inset shows the example peak at m/z 868.68, which is prominent at 30 Hz with the N_2 laser.

For healthy and mastitis sample classification with data acquired using the 337 nm laser, a model featuring 10 PCA dimensions and 1 LDA dimension was constructed from a pool of 73 healthy mass profiles and 62 mastitis mass profiles acquired in automated sequences. Each mass profile represented a single sample from a cow on a given day. Some cows were sampled up to three times for this model but always with at least a week between sampling. The model properties are shown in Table 13, and an example of the AMX sample processing window is shown in Figure 68. AMX was used rather than Progenesis QI because of its better native handling of the 'burn' chromatograms, more streamlined process for classification and the cross-validation tools. The processing window enabled manual checking of sample detection and summed spectra. This was

necessary with the mastitis model, since the TIC was lower, and the automatic processing would sometimes select samples incorrectly. Variability in ion signal peak intensity between samples was minimised through preparation of extracts with fresh solvents, reagents and liquid matrices. Samples had all encountered two freeze-thaw cycles, once during collection and once for aliquoting. Samples were all stored and extracted in the same type of 1.5 mL centrifuge tube to reduce the risk of variability due to different polymer-related background ion signal. Likewise, the same type of pipette tips was also used. Three stainless steel MALDI sample plates were used exclusively and were cleaned with the same procedure. They were rinsed with HPLC grade water, followed by HPLC methanol then sonicated in 1:1 methanol/water for 5 minutes. Finally, the sample plates were blown dry with nitrogen, something which, anecdotally, reduced the frequency of liquid droplets escaping the sample plate wells. Droplets were observed prior to analysis to ensure homogenous appearance. Droplets that did not keep the 'domed' profile on the target were spotted again. Normalisation was applied to each mass profile based on the total ion count in order to account for ion production between sample droplets. This was essential to reduce variability introduced by the sample wells on the MALDI sample plate, the condition of the heated inlet and other sources of day to day variation. The lockmass peak for each class was also monitored. Samples where this peak was not detected were rejected since this indicated a problem at some point during sample preparation.

Table 13: AMX Parameters for the mastitis detection PCA/LDA model.

Model Property	Value
PCA dimensions	10
LDA dimensions	1
Intensity limit	0
m/z range	600 – 2000
Binning (m/z)	1
Binning mode	Advanced
Lockmass (m/z)	760.5851
Spectrum Interpretation	One per burn
Lockmass	Yes
Background subtraction	Yes
Normalisation	Yes
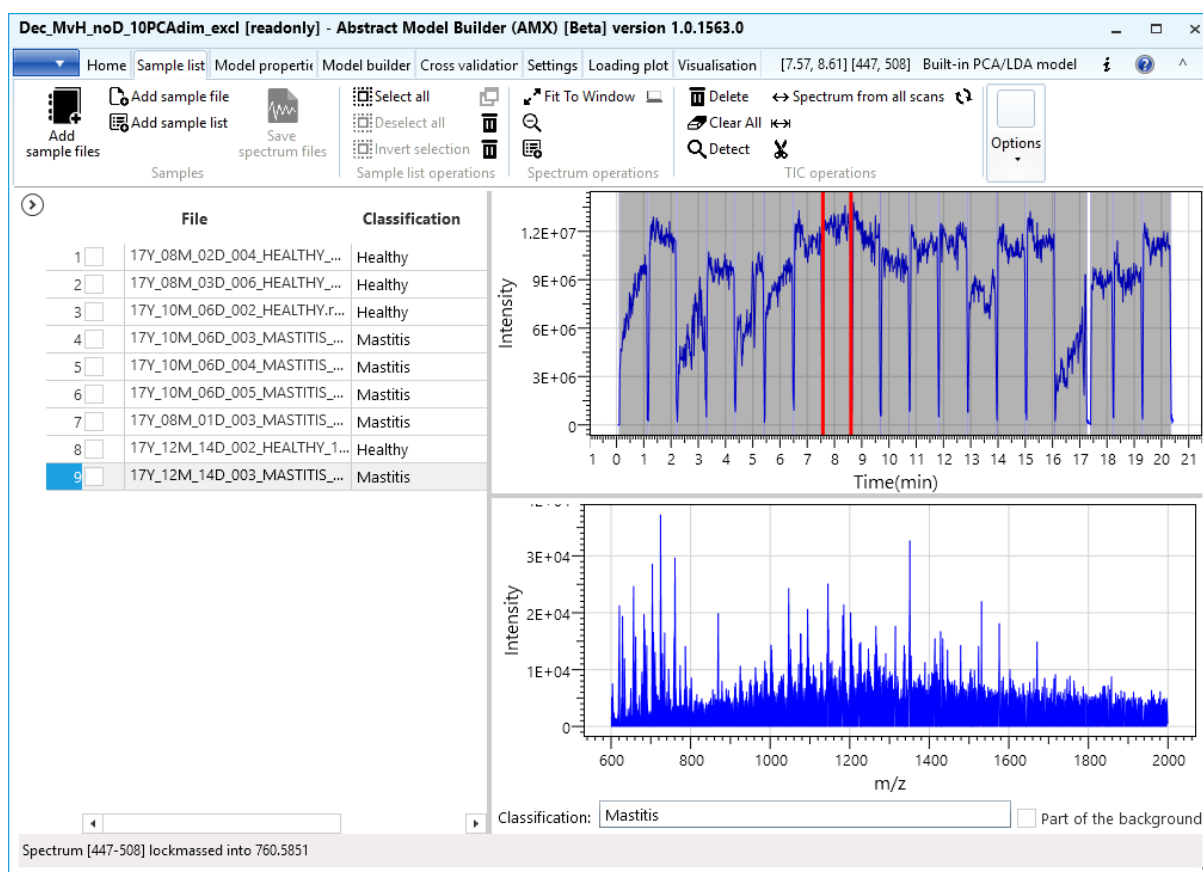
Figure 68: Screenshot of the AMX sample list, automatic sample detection and mass spectrum display.

Separation along PC2 showed the analysis detected differences between the sample groups, without additional classification information being supplied (Figure 69a). The PCA loading plot and LDA1 mass loading plot (Figure 69b and c respectively) show that the classification relies strongly on the abundant multiply charged ions detected, in addition to some phospholipids, as for the food analysis application.
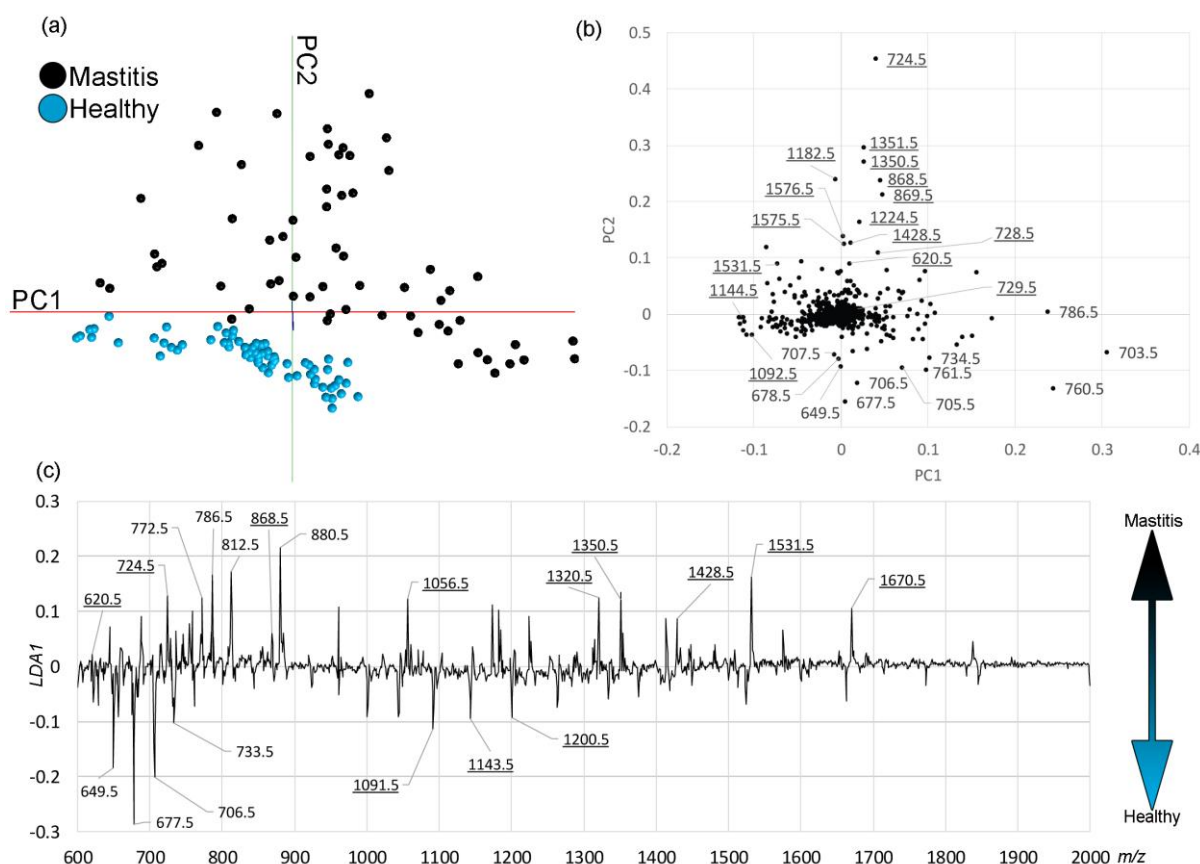
Figure 69: (a) Projection of the first two PCA dimensions shows separation between two classes. (b) Loading plot for PC1 and PC2. Underlined labels are for ions with z>1. (c) Mass loading plot for the LDA model. Peak intensity indicates ion influence on classification.

The model classification performance was assessed by the built in 'leave 20% out' cross validation in AMX. Outliers were determined by a standard deviation of 5. Sensitivity of 0.90 (56/62 mastitis spectra) and specificity of 1 (73/73 healthy spectra) were obtained, with a total correct classification rate of 95.56%. Importantly, no false positives were reported possibly due to the 'closeness' of the healthy samples in the classification space. The apparently greater level of variability for mastitis samples was tentatively attributed to the individual animal's response to infection. Attempts to establish trends based on cow age, pathogen, and SCC did not reveal much information. With this level of classification accuracy, it is worth considering how this method can be adopted for a working dairy. The model was built without replicate samples and these could be included should further confidence in the classification be required. On the other hand, this would increase the workload throughout the method to the point where it may be impractical for a large herd. An alternative might be to accept the mastitis classification after two successive classifications as such. Milking may occur two or three times a day by a semi-automated process for large herds, so mastitis classification would still be obtainable within a day.

From the loading plots in Figure 69b, multiply charged ions in bins *m/z* 724.5, *m/z* 868.5, *m/z* 869.5, *m/z* 1351.5, *m/z* 1350.5 and *m/z* 1182.5 were more abundant in mastitis samples than healthy, while for multiply charged ions in bins *m/z* 1091.5, *m/z* 1143.5 and *m/z* 1200.5 the opposite was true. Some singly charged ions corresponding to phospholipids also exhibited higher relative intensity in one class or another. The explanation of this trend in the multiply charged ions was found in the identification of these ions by spectrum deconvolution and MS/MS analysis (Section 3.3.2).

As with previous sections, AMX Recognition was demonstrated for potential online, real-time classification. Using the healthy/mastitis PCA/LDA model (Figure 69) and AMX Recognition post-processing mode, example datafiles were analysed by the software and are shown in Figure 70. Since the software is still in development, the settings for correct performance were tested to give correct sample selection. For this functionality demonstration the settings in Table 14 were used.
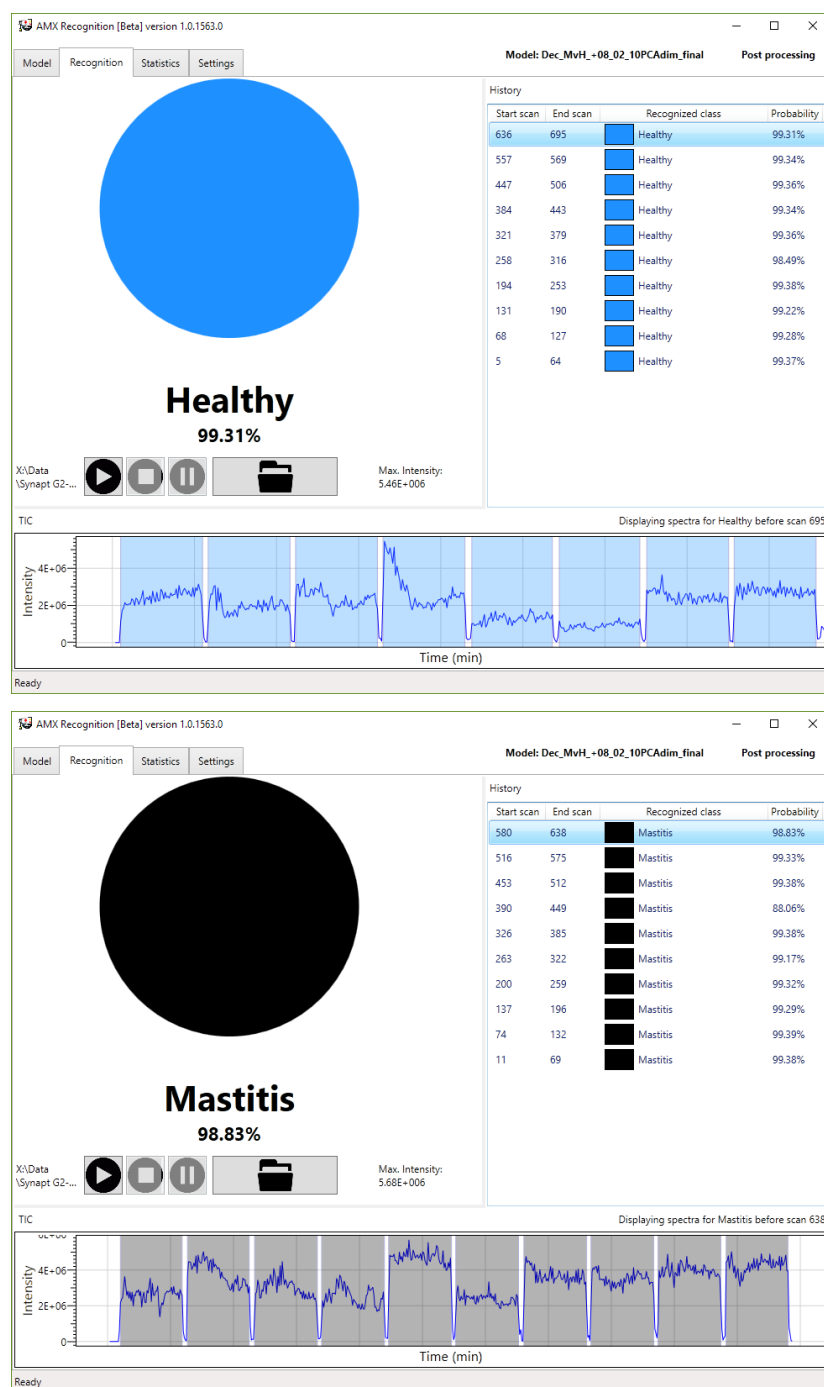
Figure 70: AMX Recognition windows showing the ability to classify healthy and mastitis samples in post-processing using the PCA/LDA model. This is transferable to online classification when samples are analysed in real time.

Table 14: Parameters for AMX Recognition demonstration with healthy and mastitis samples

Parameter	Value
Apply Lockmass correction	Yes
Apply background subtraction	Yes
Threshold	10000 counts
Lockmass	*m/z* 760.5851 (All classes)
Outlier type	Standard deviation
Standard deviation multiplier	5
Intensity limit	$1E^6$ counts
Scans per spectrum	60
Good spectrum timeout	60 s
Probability threshold	Global (0%)

3.3.2 Mastitis model compound Identification

Classification by multivariate analysis did not require the identity of compounds to be determined. However, ion identification was sought to better understand the underlying biological processes behind the classification. With this knowledge, the ability to build targeted methods and predicative models is possible enabling improvements to sample classification. Multiply charged ions detected by liquid AP-MALDI MS of milk extracts are displayed in Table 15. Some of these were of low intensity and only evident after ion mobility filtering for charge states greater than 1 (as seen in Figure 65). In the mastitis samples, peptide ions were more plentiful than in the healthy milk. This is arguably evidence of responses to infection, which result in the degradation or reduced synthesis of milk proteins.[219] These ions were detected with charge states in the region of 2^+ to 8^+.

Table 15: High charge state liquid AP-MALDI ions from healthy and mastitis samples.

Contributing ions (most abundant *m/z*)	Deconvoluted mass (Da)	Healthy	Mastitis
477.28^{4+}, 636.37^{3+}, 954.55^{2+}	1905		x
583.34^{5+}, 728.67^{4+}, 971.56^{3+}	2910		x
664.87^{6+}, 797.45^{5+}	3981	x	x
620.94^{7+}, 724.25^{6+}, 868.90^{5+}, 1085.88^{4+}, 1447.82^{3+}	4340		x
644.39^{8+}, 736.15^{7+}	5145		x
1046.82^{5+}, 1307.75^{4+}	5226		x
1224.30^{7+}, 1428.32^{6+}	8561		x
1051.01^{9+}, 1182.17^{8+}, 1351.02^{7+}, 1576.37^{6+}	9450		x
812.78^{13+}, 880.52^{12+}, 960.55^{11+}	10551		x
1183.13^{10+}, 1314.81^{9+}, 1478.88^{8+}, 1689.94^{7+}	11820	x	x
1413.74^{13+}, 1531.12^{12+}, 1670.31^{11+}	18360	x	x
1046.82^{23+}, 1092.96^{22+}, 1145.26^{21+}, 1202.24^{20+}, 1265.67^{19+}	24020	x	

Using this information, a targeted PCA/LDA model was built from the same data as the original model, over a narrow *m/z* range (*m/z* 720 – *m/z* 740). This included three peptide ions (*m/z* 728.8.67, 724.25 and 736.15) and multiple phospholipid ions. All other settings were as before. This narrow *m/z* range resulted in an increase in classification accuracy to 98.52% when assessed by 'leave 20% out' cross validation, while maintaining a specificity of 100%. Interestingly, the two samples that still failed classification were from two cows on the same date. This date (02 March 2016) was the middle sample of three consecutive mastitis samples from the two cows. The cows had received treatment after the first sampling and required further treatment after the third sampling. This suggests that the mastitis symptoms (clots) had improved by the second sampling, in which case the markers for the multivariate model may also have been affected. This relationship will be assessed by the continuation of this project over the next few years. Deconvoluted zero-charge mass spectra also show the shift to lower mass analytes in the mastitis samples (Figure 71). The dramatic difference in the relative intensity of the 24,020 Da signal suggests degradation of caseins (including β-casein and α-s$_1$-casein) was significant. The peak at 28,420 Da was tentatively assigned to α-lactalbumin (dimer), which also saw a reduction in intensity in mastitis samples. However, the signal at approx. 18300 Da was present in both sample classes, suggesting β-lactoglobulin is less affected by mastitis infection.
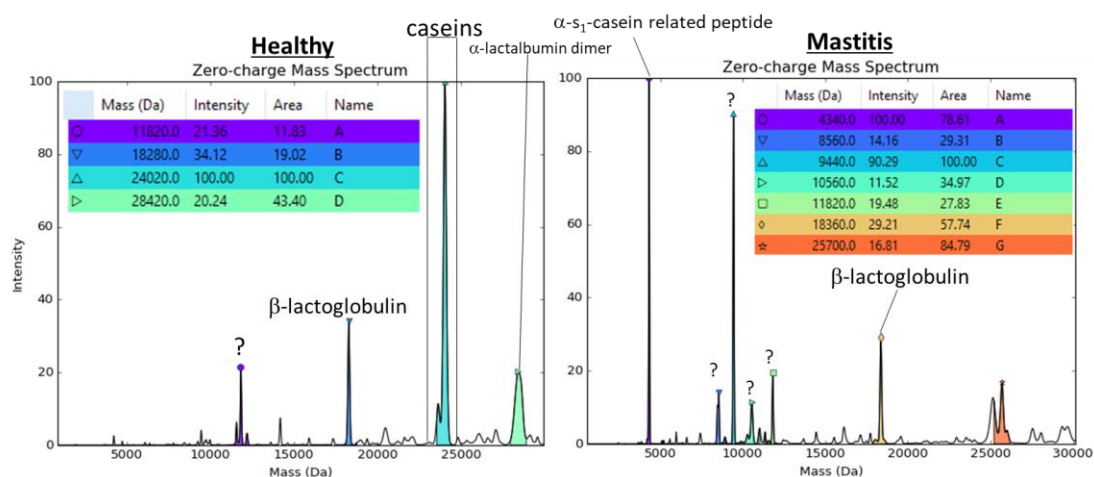
Figure 71: Example of deconvoluted mass spectra calculated from multiply charged species detected from representative healthy and mastitis samples. '?' indicates degradation products from abundant milk proteins.

Some sufficiently intense multiply charged ions were subject to CID MS/MS. Shown in Figure 72a and b, two peptides were identified as N-terminal fragments of bovine α-s$_1$-casein (accession number: P02662, without the signal peptide); (a) the 4340 Da peptide ([M+5H]$^{5+}$ *m/z* 869) was amino acids R$_1$-V$_{37}$ and (b) the 2910 Da peptide ([M+4H]$^{4+}$ *m/z* 729) was amino acids R$_1$-F$_{24}$. Cleavage of the protein occurred between V$_{37}$/N$_{38}$ and F$_{24}$/V$_{25}$ respectively. Neither of these peptides are the result of plasmin hydrolysis, which is an endogenous protease in milk that has specificity to the carboxyl side of Leucine (L) and Lysine (K).[250, 251] Plasmin has been suggested for use in mastitis detection previously.[252] The cleavage resulting in the 2910 Da peptide can be attributed to the action of Cathepsin B or D.[253] The peptide of 4340 Da was reported previously, along with a peptide of approx. 3984 Da, which is the result of plasmin cleavage between residues K$_{34}$ and E$_{35}$, that potentially could be the identity of the peptide of 3981 Da. This peptide was not fragmented here but the mass difference could be due to low signal intensity and accuracy of the deconvolution. This peptide occurred in both healthy and mastitis samples here, detected by liquid AP-MALDI MS (see Table 15, Figure 65). Previously, it was hypothesised that the 4340 Da peptide is formed by another, unidentified enzyme.[253, 254] Both peptides (4340 Da and 2910 Da) contain the sequence of a known antimicrobial peptide (isracidin, residues 1-23). Ions of the 4340 Da peptide were especially powerful for the mastitis classification.

β-casein and α-s$_1$-casein-derived peptides previously suggested as mastitis biomarkers from capillary electrophoresis (CE-) MS studies were not detected by liquid AP-MALDI MS, and conversely, peptides and protein fragments detected by AP-MALDI MS were not reported in the CE-MS work.[255] This is perhaps because AP-MALDI detects the most abundant molecules since there is no chromatographic separation, whereas CE-MS features CE separation prior to MS, enabling greater sensitivity thanks to reduced competition between ions.
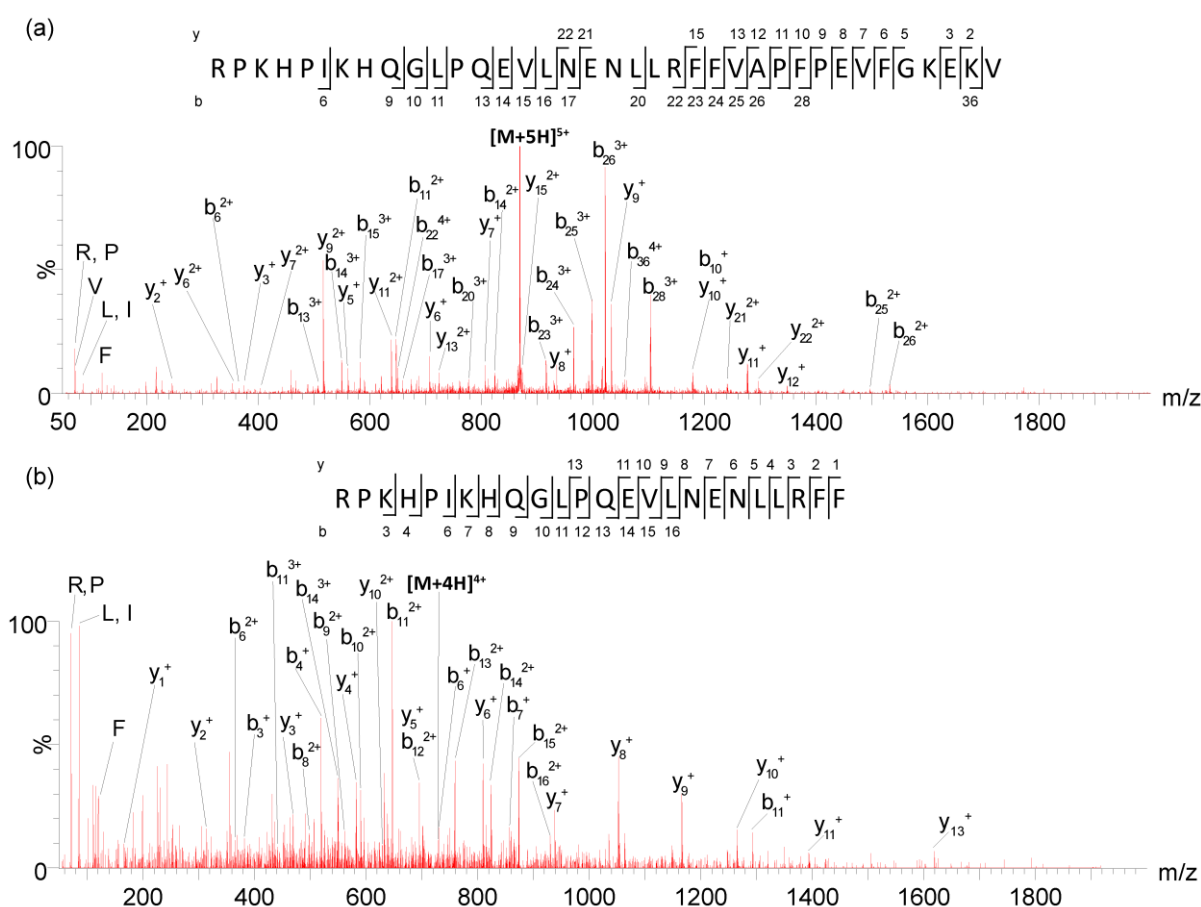
Figure 72: Liquid AP-MALDI-CID MS/MS spectrum of (a) [M+5H]$^{5+}$ (m/z 869) and (b) [M+4H]$^{4+}$ (m/z 729) annotated to highlight y, b and immonium ions. The proposed sequences are displayed above the mass spectrum and represent N-terminal fragments (amino acids R$_1$-V$_{37}$ and R$_1$-F$_{24}$ respectively) derived from bovine α-s$_1$-casein (P02662).

CID MS/MS of the [M+7H]$^{7+}$ precursor ion (*m/z* 1351, Figure 73) of the 9450 Da protein suggests identification as the C-terminal amino acids T$_{126}$-V$_{209}$ of β-casein. In addition, CID MS/MS of the [M+8H]$^{8+}$ precursor ion (*m/z* 1479, Figure 74) suggests the 11,820 Da protein is similar, consisting of β-casein (P02666) C-terminal amino acids H$_{106}$-V$_{209}$. The first β-casein fragment features an L residue at the cleavage site (L$_{125}$/T$_{126}$) and the second fragment includes a K residue (K$_{105}$/H$_{106}$), which may indicate plasmin-related activity is involved in their formation. With this information it can be suggested that the peptide of 8561 Da is the β-casein C-terminal amino acids H$_{134}$-V$_{209}$, although MS/MS data were not obtained. The MS/MS peak lists for the CID-MS/MS spectra in Figure 72-Figure 74 are included in Appendix 3. That these protein fragments are largely exclusive to mastitis samples raises the potential for them to be used as markers for more targeted tests.
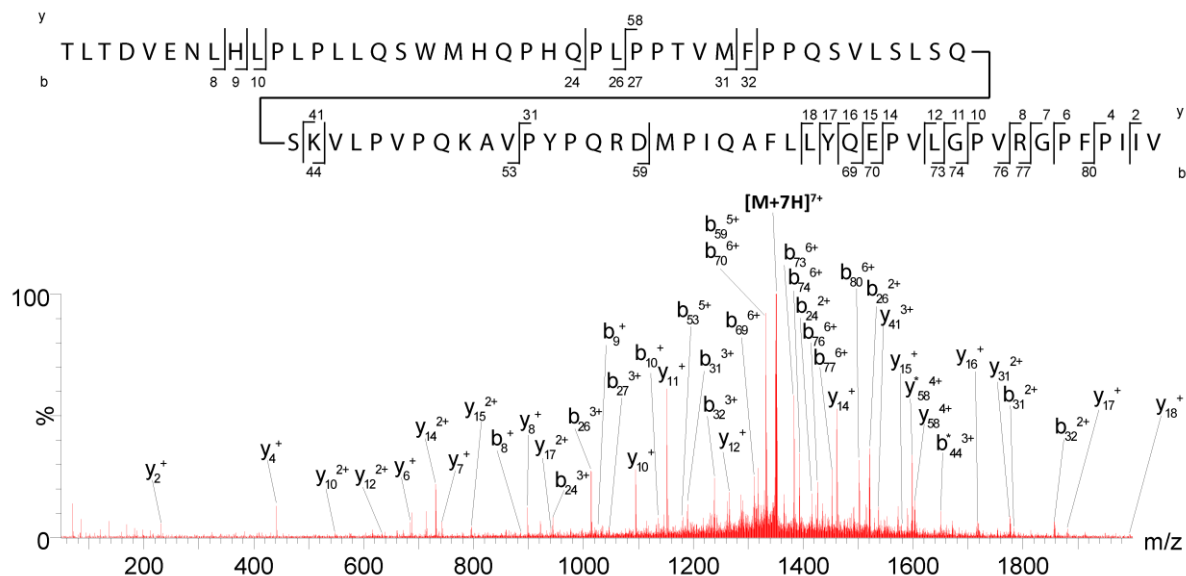
Figure 73: Liquid AP-MALDI-CID MS/MS spectrum of the [M+7H]⁷⁺ precursor ion at m/z 1351. The precursor ion can be attributed to a C-terminal fragment of β-casein (P02666), amino acid residues T₁₂₆-V₂₀₉.
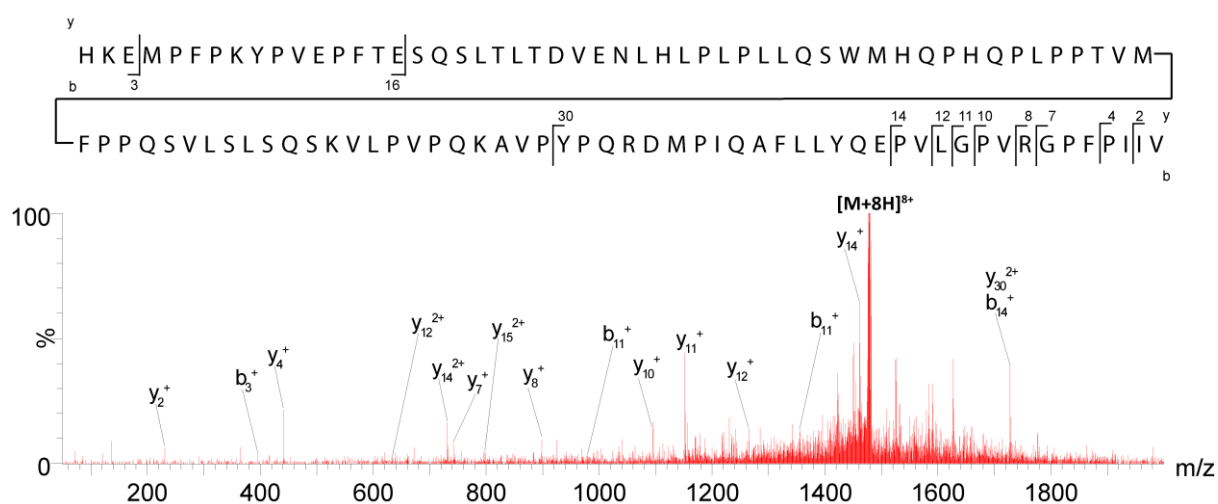
Figure 74: Liquid AP-MALDI-CID MS/MS spectrum or the [M+8H]⁸⁺ precursor ion at m/z 1479. The precursor ion can be attributed to a C-terminal fragment of β-casein, amino acid residues H₁₀₆-V₂₀₉.

3.3.3 Liquid AP-UV-MALDI profiling without a UV-chromophore

As observed by Koch *et al.* with liquid samples, laser desorption/ionisation (LDI) at AP using a laser with a wavelength of 337 nm can also be performed without the addition of a traditional UV matrix chromophore.[89] This simplifies the sample preparation through the omission of preparation of a matrix compound. Signals related to matrix chemical background are omitted, although there is still a contribution from other sources, for example the sample plate. Doubly charged metal-phospholipid ions of the type discussed in section 3.1 can be generated from liquid droplets without a chromophore. Figure 75 shows how the intensity of ions in $BaCl_2$-doped (500 pmol), PC (16:0/18:0) (1 pmol) liquid samples differs between chromophore and chromophore-lacking droplets. No acid (TFA, FA, etc) was added to either sample. With a 2,5-DHB LSM (LSM **A**), the $[L+H]^+$ ion peak was >100x as intense ($4.55e^6$ counts versus $1.38e^4$ counts) than the peak for the $[L+Ba]^{2+}$ ion. In the analysis with LSM **G**, the $[L+H]^+$ peak intensity was only 2.26e3, three orders of magnitude lower than with LSM **A**. However, the $[L+Ba]^{2+}$ peak intensity was similar to that recorded with LSM **A**. This suggests that the protonated ion formation may largely rely on processes involving the matrix e.g. proton transfer, whereas ionisation with the Ba^{2+} ion is independent of the matrix. This could indicate these ions are preformed in the droplet and are desolvated in the heated ion inlet upon ablation from the sample droplet.
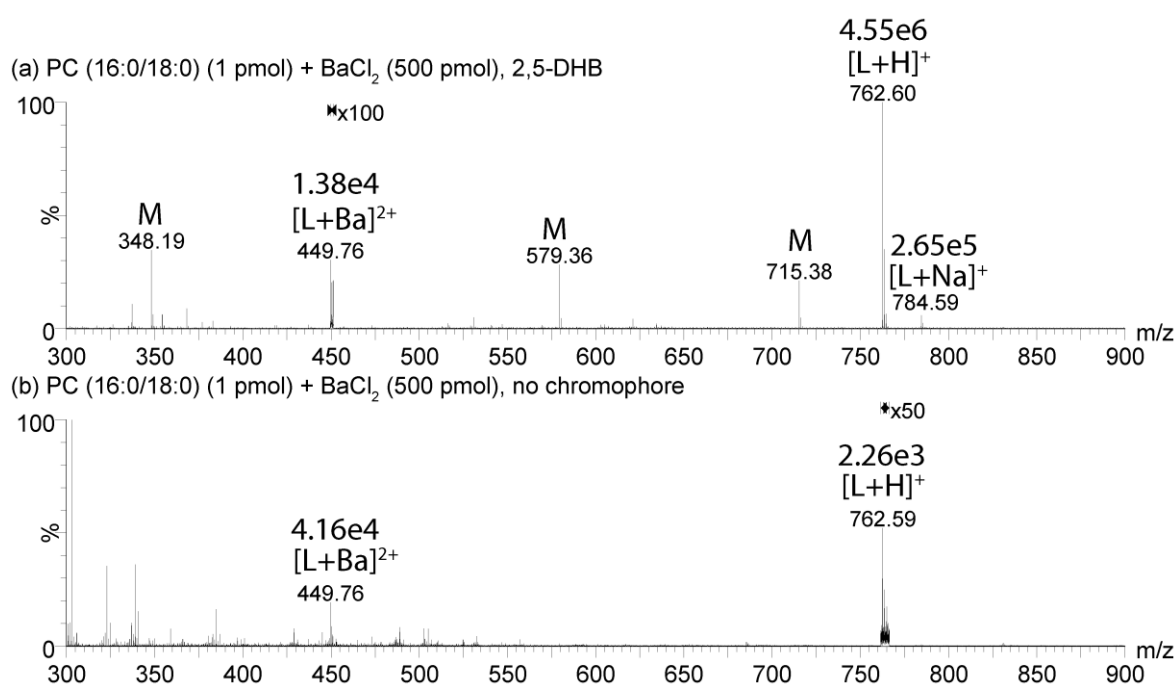
*Figure 75: BaCl₂-doped liquid samples of PC (16:0/18:0) analysed by (a) AP-MALDI with LSM **A** and (b) AP-LDI with LSM **G** for 1 minute. The signal intensity (indicated above the ion label) for the protonated analyte ion is three orders of magnitude lower without a chromophore, suggesting protonated ion formation benefits from the inclusion of a MALDI matrix. The doubly charged barium adduct ion is very similar intensity in both samples. Known matrix-related ions are marked with 'M'.*

Multiply charged ions are also generated for peptides and proteins, as with AP-MALDI LSMs and similarly to SAII, where multiply charged ions are also predominantly produced.[256] Figure 76

demonstrates the potential for rapid profiling with a mastitis milk sample. The same number of laser shots (300) for matrices **A** (Figure 46a) and **G** (Figure 46c) produce drastically different spectra. However, a MALDI sample with LSM **A** (Figure 76b) analysed for the same time period (30 seconds) but higher pulse repetition rate, and thus 900 laser shots, resulted in an improved spectrum compared to Figure 76c. However, analysis without a chromophore was more difficult to perform because of the requirement to focus on the droplet edge, and in general it appears the signal intensity per volume of sample consumed is higher.

The protein fragment ion $[M+7H]^{7+}$ at *m/z* 1531 is most resolved with the LSM **G**, allowing straightforward determination of its charge state. The ion is comparably intense in spectra (Figure 76b) and (Figure 76c). Peptide (e.g. *m/z* 724.24) and phospholipid (e.g. *m/z* 703.57) ion intensities were also similar over the 30 second period.
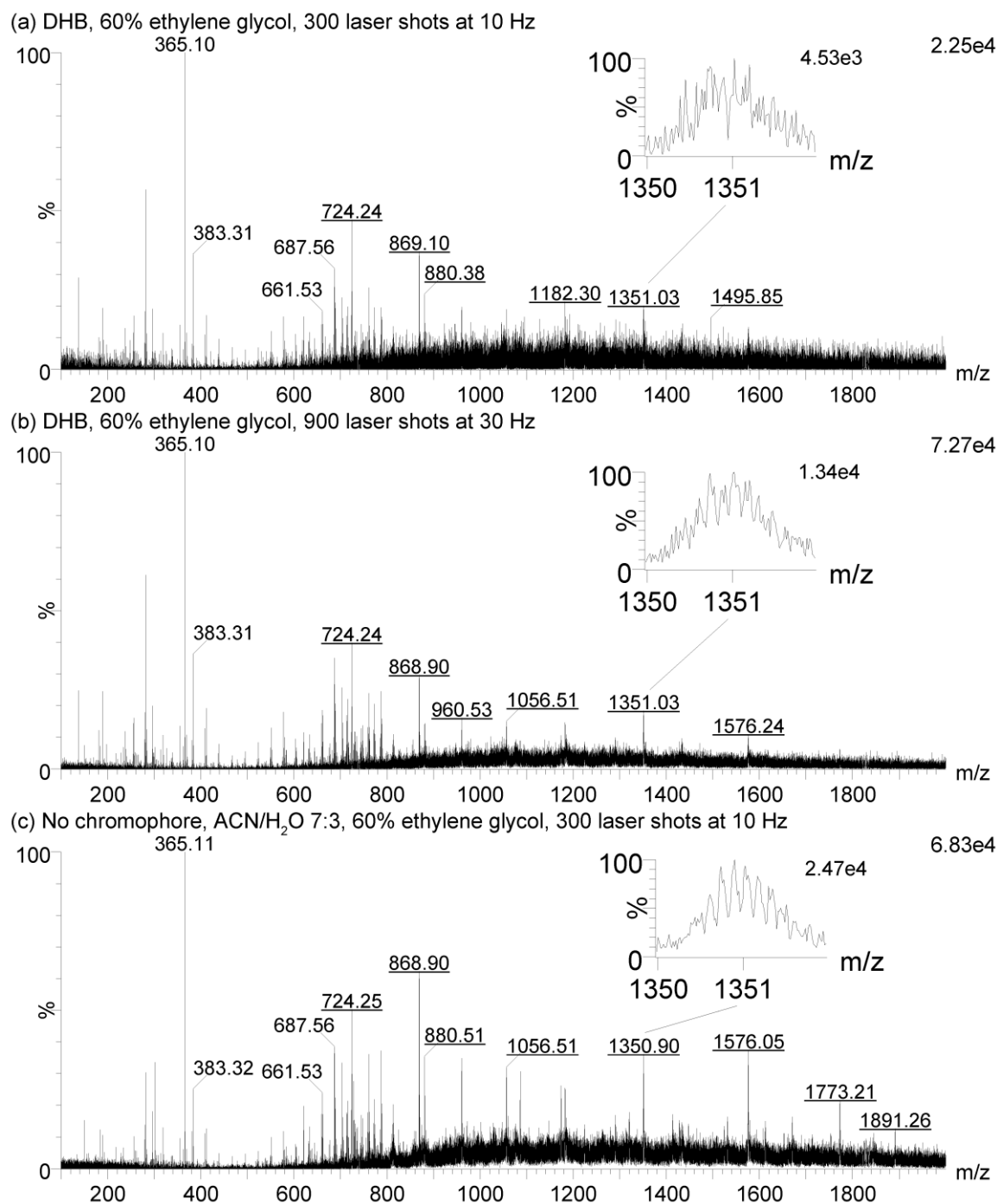
Figure 76: Liquid AP-(MA)LDI mass spectra of mastitis sample 17M09 analysed with matrix A (a) 300 laser shots at 10 Hz (b) 900 laser shots at 30 Hz and (c) matrix G with 300 laser shots at 10 Hz. Underlined labels denote examples of multiply charged ions. The [M+7H]$^{7+}$ peak at m/z 1531 is most resolved and intense without chromophore.

An obstacle to analysis without a chromophore being adopted for automated analysis is that LSM **G** requires the laser to be focused on the edge of the droplet for ions to be detected, whereas LSMs containing chromophores e.g. LSM **A** do not. A computerised system which can identify the sample edge and position the laser accordingly would be necessary for this to be applicable to high throughput analysis with the current MALDI target. Optionally, a new target geometry could

be developed that simplified producing ions from these samples, for example by channelling the sample to a specific position on the sample plate surface. Some preliminary work related to this idea has been published by Koch *et al.*.[89] The chromophore-lacking samples were limited to analysis at lower laser pulse rates (< 20 Hz). This is possibly because the 'healing' ability of the droplet is compromised by being ablated at an edge, in part because the droplet is shallower. Finally, the sample consumption compared to chromophore-containing matrices appeared greater; droplets simply did not last for as long as with other LSMs at the same laser energy and pulse rate. This might explain the somewhat better signal intensity over a short period ,although signal per ablated volume is probably worse, but comes at the expense of not being able to return to a previously analysed sample after storage, for example. Material ejected from the chromophore-lacking samples was visually evident when the laser was focused on the droplet edge. It is probable that the shorter sample lifetime was observed partly due to this material being lost from the sample well. The long lifetime of a chromophore-containing LSM is still likely to yield best overall sensitivity given enough time to consume the whole droplet.
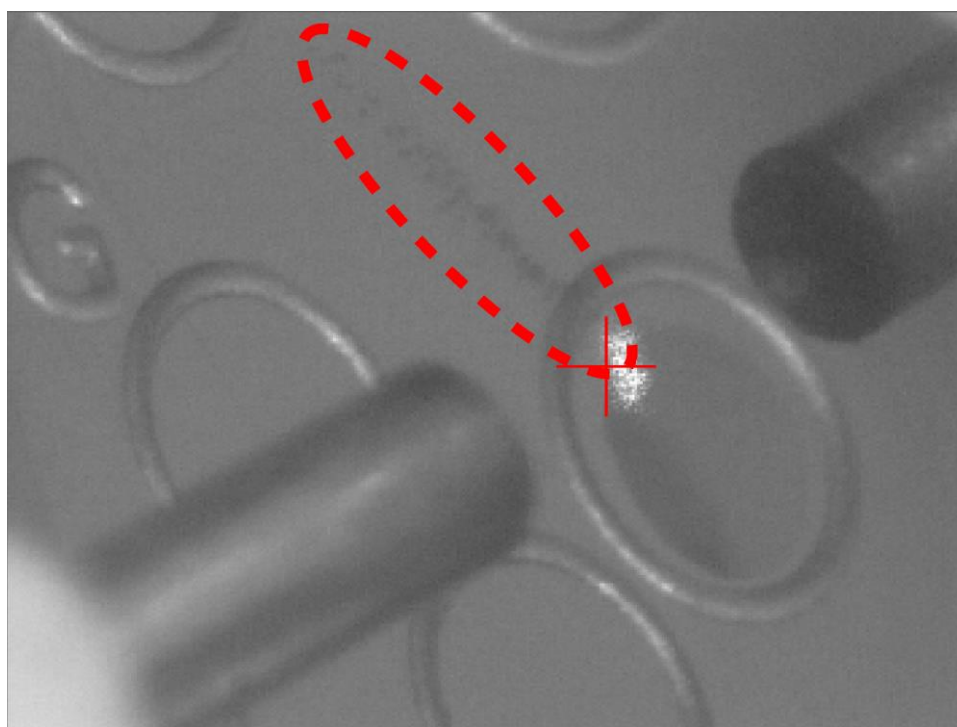
Figure 77: Liquid material (red dashed ellipse) was ejected from chromophore-lacking samples when analysed by a laser focused (approx. red cross) at the droplet edge. This material was lost from the sample well and likely contributed to the sample's relatively short lifetime.

3.4   Preliminary work for β-lactam antibiotic resistance testing.

Antibiotics are administered to dairy cattle in the event of bacterial infection, for example clinical mastitis, being diagnosed. Growth promoting antibiotics are also used to speed up animal growth and increase their economic value, despite evidence to the contrary.[257] Some commonly used antibiotics are the β-lactams, including the penicillin family of antibiotics. These antibiotics and their metabolites must fall below concentrations specified by a regional regulatory body (e.g. European Union) before the milk from treated animals can begin to be sold again.[210] The levels of antibiotics in milk may be analysed by LC-MS/MS due to its sensitivity, specificity and the ability to perform absolute quantitation.[258] These are usually multiple reaction monitoring (MRM) methods performed on highly sensitive triple quadrupole MS, which target the analytes and their fragments to produce highly specific data.

Any ion source that does not benefit from chromatographically separating analytes prior to ionisation is unlikely to unseat LC-MS/MS from being the primary choice for quantitative analysis anytime soon. However, the speed with which an ion source such as MALDI can screen samples means there are other opportunities to explore. One such opportunity is the ability to determine the antimicrobial resistance of bacteria. Resistance is usually assessed after culturing the bacteria for a day or more. However, there is a drive to detect bacteria directly from samples by MALDI-TOF MS without culturing because of the time savings possible.[221]

In this section, the addition of β-lactams to milk was explored to provide markers for detecting β-lactam resistant bacteria. It is intended as a demonstration of a methodology and requires further validation using samples of known bacteria. β-lactam resistance is exhibited in bacteria that excrete the enzyme β-lactamase, which acts to hydrolyse the four-membered ring of β-lactams (Figure 78). The hypothesis was that, in the presence of the enzyme β-lactamase, the concentration of hydrolysis products would increase, and thus their peak intensity in a mass spectrum.
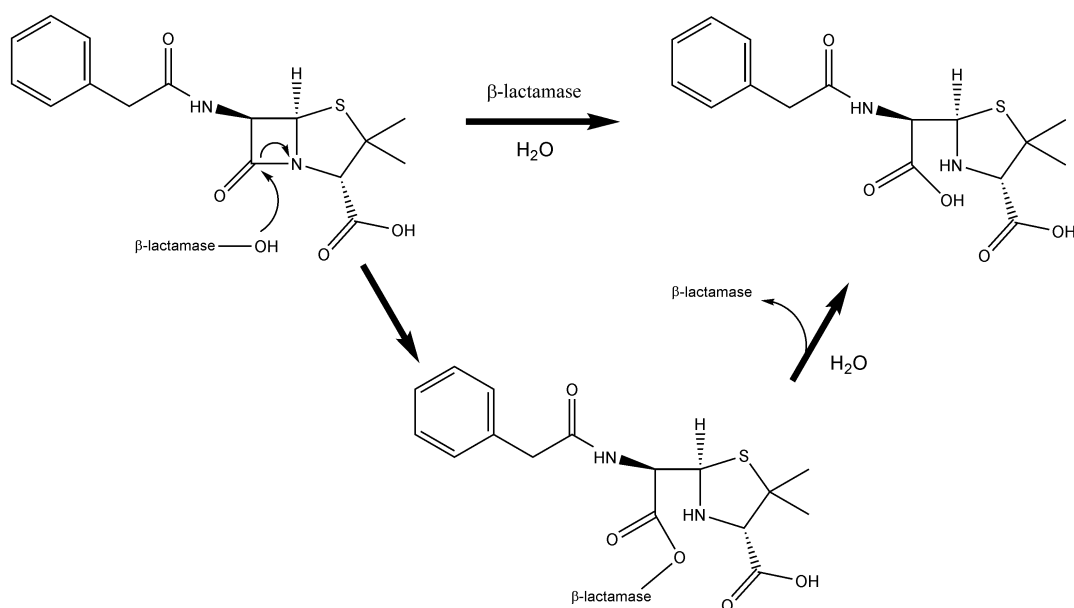
Figure 78: The mechanism for hydrolysis of penicillin G catalysed by β-lactamase. The product is 18 Da heavier. The opening of the four-membered ring removes the antibiotic property of the molecule.

As detailed in Section 2.8, a mixture of β-lactam antibiotics was added to aliquots of water and bovine whole milk and vortexed. To half of the samples from each matrix (water or milk), β-lactamase solution was added. All samples were subject to incubation at 30°C for two hours. They were then diluted with aqueous TCA solution to precipitate proteinaceous content and stop enzymatic processes, then centrifuged. The supernatant was analysed by liquid AP-MALDI MS with LSM **A**. Liquid AP-MALDI is tolerant to acids such as TFA and TCA, unlike ESI, so further clean-up was not performed.[259]

The resulting profile spectra were highly populated, including the protonated and sodiated adduct ions of the β-lactams and hydrolysis products and matrix ions from the MALDI sample components. Thus, it was decided that an MS/MS method was required to confidently detect these small molecules against the high chemical background signal present in the relevant *m/z* ranges, an example of which is shown in Figure 79. Even with the high resolving power of the TOF analyser, it was not possible to confirm the signal at *m/z* 348.10 was exclusively due to cefalexin. Ion mobility separation was not possible as the antibiotic ions overlapped with the background signal in this dimension, too.
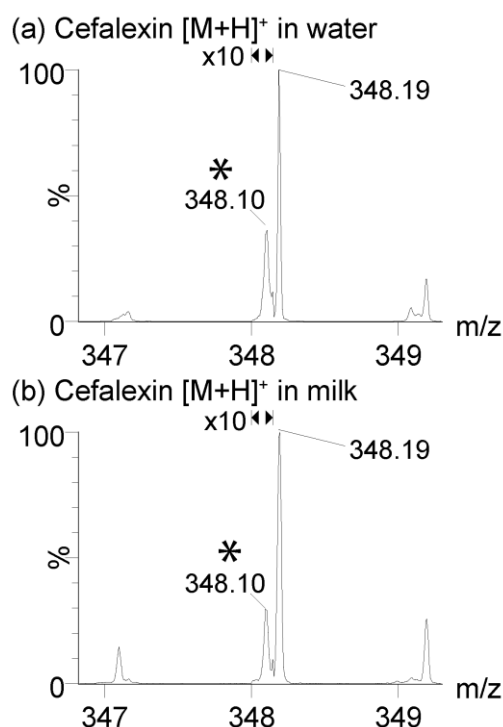
*Figure 79: The peak for protonated cefalexin in (a) water and (b) milk extracts as an example of interfering peaks. The analyte peak is marked with * and is magnified 10x. The interfering peak (m/z 348.19) only differs by m/z 0.09.*

Parallel reaction monitoring (PRM) with CID was used to monitor multiple transitions from precursor to product ions; the product ion signal was used to confirm detection and provide precursor ion abundance. On the Synapt, a targeted method like PRM, which is essentially a QTOF application of MRM, allowing many fragments to be monitored in parallel due to the TOF's non-scanning ion recording) makes use of the transfer cell ion optics to improve sensitivity for the transitions monitored.

First, standards of each antibiotic were subject to CID-MS/MS to determine which fragments to monitor. The antibiotic standards were also subject to hydrolysis by β-lactamase after which the hydrolysis products were also analysed by CID-MS/MS. By choosing fragments from experimental data, rather than fragments reported in the literature, the method was tuned for the specific instrumentation. The chosen monitored fragments are shown in Figure 80 for the non-hydrolysed antibiotics and Figure 81 for the antibiotic hydrolysis products. The structures of the fragment ions are suggested with respect to information in the literature and the METLIN *in silico* and experimental MS/MS spectral database.[260, 261]
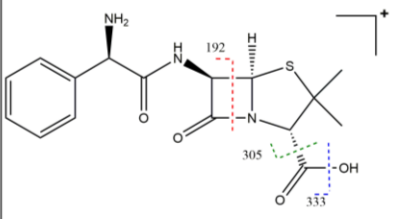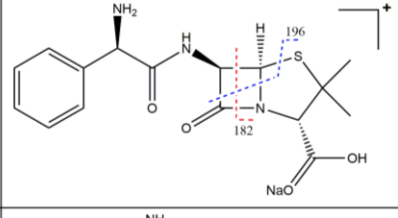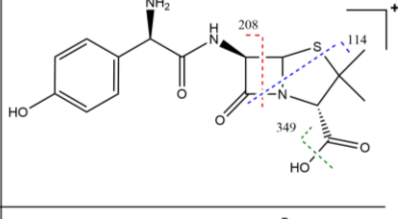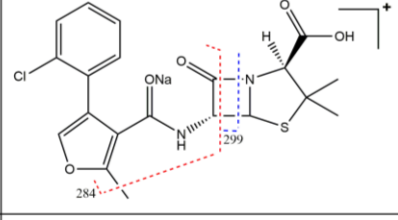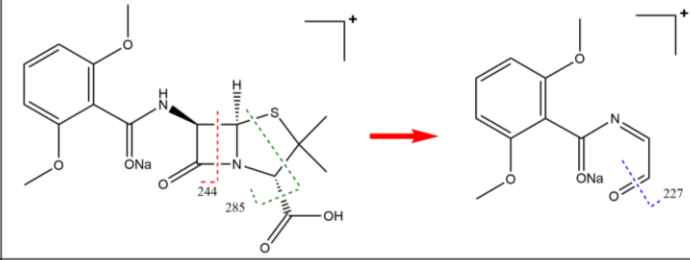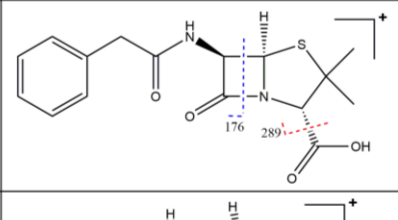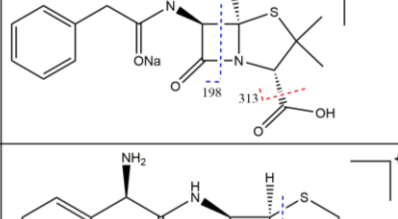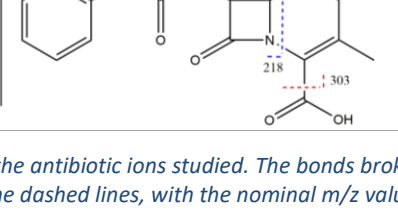
Figure 80: Structures of the antibiotic ions studied. The bonds broken by CID to produce the fragments monitored by PRM are suggested by the dashed lines, with the nominal m/z value of the fragments indicated.

Ion	Monitored Fragments
[Hydrolysed Amoxicillin + Na]+	
[Hydrolysed Ampicillin + Na]+	
[Hydrolysed Cloxacillin + Na]+	
[Hydrolysed Methicillin + Na]+	
[Hydrolysed Pencillin G + Na]+	
[Hydrolysed Cefalexin + Na]+	

Figure 81: Structures of the antibiotic hydrolysis products studied. The bonds broken by CID to produce fragments monitored by PRM are suggested by the dashed lines, with the nominal m/z value of the fragments indicated.

112

In general, sodium adduct ions were favoured as the antibiotics were supplied as sodium salts. Example CID-MS/MS spectra from the PRM method for the antibiotics and their hydrolysis products from a control and test sample are shown in Figure 82 and Figure 83, respectively. The antibiotics and hydrolysed antibiotics product ions were monitored for all samples. This was because the enzymatic hydrolysis did not convert 100% of the available antibiotic so there was always some signal in the hydrolysed samples. Similarly, other hydrolysis reactions meant the hydrolysis products could be present in the enzyme-lacking samples.
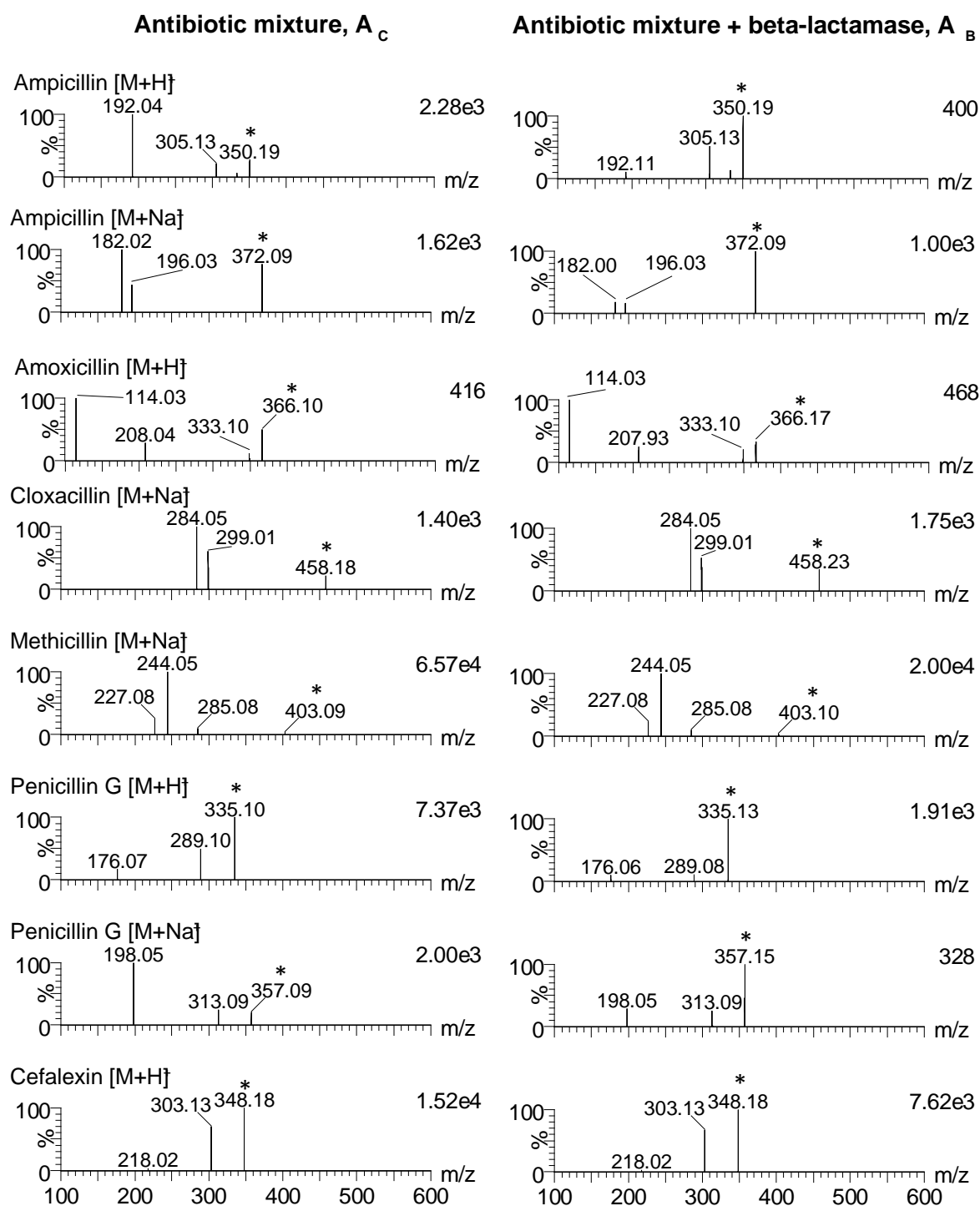
Figure 82: Liquid AP-MALDI-CID-MS/MS spectra from the PRM data of antibiotic ions in a control sample and beta-lactamase-containing test sample post-incubation at 30 °C for two hours. Precursors peaks are indicated by *.
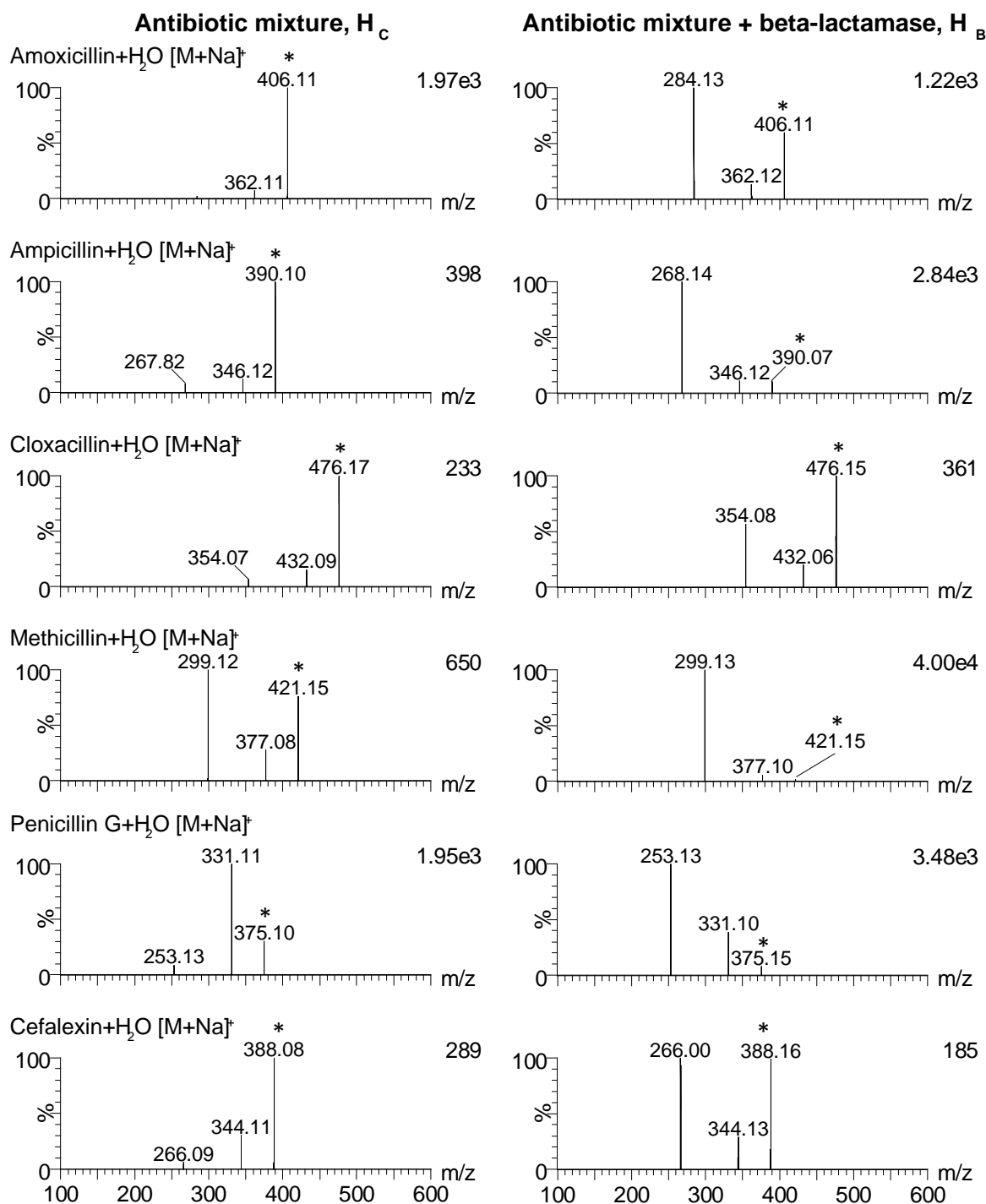
*Figure 83: Liquid AP-MALDI-CID-MS/MS spectra from the PRM data of the **hydrolysed** antibiotics in a control sample and beta-lactamase-containing test sample post-incubation at 30 °C for two hours. Precursors peaks are indicated by *.*

The workflow for comparison of the test sample to the control is shown in Figure 84. The ratio of the summed intensity of each antibiotic's CID product ions (A_C or A_B) was calculated versus the summed intensity of each hydrolysed antibiotic's CID product ions (H_C or H_B) for a control (C) and a test (B) sample. Where the ratio of 'A_B/H_B' for the test sample was less than in the control, 'A_C/H_C', this indicates that the signal for hydrolysed antibiotic product ions was greater.
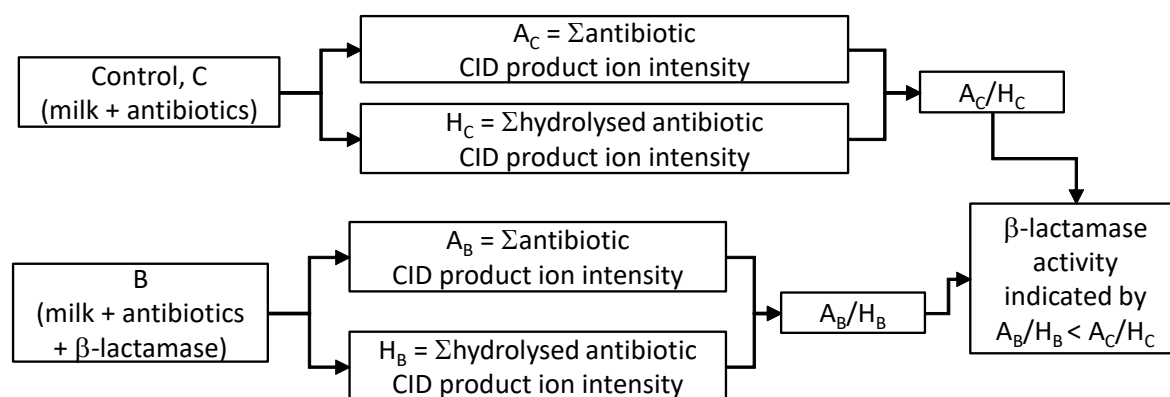
Figure 84: Workflow for comparison of a β-lactamase-containing sample to a control sample. CID products ions of antibiotic and hydrolysed antibiotic precursor ions were monitored in both samples since the hydrolysed antibiotics may be formed by processes other than β-lactamase activity.

For every antibiotic in water, and for all but amoxicillin in milk, the ratio was weighted in favour of the hydrolysis product in β-lactamase-containing samples (see Figure 85). The same change was not observed for the control samples, despite otherwise identical experimental conditions. This suggests that the enzyme was the main cause of hydrolysis of antibiotics in both water and milk. The relatively low signal intensity of the $[M+H]^+$ ion is a possible cause of the amoxicillin result.
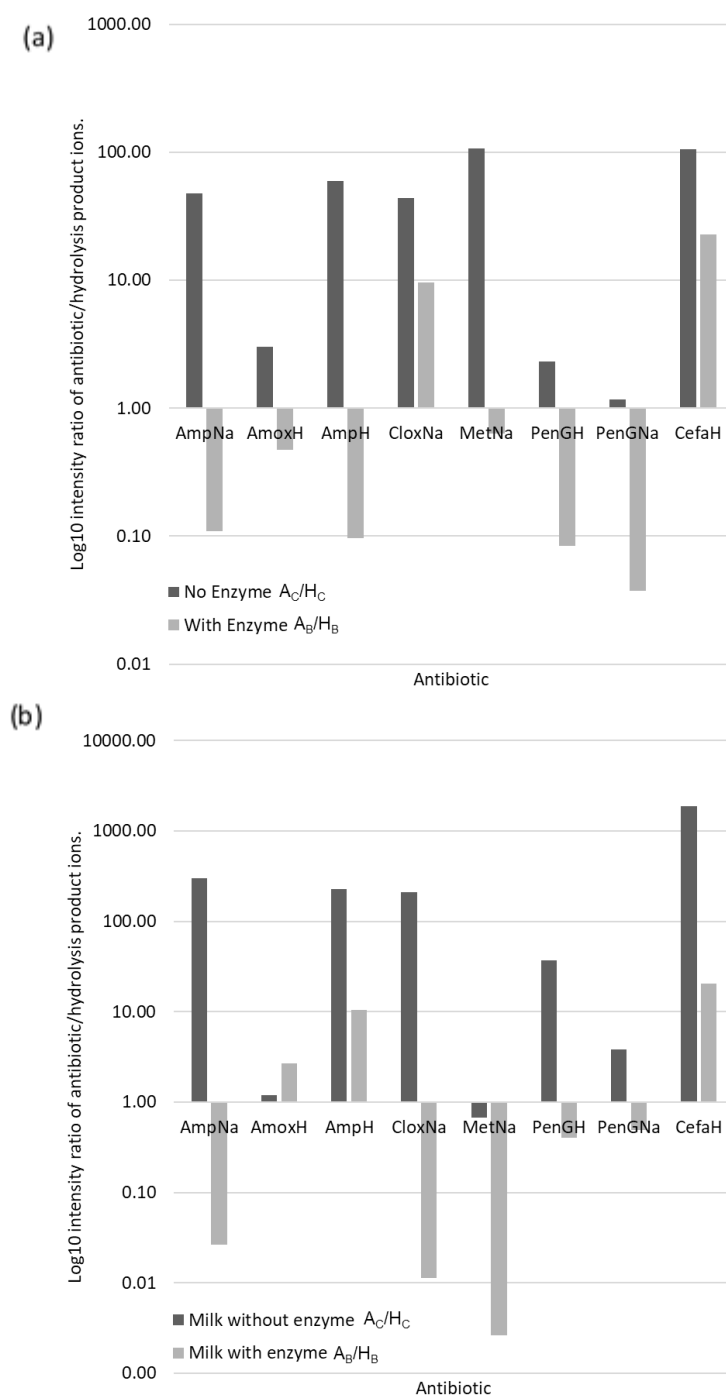
Figure 85: Bar charts showing the ratio of summed CID product ion intensity of antibiotics and their hydrolysis products for samples of (a) water and (b) bovine skimmed milk in a control sample and a test sample containing β-lactamase.

Generally, the method would benefit from improved sensitivity to provide confidence in detected ions and more rapid analysis. This could be achieved with a higher initial concentration of the antibiotic solution or multistep extraction methods such as dispersive SPE that are often used for small molecule extraction from biological matrices.[262, 263] Modification of the PRM method to expand the *m/z* window transmitted by the quadrupole may also improve signal, although would

117

result in transmission of more of the chemical background. Additionally, coupling the AP-MALDI ion source to the most sensitive MS/MS instrumentation available i.e. a triple quadrupole MS would likely yield improvements. The current method does not require a high-resolution MS, although useful for method development, so this change would not be disadvantaged by this change.

These initial results suggest that monitoring for β-lactam hydrolysis products may be useful for detecting resistance to β-lactam antibiotics, whilst avoiding the need to culture and classify the causative bacteria. Potentially, this is a reduction from days to hours for performing analysis. Liquid AP-MALDI coupled to an MS is potentially a sensitive, high throughput and semi-quantitative method for these small molecules without the need for chromatography. An unknown for further development of this technique, which should be answered as a next step, is whether the natural abundance of β-lactamase in milk from resistant bacteria would be high enough to produce a noticeable change in analyte ratio. The digestion period may require extension (thus giving greater time for enzymatic reactions) for changes to be noticeable, as the natural concentrations could be much lower than the 1.5 – 3.0 units added to the samples here. A milk sample set collected from cows with mastitis was bacteriologically characterised but did not provide enough resistant samples for a thorough analysis. As such, this remains an open question.

4 Conclusions

4.1 Determination of structural information from metal-cationised phospholipids

Structural information from PCs and SMs was obtained through liquid AP-MALDI MS generation of $[L+M]^{2+}$ ions and subsequent CID or ETD-MS/MS. Barium was found to be the metal ion which produced the greatest intensity of $[L+M]^{2+}$ signal. CID and ETD produced complementary fragment ions. The pseudo CID-MS^3 method allowed detailed structural information of lipid acyl chains to be determined through charge-remote fragmentation product ions. The methodology was simple, involving the addition of small amounts of inexpensive $BaCl_2$ (or other divalent metal salts) to liquid MALDI samples. This method complements other MS methods for lipid analysis but has been developed for compatibility with AP-MALDI as the focus rather than ESI. There is the potential to incorporate previously reported online and offline-coupled LC-MALDI for fractionation of complex samples.[259, 264]

4.2 High-throughput milk speciation

AP-MALDI-QTOF MS with liquid support matrices was demonstrated to be a flexible analytical technique for profiling complex biological samples. It was demonstrated that short analyte extractions were hugely beneficial for detecting lipids, peptides and proteins, rather than analysing directly from the crude sample. Specifically, phospholipids were readily detectable by AP-MALDI MS after a short hexane/isopropanol-based extraction. Multivariate models built from phospholipid MS profiles were suitable for classifying milk groups. The combination of data from both the positive and negative ion modes was confirmed to present the opportunity for higher classification accuracy, to as high as 98.4% with combined positive and negative ion mode data. The classification of binary milk mixtures with multivariate analysis was suggested to be more robust than classification based on the relative intensity of two 'marker' ions.

Detecting additional biological molecule classes, such as peptides and proteins, as multiply charged ions was demonstrated to contribute to the classification specificity. Improvement was found in the positive ion mode, allowing better discrimination between very similar sample classes i.e. pasteurised and unpasteurised ovine milk, and skimmed and whole bovine milk. Similar highly-charged ions were not detected in the negative ion mode. Many ion sources used for rapid profiling do not ionise molecules this large, an advantage of liquid AP-MALDI that can be used for additional food profiling and exploited for new applications in the future.

4.3 Detection of mastitis in dairy cows by liquid AP-MALDI MS of sampled milk

The combination of rapid AP-MALDI-QTOF MS analysis of liquid samples combined with multivariate analysis was demonstrated to provide the ability to classify samples from two milk sample groups, 'Healthy' and 'Mastitis'. Only short sample preparation and analysis time (1 minute) was necessary to achieve MS profiles suitable for multivariate analysis. The detection of multiply charged ions, including peptides and proteins was key to achieving a classification accuracy of 95.56% and sensitivity of 0.90, assessed by 'leave 20% out' cross validation. No false positives were reported. Furthermore, identification of some peptides and proteins in mastitis samples suggested that while plasmin proteolysis was partly responsible for casein proteins degradation, an unknown process (enzymatic or otherwise) produced some additional abundant peptides.

With this understanding gained from clinical mastitis cases in this study, there is potential for these peptides to be used as markers for targeted subclinical mastitis diagnosis. A PCA/LDA model built with a targeted range m/z 720-740, which included three peptide ions, achieved an increased classification accuracy of 98.52% and retained 100% specificity when assessed by 'leave 20% out' cross validation.

4.3.1 Profiling without a chromophore

The potential for analysis of milk samples without using a matrix chromophore was demonstrated. It was necessary to focus the laser at the edge of the droplet, like the sodium iodide ionisation described in Section 2. Peak intensity for highly charged ions appeared improved versus spectra acquired with a matrix chromophore on a short timescale. However, the chromophore-lacking samples were limited to analysis with approx. 10 Hz laser pulse frequency (λ=337 nm), possibly due to a greater ablated volume, which could require more time for droplet healing. In addition, their lifetime was much shorter than for chromophore-containing LSMs, partly owing to the ejection of droplet material across the sample plate. Chromophore-lacking samples are not yet suitable for AP-MALDI high-throughput experiments, but continued development of sample preparation and the ion source may make them viable in the future.

4.4 Antibiotic resistance testing

As an alternative approach to mastitis testing, indirect detection of the presence of β-lactamase-producing bacteria in milk could inform on the health of the cow. Samples spiked with an antibiotic mixture and β-lactamase solution were analysed by AP-MALDI-PRM MS to determine monoisotopic peak intensity ratio differences between the antibiotics and their hydrolysis products. While this has yet to be demonstrated with real samples, the proof-of-principle is in

place; samples containing the enzyme generally resulted in a greater abundance of antibiotic hydrolysis products for experiments performed in water and milk. With further development, this experiment could be used as an indicator of bacterial resistance to β-lactam antibiotics without the need for bacterial culture.

4.5   Final conclusions

Liquid AP-MALDI MS can play a significant role in the mass profiling field as the technology develops, and this thesis demonstrates some of the potential methods and applications. MALDI MS is already used for classifying bacteria in clinical environments, and its potential for food authentication has been demonstrated by research groups worldwide, as was shown in Table 1. The variety of other ambient ionisation sources within the field should mean rapid developments continue. The advantages of liquid AP-MALDI, such as dependable ion signal and multiply charged ion generation, have been demonstrated for high-throughput applications and biological molecule characterisation and alleviate some of the problems associated with traditional MALDI sample preparation. This makes it a compelling choice for tackling complex biological challenges, where other ion sources may be more limited, for example in their ability to ionise only a small selection of molecules. In addition, the option to add simple sample additives to modify the ions signals that are obtained from liquid samples provides AP-MALDI with a flexible toolset for molecular examination.

The application of liquid AP-MALDI to the detection of bovine mastitis is particularly promising since the scale of the dairy industry requires millions of samples to be tested a year.[265] Assays that maintain the high-throughput of current methods but with improved sensitivity and specificity are necessary to reduce economic losses and unnecessary antibiotic use. As such, ubiquitous, cheap but inaccurate tests, like the California mastitis test, are set to be replaced. Both the specificity and sensitivity of the liquid AP-MALDI MS test classification were higher compared to the existing testing methodologies. MALDI MS profiling for mastitis bacteria classification from animals is already established in a UK, which is an excellent indication of the potential to adopt more MS-based techniques for routine testing. Targeted assays focusing on specific biomolecular markers for diseases are likely to be the future of many medical diagnoses, for humans too. The targeting of a narrow *m/z* range for mastitis detection here has demonstrated excellent sensitivity and specificity. Rapid mass spectrometry assays of human milk have already been initiated for REIMS.[266] Routine testing of human breast milk with AP-MALDI within hospitals could improve detection of mastitis in new mothers, and the lessons learned from the cows' milk study would provide valuable insight.

5 Future Work

There are a wide range of developments that can be made to the instrumentation and methodologies employed in this research. This section discusses and suggests general modifications to the ion source and sample preparation methods, then focuses on each of the results sections for more specific improvements.

5.1 General AP-MALDI development

5.1.1 Sample preparation automation

The sample preparation procedures discussed chapters were kept simple to limit the time between crude sample collection and analysis. Another benefit is that these methods are amenable to automation using existing proteomics and MALDI-focused robotics. This could include, but is not limited to, addition of extraction solutions to samples and spotting of MALDI samples onto the sample plate. The latter has been previously demonstrated with MALDI sample preparation robots.[267] This will be especially important for future investigations involving larger cohorts of sampled organisms.

5.1.2 Source and sample modifications

Further development of the heated inlet may yield greater overall sensitivity and improved multiply charged ion production. There are many options here that could be investigated in a systematic study. For example, the inlet length could be varied; the inside of the inlet could be manufactured to provide increased surface area; multiple tubes could form the inlet, increasing the capacity. The gas could also be heated independently of the capillary. Gases other than nitrogen could be used as the counter-flow gas or modifiers, such as isopropanol, could be introduced as a vapour into the gas flow. This may affect, and possibly enhance ionisation as with the REIMS source.[268] In this context, employing a heated impactor as in REIMS in addition to the heated inlet may have sensitivity benefits, however the temperature of the impactor may need to be tuned for best ionisation performance for peptides and proteins.

Implementation of a communications setup between the mass spectrometer and MALDI laser would allow WREnS scripts to directly control the laser. The source's current form requires the laser to be controlled by separate software. Incorporation with WREnS would allow further levels of automation. For example, it was beneficial to turn the laser off between samples on an automatic run, otherwise ions were still produced from ablation of the exposed stainless-steel target. This meant the difference in ion signal used to distinguish between samples in AMX was more obvious than for chromatograms produced with constant laser pulsing. In addition, it may be possible to synchronise the laser pulses and ion optics to further improve overall system

sensitivity by more efficiently transferring ions from source to detector. Splitting of the laser beam to form multiple foci, and thus ablation areas, on the droplet may be another way to improve sensitivity.

The benefits of other lasers could be explored. A λ=2.94 μm IR laser would be interesting to incorporate since water and glycerol absorb this wavelength readily, and UV-MALDI matrices could be omitted from the workflow. IR lasers have already been demonstrated for remote sampling with a probe design like that of the iKnife.[269, 270] Continuation of the work with high repetition rate lasers may enable more rapid analysis and greater sensitivity, but the conditions will need further optimisation. Other UV-MALDI matrix compounds may be more suitable than 2,5-DHB for lasers with a wavelength other than λ=337 nm. The continuing work of group member Jeffrey Brown has focused on a solid-state laser (λ=349 nm). One interesting experiment would be to use a higher (>30 Hz) repetition rate λ=337 nm laser to see how signal intensity scales for this wavelength and beam profile.

5.2   Determination of structural information from metal-cationised phospholipids

This work could continue to be developed to include other species of phospholipids. For instance, it would be interesting to see how species such as phosphatidylserine and phosphatidylinositol behave since they are typically analysed in the negative ion mode. Confirmation of the lack of $[PE+Ba]^{2+}$ formation is also important, since the current hypothesis is only based on a single compound. Finding a way to keep Ba^{2+} coordinated to the fatty acids would be beneficial for MS^3 sensitivity, since much of it is dissociated as $Ba^{+\bullet}$ or $BaOH^+$. The performance of this method on an ion trap instrument capable of 'real' MS^3 (i.e. not ion mobility-enabled) would be interesting to assess.

It may be interesting to attempt Paternò-Büchi reactions with a modified capillary and acetone vapour in the gas flow. Using a quartz capillary and UV lamp (λ=193 nm) this could be an alternative approach to characterisation with the Ba^{2+} adduct ions and be a demonstration of reactions in the ablated microdroplets. Alternatively, ozone could be used as the counter-flow gas, similarly to OzESI, to assess another potential gas phase reaction for phospholipid structure determination.

5.3   High-throughput milk speciation

The analysis and discrimination of multiple milk classes was demonstrated, but there is scope to expand this methodology to other dairy products and food groups in general. Liquid products are an obvious target since the source is natively compatible with this sample state. It may be

possible to use the information obtained from this research to learn how to analyse liquid food samples without sample preparation. However, a short extraction procedure has been shown to be very beneficial for easily detecting many analytes, so this should not be discounted. Other samples, such as fruit juices, could be investigated in this way.

The throughput of this method reported here was approximately 10 s/sample, but it was demonstrated that suitable spectra could be obtained at a rate of 1 s/sample. This makes for a very high-throughput method on the standard 96-well MALDI sample plate. Since the source is limited by the XY-stage translation speed, an obvious and straightforward modification would be to use commercially available 386-well sample plates. The distance between each well is much smaller (approximately 2 mm), meaning shorter stage travel distance would be required. The small volume of wells on these plates (estimated to be approx. 0.25 μL) would not be an issue, since, a liquid AP-MALDI analysis (performed on the order of seconds) consumes such a small volume of sample. Practically, the only required change to the method would be the coordinates in the WREnS script, to adapt for the higher density and different well locations on the 386-well sample plate.

5.4   High-throughput mastitis detection

Generally, the method is already successful. However, there are developments that should be targeted for further validation and increased rapidity.

First, developments to the ion source described above could increase the sensitivity. This would improve peak intensity per scan and allow reduced sample analysis time, which ultimately would improve throughput. Improved sensitivity may also allow detection of lower abundance peptides, which in turn could improve classification accuracy. Sensitivity improvements may also be achieved from modifications to the sample preparation procedure. For example, replacement of sodium hydroxide with ammonium hydroxide may reduce ion suppression and adduct ion formation. Modifications should not impact the high-throughput focus of the method however, so time-consuming procedures like solid phase extraction should be avoided.

The software used for building the classification models and online recognition (i.e. AMX and AMX Recognition) both require some usability improvements for this method. AMX would benefit from the ability to name individual samples on the chromatogram. At the time of writing, only the scan numbers are reported which makes relating spectra to samples cumbersome. This could be implemented in a standard sample list format, with support for importing .csv files for automatic spectrum assignment. This will be especially important for large datasets.

The first principal component of a model is often the result of a combination of experimental variables (for instance, sample age on target, sample position on target, etc.). It would be useful to be able to exclude specific components from the LDA calculation when they are not relevant to the classification of the sample groups. It would also be useful to be able to utilise the ion mobility data recorded during analysis. The mobility function is not currently used to allow filtering in AMX. The ability to import and apply drift time rules would enable subtraction of the signal for MALDI matrix ions, ubiquitous synthetic polymers etc. Currently this can be performed with DriftScope, but the exported data is often unreadable or appears corrupted when added to an AMX project.

The AMX Recognition software requires the addition of a method more suitable for the liquid AP-MALDI analyses. The software was developed primarily for use with the iKnife/REIMS ion source, where sampling is performed for an undefined amount of time. Thus, the sample is detected by the change in TIC. The inclusion of a mode where the sample is detected by the change in intensity of a certain ion (for example a MALDI matrix peak) could make for a more reliable detection system for liquid AP-MALDI. The matrix ion would not be detected outside of the sample wells and so ions produced from the MALDI sample plate would not interfere with the detection. This was an issue when the laser was not switched off between samples, where sample TIC was of similar intensity as the TIC from ablation of sample plate material. Alternatively, since each liquid AP-MALDI sample is analysed for a predetermined amount of time, the software could be told to expect samples during set scan ranges. These could be calculated by the software after the first sample scan is defined by the user.

5.5   Chromophore-lacking liquid AP-LDI MS analysis

An obstacle to this being adopted for automated analysis is that LSM **G** requires the laser to be focused on the edge of the droplet for ions to be detected, whereas LSMs containing a chromophore compound do not. A computerised system which can identify the sample edge and position the laser accordingly would be necessary for this to be applicable to high-throughput analysis with the current MALDI sample plate design. This could be done optically i.e. with a camera system. Alternatively, this could be achieved using a WREnS script to identify where signal intensity is greatest as the laser focus is scanned across the sample, although this would likely add a few seconds per sample to any analysis. Optionally, a new target geometry could be developed that simplified producing ions from these samples. Finally, the sample consumption compared to chromophore-containing matrices appeared greater; this might explain the higher signal intensity

but comes at the expense of sample longevity. Further optimisation of the analytical conditions may help to extend sample lifetime.

As discussed in section 5.1.2, the use of a laser with a different wavelength might be a suitable solution. Water, glycerol and other diols would readily absorb IR radiation, and the recent development of picosecond pulse duration IR lasers could limit thermal effects. This modification has already been demonstrated to be a compelling soft ionisation technique with a similar ionisation inlet to the one used for this research project.[271]

5.6   Antibiotic resistance testing

A proof of concept has been demonstrated for using the conversion of β-lactam antibiotics to their respective hydrolysis as an indicator for β-lactamase activity in milk. This has yet to be demonstrated for samples containing β-lactamase-producing bacteria. As future projects within the research group will continue with a more specific focus on microorganism profiling, there is opportunity for the applicability of this method to be assessed.

As with all parts of this project, overall sensitivity improvements would benefit the method. Specific to this section, this could be achieved by performing a more efficient analyte extraction, for example by adding a dispersive SPE step. This would add time to the procedure but would still be faster than bacteria culturing. This should always be a consideration when modifying this method in the future.

The MS method would be easy to adapt to highly sensitive triple-quadrupole mass spectrometers. It would be written as an MRM experiment in much the same way as the PRM experiment here. Tuning of transitions to the specific instrument used might be necessary. In theory, the ion source itself can be mounted to all Waters mass spectrometers manufactured within the last ten years with ease.

The method could be expanded to assess additional β-lactam antibiotics. Unfortunately, other classes of antibiotics, such as macrolides, are not modified to be less potent by resistant bacteria; rather the drug target is modified to be less susceptible to attack. For example, methylation of the bacterial ribosome is an erythromycin resistance pathway.[272] This modification would be much more challenging to detect than the β-lactam hydrolysis because the modified molecule is so much larger. A more specific analyte extraction procedure would likely be required for this to work. The proteinaceous content of bacteria in the samples could also be analysed, perhaps by adoption of the method described by Barreiro *et al.*[221]

Bibliography

Bibliography

1.   K. Wien, *Brazilian Journal of Physics*, 1999, **29**, 401-414.
2.   L. W. L. W. Austin, *The positive charges carried by the canal rays*, U.S. Dept. of Commerce and Labor, Bureau of Standards : U.S. Govt. Print. Off, 1905.
3.   J. J. Thomson, *Proceedings of the Royal Society of London. Series A*, 1913, **89**, 1.
4.   A. J. Dempster, *Physical Review*, 1918, **11**, 316-325.
5.   F. W. Aston, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 1919, **38**, 707-714.
6.   A. J. Dempster, *Nature*, 1935, **136**, 180.
7.   A. O. Nier, E. T. Booth, J. R. Dunning and A. V. Grosse, *Physical Review*, 1940, **57**, 546-546.
8.   J. Coster-Mullen, *Atom bombs : the top secret inside story of Little Boy and Fat Man*, J. Coster-Mullen, [United States], 2008.
9.   W. E. Stephens, *Bulletin of the American Physical Society*, 1946, **21**, 22.
10.   A. E. Cameron and D. F. E. Jr., *Review of Scientific Instruments*, 1948, **19**, 605-607.
11.   M. M. Wolff and W. E. Stephens, *Review of Scientific Instruments*, 1953, **24**, 616-617.
12.   H. R. Morris, T. Paxton, A. Dell, J. Langhorne, M. Berg, R. S. Bordoli, J. Hoyes and R. H. Bateman, *Rapid Communications in Mass Spectrometry*, 1996, **10**, 889-896.
13.   G. L. Andrews, B. L. Simons, J. B. Young, A. M. Hawkridge and D. C. Muddiman, *Analytical Chemistry*, 2011, **83**, 5442-5446.
14.   M. A. Park, *Genetic Engineering & Biotechnology News*, 2016, **36**, 16-17.
15.   A. Wallace, http://www.waters.com/webassets/cms/library/docs/local_seminar_presentations/SW_NMS2011_Product_intr_Vendors_sem_44_Alistair_Wallacepdf.pdf, 2011, [2 February 2015]
16.   C. Xie, D. Zhong, K. Yu and X. Chen, *Bioanalysis*, 2012, **4**, 937-959.
17.   http://www.nobelprize.org/nobel_prizes/physics/laureates/1989/ [24 May 2018]
18.   A. G. Marshall, in *eMagRes*.
19.   K. H. Kingdon, *Physical Review*, 1923, **21**, 408-418.
20.   R. D. Knight, *Applied Physics Letters*, 1981, **38**, 221-223.
21.   A. Makarov, *Analytical Chemistry*, 2000, **72**, 1156-1162.
22.   http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2002/ [19 January 2015]
23.   K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida and T. Matsuo, *Rapid Communications in Mass Spectrometry*, 1988, **2**, 151-153.
24.   C. M. Whitehouse, R. N. Dreyer, M. Yamashita and J. B. Fenn, *Analytical Chemistry*, 1985, **57**, 675-679.
25.   J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse, *Science*, 1989, **246**, 64-71.
26.   I. V. Chernushevich, A. V. Loboda and B. A. Thomson, *Journal of mass spectrometry : JMS*, 2001, **36**, 849-865.
27.   S. Guan and A. G. Marshall, *Journal of the American Society for Mass Spectrometry*, 1996, **7**, 101-106.

28. A. Radionova, I. Filippov and P. J. Derrick, *Mass Spectrom. Rev.*, 2016, **35**, 738-757.

29. R. J. Beuhler and L. Friedman, *International Journal of Mass Spectrometry and Ion Physics*, 1977, **23**, 81-97.

30. E. A. Dennis, A. W. Gundlach-Graham, S. J. Ray, C. G. Enke and G. M. Hieftje, *Journal of the American Society for Mass Spectrometry*, 2016, **27**, 1772-1786.

31. E. de Hoffmann and V. Stroobant, *Mass Spectrometry: Principles and Applications, 3rd Edition*, John WIley and Sons Limited, 2007.

32. T. I. Wang, C. W. Chu, H. M. Hung, G. S. Kuo and C. C. Han, *Review of Scientific Instruments*, 1994, **65**, 1585-1589.

33. T. Satoh, T. Sato, A. Kubo and J. Tamura, *Journal of the American Society for Mass Spectrometry*, 2011, **22**, 797-803.

34. D. M. Murphy, *Journal of the American Society for Mass Spectrometry*, 2017, **28**, 242-246.

35. I. V. Chernushevich, S. I. Merenbloom, S. Liu and N. Bloomfield, *Journal of the American Society for Mass Spectrometry*, 2017, **28**, 2143-2150.

36. M. Fernandez Ocaña, J. Jarvis, R. Parker, M. Bramley Peter, M. Halket John, K. P. Patel Raj and H. Neubert, *Proteomics*, 2005, **5**, 1209-1216.

37. M. Smith and C. Thickitt, *Rapid communications in mass spectrometry : RCM*, 2009, **23**, 3018-3022.

38. F. W. McLafferty, P. F. Bente, R. Kornfeld, S. C. Tsai and I. Howe, *Journal of Mass Spectrometry*, 1995, **30**, 797-806.

39. P. Roepstorff and J. Fohlman, *Biological Mass Spectrometry*, 1984, **11**, 601-601.

40. R. S. Johnson, S. A. Martin, K. Biemann, J. T. Stults and J. T. Watson, *Analytical Chemistry*, 1987, **59**, 2621-2625.

41. N. M. Riley and J. J. Coon, *Analytical Chemistry*, 2018, **90**, 40-64.

42. M. Huzarska, I. Ugalde, D. A. Kaplan, R. Hartmer, M. L. Easterling and N. C. Polfer, *Analytical Chemistry*, 2010, **82**, 2873-2878.

43. J. R. Jhingree, R. Beveridge, E. R. Dickinson, J. P. Williams, J. M. Brown, B. Bellina and P. E. Barran, *International Journal of Mass Spectrometry*, 2017, **413**, 43-51.

44. J. A. Mosely, M. J. P. Smith, A. S. Prakash, M. Sims and A. W. T. Bristow, *Analytical Chemistry*, 2011, **83**, 4068-4075.

45. R. A. Zubarev, K. F. Haselmann, B. Budnik, F. Kjeldsen and F. Jensen, *European Journal of Mass Spectrometry*, 2017, **8**, 337-349.

46. J. E. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz and D. F. Hunt, *Proceedings of the National Academy of Sciences of the United States of America*, 2004, **101**, 9528-9533.

47. N. D. Udeshi, J. Shabanowitz, D. F. Hunt and K. L. Rose, *The FEBS journal*, 2007, **274**, 6269-6276.

48. X. Liang, J. Liu, Y. LeBlanc, T. Covey, A. C. Ptak, J. T. Brenna and S. A. McLuckey, *Journal of the American Society for Mass Spectrometry*, 2007, **18**, 1783-1788.

49. R. L.-K. Griffiths, http://etheses.bham.ac.uk/5233/ University of Birmingham, 2015.

50. F. Lermyte, A. Konijnenberg, J. P Williams, J. Brown, D. Valkenborg and F. Sobott, *Journal of The American Society for Mass Spectrometry*, 2014.

51. A. Theisen, B. Yan, J. M. Brown, M. Morris, B. Bellina and P. E. Barran, *Analytical Chemistry*, 2016, **88**, 9964-9971.

52. K. L. Fort, A. Dyachenko, C. M. Potel, E. Corradini, F. Marino, A. Barendregt, A. A. Makarov, R. A. Scheltema and A. J. R. Heck, *Analytical Chemistry*, 2016, **88**, 2303-2310.
53. D. R. Klein and J. S. Brodbelt, *Analytical Chemistry*, 2017, **89**, 1516-1522.
54. J. Laskin and J. H. Futrell, *Mass Spectrom. Rev.*, 2005, **24**, 135-167.
55. S. Becher, B. Spengler and S. Heiles, *European Journal of Mass Spectrometry (Chichester)*, 2018, **24**, 54-65.
56. J. A. Madsen, T. W. Cullen, M. S. Trent and J. S. Brodbelt, *Analytical Chemistry*, 2011, **83**, 5107-5113.
57. T. Ly and R. R. Julian, *Angewandte Chemie*, 2009, **48**, 7130-7137.
58. L. E. Talbert and R. R. Julian, *Journal of the American Society for Mass Spectrometry*, 2018.
59. J. P. O'Brien, B. D. Needham, J. C. Henderson, E. M. Nowicki, M. S. Trent and J. S. Brodbelt, *Analytical Chemistry*, 2014, **86**, 2138-2145.
60. E. Ryan, C. Q. N. Nguyen, C. Shiea and G. E. Reid, *Journal of the American Society for Mass Spectrometry*, 2017, **28**, 1406-1419.
61. D. R. Hernandez, J. D. Debord, M. E. Ridgeway, D. A. Kaplan, M. A. Park and F. Fernandez-Lima, *The Analyst*, 2014, **139**, 1913-1921.
62. K. Giles, J. P. Williams and I. Campuzano, *Rapid communications in mass spectrometry : RCM*, 2011, **25**, 1559-1566.
63. J. N. Dodds, J. C. May and J. A. McLean, *Analytical Chemistry*, 2017, **89**, 12176-12184.
64. http://www.waters.com/waters/en_GB/Travelling-wave-%28T-Wave™%29-/nav.htm?locale=en_GB&cid=134663694 [03 May 2018]
65. G. Paglia, P. Angel, J. P. Williams, K. Richardson, H. J. Olivos, J. W. Thompson, L. Menikarachchi, S. Lai, C. Walsh, A. Moseley, R. S. Plumb, D. F. Grant, B. O. Palsson, J. Langridge, S. Geromanos and G. Astarita, *Analytical Chemistry*, 2015, **87**, 1137-1144.
66. R. Kaufmann, F. Hillenkamp, R. Nitsche, M. Schurmann and R. Wechsung, *Microscopica acta. Supplement*, 1978, 297-306.
67. R. Kaufmann, *Scanning electron microscopy*, 1980, 641-646.
68. A. Orsulakova, C. Morgenstern, R. Kaufmann and M. D'Haese, *Scanning electron microscopy*, 1982, 1763-1766.
69. J. F. Osborn and E. Gabriel, *Deutsche zahnarztliche Zeitschrift*, 1982, **37**, 699-700.
70. U. Seydel, B. Lindner, U. Zahringer, E. T. Rietschel, S. Kusumoto and T. Shiba, *Biomedical mass spectrometry*, 1984, **11**, 132-141.
71. R. Kaufmann, F. Hillenkamp and R. Wechsung, *Medical progress through technology*, 1979, **6**, 109-121.
72. M. Karas, D. Bachmann and F. Hillenkamp, *Analytical Chemistry*, 1985, **57**, 2935-2939.
73. M. Karas and F. Hillenkamp, *Analytical Chemistry*, 1988, **60**, 2299-2301.
74. T. Wenzel, K. Sparbier, T. Mieruch and M. Kostrzewa, *Rapid communications in mass spectrometry : RCM*, 2006, **20**, 785-789.
75. C. H. Le, J. Han and C. H. Borchers, *Analytical Chemistry*, 2012, **84**, 8391-8398.
76. G. Stübiger and O. Belgacem, *Analytical Chemistry*, 2007, **79**, 3206-3213.
77. T. W. Jaskolla, W. D. Lehmann and M. Karas, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 12200-12205.

78. K. Breuker, R. Knochenmuss and R. Zenobi, *International Journal of Mass Spectrometry*, 1999, **184**, 25-38.

79. B. Stahl, S. Thurl, J. R. Zeng, M. Karas, F. Hillenkamp, M. Steup and G. Sawatzki, *Analytical Biochemistry*, 1994, **223**, 218-226.

80. R. L. Vermillion-Salsbury, D. M. Hercules and G. Cooks, *Rapid Communications in Mass Spectrometry*, 2002, **16**, 1575-1581.

81. R. Shroff and A. Svatos, *Analytical Chemistry*, 2009, **81**, 7954-7959.

82. S. Trimpin, D. E. Clemmer and C. N. McEwen, *Journal of the American Society for Mass Spectrometry*, 2007, **18**, 1967-1972.

83. C. K. L. Wong and T. W. Dominic Chan, *Rapid Communications in Mass Spectrometry*, 1997, **11**, 513-519.

84. C. Yang, X. Hu, A. V. Loboda and R. H. Lipson, *Journal of the American Society for Mass Spectrometry*, 2010, **21**, 294-299.

85. A. Overberg, M. Karas, U. Bahr, R. Kaufmann and F. Hillenkamp, *Rapid Communications in Mass Spectrometry*, 1990, **4**, 293-296.

86. R. Cramer, W. J. Richter, E. Stimson and A. L. Burlingame, *Analytical Chemistry*, 1998, **70**, 4939-4944.

87. A. Pirkl, J. Soltwisch, F. Draude and K. Dreisewerd, *Analytical Chemistry*, 2012, **84**, 5669-5676.

88. E. Nordhoff, F. Kirpekar, M. Karas, R. Cramer, S. Hahner, F. Hillenkamp, K. Kristiansen, P. Roepstroff and A. Lezius, *Nucleic Acids Research*, 1994, **22**, 2460-2465.

89. A. Koch, A. Schnapp, J. Soltwisch and K. Dreisewerd, *International Journal of Mass Spectrometry*, 2016, **416**, 61-70.

90. T. N. Peel, N. C. Cole, B. L. Dylla and R. Patel, *Diagnostic microbiology and infectious disease*, 2014.

91. K. De Bruyne, B. Slabbinck, W. Waegeman, P. Vauterin, B. De Baets and P. Vandamme, *Systematic and applied microbiology*, 2011, **34**, 20-29.

92. A. Calderaro, M. C. Arcangeletti, I. Rodighiero, M. Buttrini, C. Gorrini, F. Motta, D. Germini, M. C. Medici, C. Chezzi and F. De Conto, *Scientific reports*, 2014, **4**, 6803.

93. B. Hammarstrom, B. Nilson, T. Laurell, J. Nilsson and S. Ekstrom, *Analytical Chemistry*, 2014, **86**, 10560-10567.

94. M. Karas, U. Bahr, A. Ingendoh and F. Hillenkamp, *Angewandte Chemie International Edition in English*, 1989, **28**, 760-761.

95. M. W. Towers, J. E. McKendrick and R. Cramer, *Journal of Proteome Research*, 2010, **9**, 1931-1940.

96. K. M. Kirmess, R. Knochenmuss, G. J. Blanchard and G. R. Kinsel, *Journal of mass spectrometry : JMS*, 2016, **51**, 79-85.

97. H. Ehring, M. Karas and F. Hillenkamp, *Organic Mass Spectrometry*, 1992, **27**, 472-480.

98. R. Knochenmuss and R. Zenobi, *Chemical Reviews*, 2003, **103**, 441-452.

99. Y. Bae, Y. Shin, J. Moon and M. Kim, *Journal of the American Society for Mass Spectrometry*, 2012, **23**, 1326-1335.

100. M. Karas, M. Glückmann and J. Schäfer, *Journal of Mass Spectrometry*, 2000, **35**, 1-12.

101. K. Dreisewerd, *Chemical Reviews*, 2003, **103**, 395-426.

102. M. Karas and R. Krüger, *Chemical Reviews*, 2003, **103**, 427-440.

103. R. Knochenmuss, *The Analyst*, 2006, **131**, 966-986.
104. R. Knochenmuss, *Mass Spectrometry*, 2013, **2**, S0006.
105. R. Knochenmuss, *Analytical Chemistry*, 2003, **75**, 2199-2207.
106. R. Knochenmuss, *Journal of Mass Spectrometry*, 2002, **37**, 867-877.
107. L. Konermann, E. Ahadi, A. D. Rodriguez and S. Vahidi, *Analytical Chemistry*, 2013, **85**, 2-9.
108. R. Cramer, A. Pirkl, F. Hillenkamp and K. Dreisewerd, *Angewandte Chemie*, 2013, **52**, 2364-2367.
109. P. Ryumin, J. Brown, M. Morris and R. Cramer, *Methods*, 2016, **104**, 11-20.
110. S. F. Teunissen and M. N. Eberlin, *Journal of the American Society for Mass Spectrometry*, 2017.
111. M. Wilm, *Molecular & cellular proteomics : MCP*, 2011, **10**, M111.009407-M009111.009407.
112. O. J. Hale and R. Cramer, *Analytical and Bioanalytical Chemistry*, 2018, **410**, 1435-1444.
113. V. S. Pagnotti, N. D. Chubatyi and C. N. McEwen, *Analytical Chemistry*, 2011, **83**, 3981-3985.
114. S. Trimpin, D. W. Woodall, S. Thawoos, C. D. Foley, J. Li, E. D. Inutan and P. M. Stemmer, *Methods*, 2016.
115. S. Trimpin, C. Lee, S. M. Weidner, T. J. El-Baba, C. A. Lutomski, E. D. Inutan, C. D. Foley, C. K. Ni and C. N. McEwen, *Chemphyschem : a European journal of chemical physics and physical chemistry*, 2018, **19**, 581-589.
116. A. D. Catherman, O. S. Skinner and N. L. Kelleher, *Biochemical and Biophysical Research Communications.*, 2014, **445**, 683-693.
117. K. Jurowski, K. Kochan, J. Walczak, M. Baranska, W. Piekoszewski and B. Buszewski, *Critical reviews in Analytical Chemistry*, 2017, **47**, 418-437.
118. A. Sloan, G. Wang and K. Cheng, *Clinica chimica acta; international journal of clinical chemistry*, 2017, **473**, 180-185.
119. O. Vorm, P. Roepstorff and M. Mann, *Analytical Chemistry (Washington)*, 1994, **66**, 3281-3287.
120. Y. Dai, R. M. Whittal and L. Li, *Analytical Chemistry*, 1999, **71**, 1087-1091.
121. F. Xiang, C. Beavis Ronald and W. Ens, *Rapid Communications in Mass Spectrometry*, 1994, **8**, 199-204.
122. Ö. Patrik, E. Simon, B. Jonas, N. Johan, L. Thomas and M. V. György, *Rapid Communications in Mass Spectrometry*, 1999, **13**, 315-322.
123. J. D. Leszyk, *Journal of biomolecular techniques : JBT*, 2010, **21**, 81-91.
124. Y. L. Li, M. L. Gross and F. F. Hsu, *Journal of the American Society for Mass Spectrometry*, 2005, **16**, 679-682.
125. R. W. Garden and J. V. Sweedler, *Analytical Chemistry*, 2000, **72**, 30-36.
126. A. Tholey and E. Heinzle, *Analytical and Bioanalytical Chemistry*, 2006, **386**, 24-37.
127. S. L. Cohen and B. T. Chait, *Analytical Chemistry*, 1996, **68**, 31-37.
128. D. W. Armstrong, L.-K. Zhang, L. He and M. L. Gross, *Analytical Chemistry*, 2001, **73**, 3679-3686.
129. R. Cramer and S. Corless, *Proteomics*, 2005, **5**.
130. R. Cramer and M. Towers, *Spectroscopy*, 2007, **22**.
131. R. Cramer, M. Karas and T. W. Jaskolla, *Analytical Chemistry*, 2014, **86**, 744-751.

132.   Jeffrey Brown, Michael Morris and Rainer Cramer, 65th ASMS Conference on Mass Spectrometry and Allied Topics, Indianapolis, IN, USA, 2017.

133.   V. M. Doroshenko, V. V. Laiko, N. I. Taranenko, V. D. Berkout and H. S. Lee, *International Journal of Mass Spectrometry*, 2002, **221**, 39-58.

134.   E. Moskovets, *Rapid communications in mass spectrometry : RCM*, 2015, **29**, 1501-1512.

135.   E. Moskovets, A. Misharin, V. Laiko and V. Doroshenko, *Methods*, 2016, **104**, 21-32.

136.   M. Kompauer, S. Heiles and B. Spengler, *Nature Methods*, 2016.

137.   Y. Li, B. Shrestha and A. Vertes, *Analytical Chemistry*, 2007, **79**, 523-532.

138.   J. Balog, L. Sasi-Szabo, J. Kinross, M. R. Lewis, L. J. Muirhead, K. Veselkov, R. Mirnezami, B. Dezso, L. Damjanovich, A. Darzi, J. K. Nicholson and Z. Takats, *Science translational medicine*, 2013, **5**, 194ra193.

139.   E. Jones, T. KARANCSI, S. D. Pringle, J. BALOG, D. SIMON, L. GÖDÖRHÁZY, D. SZALAY and Z. TAKÁTS, Google Patents, 2016.

140.   K. C. Sachfer, T. Szaniszlo, S. Gunther, J. Balog, J. Denes, M. Keseru, B. Dezso, M. Toth, B. Spengler and Z. Takats, *Analytical Chemistry*, 2011, **83**, 1632-1640.

141.   K. C. Schafer, J. Balog, T. Szaniszlo, D. Szalay, G. Mezey, J. Denes, L. Bognar, M. Oertel and Z. Takats, *Analytical Chemistry*, 2011, **83**, 7729-7735.

142.   M. R. Wenk, *Nature Reviews Drug Discovery*, 2005, **4**, 594.

143.   T. Hu and J. L. Zhang, 2018, **41**, 351-372.

144.   R. W. Gross, *Biochimica et biophysica acta*, 2017, **1862**, 731-739.

145.   C. Hinz, S. Liggi and J. L. Griffin, *Current opinion in chemical biology*, 2018, **42**, 42-50.

146.   F.-F. Hsu, A. Bohrer and J. Turk, *Journal of the American Society for Mass Spectrometry*, 1998, **9**, 516-526.

147.   H. J. Yoo and K. Haåkansson, *Analytical Chemistry*, 2011, **83**, 1275-1283.

148.   F. F. Hsu and J. Turk, *Journal of the American Society for Mass Spectrometry*, 2008, **19**, 1681-1691.

149.   M. Kliman, J. C. May and J. A. McLean, *Biochimica et biophysica acta*, 2011, **1811**, 935-945.

150.   G. Sun, K. Yang, Z. Zhao, S. Guan, X. Han and R. W. Gross, *Analytical Chemistry*, 2008, **80**, 7576-7585.

151.   F. F. Hsu and J. Turk, *Journal of the American Society for Mass Spectrometry*, 2008, **19**, 1673-1680.

152.   P. F. James, M. A. Perugini and R. A. O'Hair, *Rapid communications in mass spectrometry : RCM*, 2007, **21**, 757-763.

153.   R. L. Kozlowski, J. L. Campbell, T. W. Mitchell and S. J. Blanksby, *Analytical and Bioanalytical Chemistry*, 2015, **407**, 5053-5064.

154.   B. L. Poad, M. R. Green, J. M. Kirk, N. Tomczyk, T. W. Mitchell and S. J. Blanksby, *Analytical Chemistry*, 2017, **89**, 4223-4229.

155.   M. C. Thomas, T. W. Mitchell, D. G. Harman, J. M. Deeley, J. R. Nealon and S. J. Blanksby, *Analytical Chemistry*, 2008, **80**, 303-311.

156.   N. Vu, J. Brown, K. Giles and Q. Zhang, *Rapid communications in mass spectrometry : RCM*, 2017, **31**, 1415-1423.

157.   H. T. Pham, A. T. Maccarone, J. L. Campbell, T. W. Mitchell and S. J. Blanksby, *Journal of the American Society for Mass Spectrometry*, 2013, **24**, 286-296.

158. H. T. Pham, A. T. Maccarone, M. C. Thomas, J. L. Campbell, T. W. Mitchell and S. J. Blanksby, *The Analyst*, 2014, **139**, 204-214.
159. D. L. Marshall, H. T. Pham, M. Bhujel, J. S. Chin, J. Y. Yew, K. Mori, T. W. Mitchell and S. J. Blanksby, *Analytical Chemistry*, 2016, **88**, 2685-2692.
160. R. L. Kozlowski, T. W. Mitchell and S. J. Blanksby, *Scientific reports*, 2015, **5**, 9243.
161. M. C. Thomas, T. W. Mitchell and S. J. Blanksby, in *Lipidomics: Volume 1: Methods and Protocols*, ed. D. Armstrong, Humana Press, Totowa, NJ, 2009, pp. 413-441.
162. X. Ma and Y. Xia, *Angewandte Chemie*, 2014, **53**, 2592-2596.
163. J. Ren, E. T. Franklin and Y. Xia, *Journal of the American Society for Mass Spectrometry*, 2017, **28**, 1432-1441.
164. P. Li and G. P. Jackson, *Journal of mass spectrometry : JMS*, 2017, **52**, 271-282.
165. T. Baba, J. L. Campbell, J. C. Y. Le Blanc and P. R. S. Baker, *Analytical Chemistry*, 2017.
166. K. Sharma and M. Paradakar, *Food Security*, 2010, **2**, 97-107.
167. J. D. G. McEvoy, *Drug Testing and Analysis*, 2016, **8**, 511-520.
168. J. C. Moore, J. Spink and M. Lipp, *Journal of food science*, 2012, **77**, R118-126.
169. S. Gerosa and J. Skoet, *Milk availability : trends in production and demand and medium-term outlook*, [Rome] : FAO, Agricultural Development Economics Div., 2012.
170. T. Azad and S. Ahmed, *International Journal of Food Contamination*, 2016, **3**.
171. R.-K. Chen, *Rapid communications in mass spectrometry*, 2004, **18**, 1167-1171.
172. P. F. Scholl, S. M. Farris and M. M. Mossoba, *J. Agric. Food. Chem.*, 2014, **62**, 1498-1505.
173. V. Hrbek, L. Vaclavik, O. Elich and J. Hajslova, *Food Control*, 2014, **36**, 138-145.
174. G. J. Van Berkel and V. Kertesz, *Rapid communications in mass spectrometry : RCM*, 2017, **31**, 281-291.
175. J. Schiller, R. SÜß, M. Petkovi and K. Arnold, *Journal of Food Lipids*, 2002, **9**, 185-200.
176. T. T. Ng, P. K. So, B. Zheng and Z. P. Yao, *Analytica chimica acta*, 2015, **884**, 70-76.
177. J. S. Garcia, G. B. Sanvido, S. A. Saraiva, J. J. Zacca, R. G. Cosso and M. N. Eberlin, *Food Chem.*, 2012, **131**, 722-726.
178. C. Calvano, C. De Ceglie, A. Aresta, L. Facchini and C. Zambonin, *Anal. BioAnalytical Chemistry*, 2013, **405**, 1641-1649.
179. C. D. Calvano, C. D. Ceglie, L. D'Accolti and C. G. Zambonin, *Food chemistry*, 2012, **134**, 1192-1198.
180. C. D. Calvano, C. De Ceglie, A. Monopoli and C. G. Zambonin, *Journal of mass spectrometry : JMS*, 2012, **47**, 1141-1149.
181. R. Cozzolino, S. Passalacqua, S. Salemi, P. Malvagna, E. Spina and D. Garozzo, *Journal of mass spectrometry : JMS*, 2001, **36**, 1031-1037.
182. R. Cozzolino, S. Passalacqua, S. Salemi and D. Garozzo, *Journal of mass spectrometry : JMS*, 2002, **37**, 985-991.
183. V. Cunsolo, V. Muccilli, R. Saletti and S. Foti, *Journal of mass spectrometry : JMS*, 2013, **48**, 148-153.
184. M. Sassi, S. Arena and A. Scaloni, *J. Agric. Food. Chem.*, 2015, **63**, 6157-6171.
185. J. Balog, D. Perenyi, C. Guallar-Hoyas, A. Egri, S. D. Pringle, S. Stead, O. P. Chevallier, C. T. Elliott and Z. Takats, *J. Agric. Food. Chem.*, 2016, **64**, 4793-4800.

186. K. Verplanken, S. Stead, R. Jandova, C. V. Poucke, J. Claereboudt, J. V. Bussche, S. Saeger, Z. Takats, J. Wauters and L. Vanhaecke, *Talanta*, 2017, **169**, 30-36.
187. Y. Guitton, G. Dervilly-Pinel, R. Jandova, S. Stead, Z. Takats and B. Le Bizec, *Food Additives & Contaminants: Part A*, 2017, 1-11.
188. C. Black, O. P. Chevallier, S. A. Haughey, J. Balog, S. Stead, S. D. Pringle, M. V. Riina, F. Martucci, P. L. Acutis, M. Morris, D. S. Nikolopoulos, Z. Takats and C. T. Elliott, *Metabolomics*, 2017, **13**, 153.
189. L. Du, W. Lu, Z. J. Cai, L. Bao, C. Hartmann, B. Gao and L. L. Yu, *Food chemistry*, 2018, **240**, 573-578.
190. T. M. Guerreiro, D. N. de Oliveira, C. Melo, E. de Oliveira Lima, M. D. S. Ribeiro and R. R. Catharino, *Food research international*, 2018, **108**, 498-504.
191. M. Driffield, E. Bradley, L. Castle, A. Lloyd, M. Parmar, D. Speck, D. Roberts and S. Stead, *Rapid Communications in Mass Spectrometry*, 2015, **29**, 1603-1610.
192. A. Dalmia, *Journal of AOAC International*, 2017, **100**, 573-575.
193. B. T. Ruotolo, J. L. P. Benesch, A. M. Sandercock, S.-J. Hyung and C. V. Robinson, *Nature Protocols*, 2008, **3**, 1139.
194. Y. Yang, N. Zheng, J. Yang, D. Bu, J. Wang, L. Ma and P. Sun, *International Dairy Journal*, 2014, **35**, 15-20.
195. M. Lees, *The Analytical Scientist*, 2016, 18-19.
196. A. Haug, A. T. Hostmark and O. M. Harstad, *Lipids in health and disease*, 2007, **6**, 25.
197. O. E. Mäkinen, V. Wanhalinna, E. Zannini and E. K. Arendt, *Critical Reviews in Food Science and Nutrition*, 2016, **56**, 339-349.
198. A. K. Anagnostopoulos, A. Katsafadou, V. Pierros, E. Kontopodis, G. C. Fthenakis, G. Arsenos, S. C. Karkabounas, AthinaTzora, I. Skoufos and G. T. Tsangaris, *Journal of Proteomics*, 2016.
199. R. K. Chen, L. W. Chang, Y. Y. Chung, M. H. Lee and Y. C. Ling, *Rapid communications in mass spectrometry : RCM*, 2004, **18**, 1167-1171.
200. J. Lu, L. Liu, X. Pang, S. Zhang, Z. Jia, C. Ma, L. Zhao and J. Lv, *Food Chem.*, 2016.
201. N. Nicolaou, Y. Xu and R. Goodacre, *Analytical and Bioanalytical Chemistry*, 2011, **399**, 3491-3502.
202. G. Picariello, R. Sacchi and F. Addeo, *European Journal of Lipid Science and Technology*, 2007, **109**, 511-524.
203. M. Stahlman, C. S. Ejsing, K. Tarasov, J. Perman, J. Boren and K. Ekroos, *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences*, 2009, **877**, 2664-2672.
204. E. Sokol, T. Ulven, N. J. Faergeman and C. S. Ejsing, *European journal of lipid science and technology : EJLST*, 2015, **117**, 751-759.
205. Q. Zhou, B. Gao, X. Zhang, Y. Xu, H. Shi and L. L. Yu, *Food chemistry*, 2014, **143**, 199-204.
206. I. Haddad, M. Mozzon, R. Strabbioli and N. G. Frega, *International Dairy Journal*, 2011, **21**, 119-127.
207. H. Hogeveen, K. Huijps and T. J. Lam, *New Zealand veterinary journal*, 2011, **59**, 16-23.
208. S. More, *Irish veterinary journal*, 2009, **62 Suppl 4**, S5-14.
209. P. L. Ruegg, *Journal of Dairy Science*, 2017, **100**, 10381-10397.
210. The European Commission, *Official Journal of the European Union*, 2009, **37/2010**.

211. C. M. Duarte, P. P. Freitas and R. Bexiga, *Journal of Veterinary Diagnostic Investigation*, 2015.
212. S. C. Archer, A. J. Bradley, S. Cooper, P. L. Davies and M. J. Green, *Preventive veterinary medicine*, 2017, **144**, 1-6.
213. http://qmms.co.uk/services/bacteriology [27 March 2018]
214. C. Kamphuis, R. Sherlock, J. Jago, G. Mein and H. Hogeveen, *Journal of Dairy Science*, 2008, **91**, 4560-4570.
215. A. Langer, S. Sharma, N. K. Sharma and D. S. Nauriyal, *International Journal of Applied Sciences and Biotechnology*, 2014, **2**.
216. Supriya, S. Vivek and L. Deepika, *Haryana Veterinarian*, 2010, **49**, 64-65.
217. R. Baeker, S. Haebel, K. Schlatterer and B. Schlatterer, *Prostaglandins Other Lipid Mediat.*, 2002, **67**, 75-88.
218. P. Roncada, C. Piras, A. Soggiu, R. Turk, A. Urbani and L. Bonizzi, *Journal of Proteomics*, 2012, **75**, 4259-4274.
219. C. J. Hogarth, J. L. Fitzpatrick, A. M. Nolan, F. J. Young, A. Pitt and P. D. Eckersall, *Proteomics*, 2004, **4**, 2094-2100.
220. L. B. Larsen, P. L. H. McSweeney, M. G. Hayes, J. B. Andersen, K. L. Ingvartsen and A. L. Kelly, *International Dairy Journal*, 2006, **16**, 1-8.
221. J. R. Barreiro, J. L. Goncalves, P. A. C. Braga, A. G. Dibbern, M. N. Eberlin and M. Veiga Dos Santos, *Journal of Dairy Science*, 2017, **100**, 2928-2934.
222. P. Ryumin and R. Cramer, *Analytica chimica acta*, 2018, **1013**, 43-53.
223. O. J. Hale, P. Ryumin, J. M. Brown, M. Morris and R. Cramer, *Rapid communications in mass spectrometry : RCM*, 2018, **0**.
224. https://www.slideshare.net/WatersChemical/ion-mobility-petro-org-software-novel-techniques-for-petroleomics-investigations [19 June 2018]
225. https://www.spectra-physics.com/service/obsolete-product-manuals [21 June 2018]
226. I. GmbH, InnoLight GmbH, 2009, vol. 2018.
227. M. Sud, E. Fahy, D. Cotter, A. Brown, E. A. Dennis, C. K. Glass, A. H. Merrill, R. C. Murphy, C. R. H. Raetz, D. W. Russell and S. Subramaniam, *Nucleic Acids Research*, 2007, **35**, D527-D532.
228. M. T. Marty, A. J. Baldwin, E. G. Marklund, G. K. A. Hochberg, J. L. P. Benesch and C. V. Robinson, *Analytical Chemistry*, 2015, **87**, 4370-4376.
229. M. Strohalm, M. Hassman, B. Košata and M. Kodíček, *Rapid Communications in Mass Spectrometry*, 2008, **22**, 905-908.
230. J. Zhou, C. Liu, D. Si, B. Jia, L. Zhong and Y. Yin, *Analytica chimica acta*, 2017, **972**, 62-72.
231. R. Shannon, *Acta Crystallographica Section A*, 1976, **32**, 751-767.
232. F.-F. Hsu and J. Turk, *Journal of the American Society for Mass Spectrometry*, 2003, **14**, 352-363.
233. J. L. Peñalvo, M. C. Castilho, M. I. N. Silveira, M. C. Matallana and M. E. Torija, *European Food Research and Technology*, 2004, **219**, 251-253.
234. J. Balog, T. Szaniszlo, K. C. Schaefer, J. Denes, A. Lopata, L. Godorhazy, D. Szalay, L. Balogh, L. Sasi-Szabo, M. Toth and Z. Takats, *Analytical Chemistry*, 2010, **82**, 7343-7350.

235. K.-C. Schäfer, J. Dénes, K. Albrecht, T. Szaniszló, J. Balog, R. Skoumal, M. Katona, M. Tóth, L. Balogh and Z. Takáts, *Angewandte Chemie International Edition*, 2009, **48**, 8240-8242.

236. E. R. St John, J. Balog, J. S. McKenzie, M. Rossi, A. Covington, L. Muirhead, Z. Bodai, F. Rosini, A. V. M. Speller, S. Shousha, R. Ramakrishnan, A. Darzi, Z. Takats and D. R. Leff, *Breast cancer research : BCR*, 2017, **19**, 59.

237. D. Simon, T. Karansci, S. Pringle and Z. Takats, 65th ASMS Conference on Mass Spectrometry and Allied Topics, Indianapolis, IN, USA, 2017.

238. B. Emerson, J. Gidden, J. O. Lay, Jr. and B. Durham, *Jounal of Lipid Research*, 2010, **51**, 2428-2434.

239. K. Pearson, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 1901, **2**, 559-572.

240. H. Han, *BMC Bioinformatics*, 2010, **11**, S1-S1.

241. J. Lever, M. Krzywinski and N. Altman, *Nature Methods*, 2017, **14**, 641.

242. A. M. Martinez and A. C. Kak, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**, 228-233.

243. J. Ponthus and E. Riches, *International Journal for Ion Mobility Spectrometry*, 2013, **16**, 95-103.

244. L. A. Lerno, Jr., J. B. German and C. B. Lebrilla, *Analytical Chemistry*, 2010, **82**, 4236-4245.

245. J. P. Williams, R. Lock, V. J. Patel and J. H. Scrivens, *Analytical Chemistry*, 2006, **78**, 7440-7445.

246. R. G. Brereton and G. R. Lloyd, *J. Chemom.*, 2014, **28**, 213-225.

247. S.-T. Tsai, C. W. Chen, L. C. L. Huang, M.-C. Huang, C.-H. Chen and Y.-S. Wang, *Analytical Chemistry*, 2006, **78**, 7729-7734.

248. C. H. Hsiao, C. W. Hong, B. H. Liu, C. W. Chen, C. C. Wu and Y. S. Wang, *Rapid communications in mass spectrometry : RCM*, 2011, **25**, 834-842.

249. J. Trygg and S. Wold, *J. Chemom.*, 2002, **16**, 119-128.

250. F. J. Bikker, G. Koop, N. B. Leusink, K. Nazmi, W. E. Kaman, H. S. Brand and E. C. I. Veerman, *Veterinary Research Communications*, 2014, **38**, 271-277.

251. H. Yuan, K. M. Vance, C. E. Junge, M. T. Geballe, J. P. Snyder, J. R. Hepler, M. Yepes, C.-M. Low and S. F. Traynelis, *Journal of Biological Chemistry*, 2009, **284**, 12862-12873.

252. R. de Vries, M. Brandt, Å. Lundh, K. Holtenius, K. Hettinga and M. Johansson, *Journal of Dairy Science*, 2016, **99**, 9300-9306.

253. F. Baum, M. Fedorova, J. Ebner, R. Hoffmann and M. Pischetsrieder, *Journal of Proteome Research*, 2013, **12**, 5447-5462.

254. J. Meltretter, A. Schmidt, A. Humeny, C. M. Becker and M. Pischetsrieder, *Journal of agricultural and food chemistry*, 2008, **56**, 2899-2906.

255. R. Mansor, W. Mullen, A. Albalat, P. Zerefos, H. Mischak, D. C. Barrett, A. Biggs and P. D. Eckersall, *Journal of Proteomics*, 2013, **85**, 89-98.

256. B. Wang and S. Trimpin, *Analytical Chemistry*, 2014, **86**, 1000-1006.

257. J. P. Graham, J. J. Boland and E. Silbergeld, *Public Health Rep.*, 2007, **122**, 79-87.

258. L. Jank, M. T. Martins, J. B. Arsand, R. B. Hoff, F. Barreto and T. M. Pizzolato, *Food Additives & Contaminants: Part A*, 2015, **32**, 1992-2001.

259. P. Ryumin, J. Brown, M. Morris and R. Cramer, *International Journal of Mass Spectrometry*, 2017, **416**, 20-28.

260. S. Suwanrumpha, A. Flory Donald, B. Freas Royal and L. Vestal Marvin, *Biomedical & Environmental Mass Spectrometry*, 1988, **16**, 381-386.

261. C. Guijas, J. R. Montenegro-Burke, X. Domingo-Almenara, A. Palermo, B. Warth, G. Hermann, G. Koellensperger, T. Huan, W. Uritboonthai, A. E. Aisporna, D. W. Wolan, M. E. Spilker, H. P. Benton and G. Siuzdak, *Analytical Chemistry*, 2018, **90**, 3156-3164.

262. C. K. Fagerquist, A. R. Lightfield and S. J. Lehotay, *Analytical Chemistry*, 2005, **77**, 1473-1482.

263. T. Bessaire, C. Mujahid, A. Beck, A. Tarres, M.-C. Savoy, P.-M. Woo, P. Mottier and A. Desmarchelier, *Food Additives & Contaminants: Part A*, 2018, **35**, 661-673.

264. J. M. Daniel, V. V. Laiko, V. M. Doroshenko and R. Zenobi, *Analytical and Bioanalytical Chemistry*, 2005, **383**, 895-902.

265. https://www.nmr.co.uk [18 November 2018]

266. https://www.imperial.ac.uk/news/185333/mothers-donate-breast-milk-pioneering-citizen/ [19 November 2018]

267. K. Wiangnon and R. Cramer, in *Advances in MALDI and Laser-Induced Soft Ionization Mass Spectrometry*, ed. R. Cramer, Springer International Publishing, Cham, 2016, pp. 65-76.

268. D. Simon, T. Karancsi, J. Balog, E. Jones, S. D. Pringle and Z. Takats, 64th ASMS Conference on Mass Spectrometry and Allied Topics, San Antonio, TX, USA, 2016.

269. B. Fatou, P. Saudemont, E. Leblanc, D. Vinatier, V. Mesdag, M. Wisztorski, C. Focsa, M. Salzet, M. Ziskind and I. Fournier, *Scientific reports*, 2016, **6**, 25919.

270. B. Fatou, M. Ziskind, P. Saudemont, J. Quanico, C. Focsa, M. Salzet and I. Fournier, *Molecular & Cellular Proteomics*, 2018.

271. Y. Lu, C. L. Pieterse, W. D. Robertson and R. J. D. Miller, *Analytical Chemistry*, 2018, **90**, 4422-4428.

272. R. Leclercq and P. Courvalin, *Antimicrobial Agents and Chemotherapy*, 2002, **46**, 2727-2734.

6 Appendix

6.1 WREnS Script – Arbitrary Wells Acquisition:

Bold text would be replaced with values as required for the acquisition.

```csharp
using System;
using System.IO;
using System.Threading;
using System.Collections.Generic;
using MLReader;
using MassSupport;
using GenericInterface;
using AcqWrensClient;
using WrensDataProcessing;
using RunCode;
using DataConvertor;
using Solid.Arduino;
using Stage;
using StageLogic;
using SumClassLib;
using Zaber.Serial.Core;

namespace RunCodePartial
{
public class RunCodeHeader : RunCodeScript
{
public override void main()
{
    connect("epc");
      //this will start the script when acquisition is started from MassLynx
      //watch_file(@"C:\Masslynx\ProjectFolder\Data");

      //less specific as it watches a whole folder,
      //watch out for data processing taking place while you run as this will
      change the files and kick off the script

      watch_folder(@"D:\data\AP MALDI\Month YEAR.PRO\Data");
      //watch_folder(@"D:\data\AP MALDI\name.PRO\Data");

}
public override void MainMethod()
{

    var port = new Zaber.Serial.Core.ZaberBinaryPort("COM4");
      port.Open();
      //StageLogic.Primitives.runPatternPlate(port,
@"C:\wrens_data\88targetplatea.txt", 1000);
      var dict = StageLogic.Helper.populateDictionary(X_coordinate,Y_cooridinate);
      //sample well position dictionary. Requires A1 position, use A2 position
coordinate
  // knowing 1 mm equals 20997 steps and 4.5 mm distance between each well. I.e. A2
X-coordinate - (4.5*20997) = A1 X-coordinate
      dict = StageLogic.Helper.addLMDictionary(dict, X_coordinate,Y_cooridinate);
      var SampleWell = new List<Tuple<string, int>>
      {
          Tuple.Create( "LMA1", time ms),
          Tuple.Create( "A2", time ms),
//etc for wells as required.
```

```csharp
            // LM = lockmass well, time in milliseconds. Modify as required. Can
analyse all wells except for Column 1.

        };
         foreach (var well in SampleWell) {
         StageLogic.Primitives.move_sample_well(port, dict, well.Item1);
         wait(well.Item2);
         }

         port.Close();
}

}//class
}//namespace
```

6.2 Appendix 2: Identified milk lipids from liquid AP-MALDI MS and ion mobility data

Table 16: Milk lipids identified by m/z and collision cross section.

Description[a]	m/z	Mass Error (ppm)	Ion	CCS (Å²)[b]	ΔCCS (%)[c]	Ref. CCS (Å²)[d]
Cer(d40:1(2OH))	638.605	-4.7	M+H			
PC(25:0(CHO))	650.439	0.4	M+H			
PA(O-34:1)	661.520	4.5	M+H			
PA(P-34:0)	661.520	4.5	M+H			
PS(P-28:0)	664.456	2.4	M+H			
PA(34:5)	665.425	9.3	M-H			
SM(d32:1)	675.539	-6.7	M+H			
PE-Cer(d35:1)	675.539	-6.7	M+H			
PE-Cer(d34:1(2OH))	677.523	0.3	M+H			
PS(28:0)	678.430	-7.6	M-H			
PA(35:5)	679.431	-5.5	M-H			
PG(O-30:1)	679.494	4.5	M+H			
PG(P-30:0)	679.494	4.5	M+H			
SM(d16:1/17:0)	689.555	-6.3	M+H			
PE-Cer(d36:1)	689.555	-6.3	M+H			
PS(P-30:1)	690.475	5.9	M+H			
PI(24:0)	697.391	-4.1	M-H			
PA(O-37:2)	699.533	-0.3	M-H			
PA(P-37:1)	699.533	-0.3	M-H			
PA(P-37:0)	701.546	-4.6	M-H			
SM(d34:1)	703.573	-2.5	M+H	293	0.8%	291
PE-Cer(d37:1)	703.573	-2.5	M+H			
PG(31:1)	705.469	-2.6	M-H			
PA(O-37:0)	705.581	2.2	M+H			
PC(30:0)	706.538	-0.4	M+H			
PE-NMe(32:0)	706.538	-0.4	M+H			
PE(33:0)	706.538	-0.4	M+H			
PA(O-38:1)	717.586	9.2	M+H			
PA(P-38:0)	717.586	9.2	M+H			
SM(d35:1)	717.586	-6.5	M+H			
PE-Cer(d38:1)	717.586	-6.5	M+H			
PC(31:1)	718.545	9.5	M+H			
PE(34:1)	718.545	9.5	M+H			
PC(31:0)	720.553	-1.5	M+H			
PE-NMe2(32:0)	720.553	-1.5	M+H			
PE(34:0)	720.553	-1.5	M+H			
PI(26:0)	725.432	9.5	M-H			
PC(32:3)	728.518	-6.8	M+H			
PE(35:3)	728.518	-6.8	M+H			
SM(d36:1)	731.600	-8.8	M+H	299	0.7%	297
PA(O-39:1)	731.600	6.6	M+H			
PA(P-39:0)	731.600	6.6	M+H			
PE-Cer(d39:1)	731.600	-8.8	M+H			
PC(32:1)	732.554	0.6	M+H	291	1.4%	287
PE-NMe(34:1)	732.554	0.6	M+H			
PE(35:1)	732.554	0.6	M+H			
PC(32:0)	734.569	-1.2	M+H	295	1.2%	291
PE(35:0)	734.569	-1.2	M+H			
PI(37:0)	739.438	-2.5	M-H			
PA(40:10))	739.438	5.4	M-H			
PC(O-34:2)	742.571	-5.7	M-H			
PC(P-24:1)	742.571	-5.7	M-H			
PE(O-37:2)	742.571	-5.7	M-H			
PnC(34:1)	742.571	-5.7	M-H			
PG(34:3)	745.506	5.8	M+H			

PC(33:1)	746.571	2.7	M+H			
PE(36:1)	746.571	2.7	M+H			
PE-NMe2(34:1)	746.571	2.7	M+H			
PA(39:0)	747.584	-7.8	M+H			
PC(33:0)	748.586	1.8	M+H			
PE(36:0)	748.586	1.8	M+H			
PI(28:1)	751.446	7.5	M-H			
PI(28:0)	753.454	-2.6	M-H			
PPA(34:1)	753.454	8.3	M-H			
PC(34:3)	756.549	-6.3	M+H	297	1.3%	293
PE(37:3)	756.549	-6.3	M+H			
PC(34:2)	758.565	-5.3	M+H	294	0.5%	293
PE-NMe(36:2)	758.565	-5.3	M+H			
PE(27:2)	758.565	-5.3	M+H			
SM(d38:1)	759.630	-9.4	M+H	306	0.2%	305
PA(P-41:0)	759.630	5.4	M+H			
PC(34:1)	760.585	0.1	M+H	298	0.9%	295
PE(37:1)	760.585	0.1	M+H			
PG(35:1)	761.539	6.7	M-H			
PC(34:0)	762.599	-2.4	M+H	300	1.1%	297
PE(37:0)	762.599	-2.4	M+H			
PE-NMe(36:0)	762.599	-2.4	M+H			
SM(d39:1)	773.652	-1.8	M+H			
PA(O-42:0)	775.663	6.6	M+H			
PI(30:0)	781.493	6.9	M-H			
PC(36:4)	782.566	-5.0	M+H	299	0.4%	298
PE(39:4)	782.566	-5.0	M+H			
PC(36:3)	784.577	-9.8	M+H	300	0.3%	299
PS(O-37:2)	786.565	-1.2	M-H			
PS(P-37:1)	786.565	-1.2	M-H			
PC(36:2)	786.598	-3.1	M+H	301	0.0%	301
PE(39:2)	786.598	-3.1	M+H			
SM(d40:1)	787.667	-2.3	M+H	311	0.1%	311
PS(O-37:1)	788.579	-3.1	M-H			
PS(P-37:0)	788.579	-3.1	M-H			
PC(O-37:1)	788.659	8.1	M+H			
PC(P-37:0)	788.659	8.1	M+H			
PE(O-40:1)	788.659	8.1	M+H			
PE(P-40:0)	788.659	8.1	M+H			
PC(O-38:6)	790.581	6.2	M-H			
PC(P-38:5)	790.581	6.2	M-H			
PC(O-37:0)	790.670	2.4	M+H			
PE(O-40:0)	790.670	2.4	M+H			
MGDG(38:8)	799.542	8.5	M+H			
PG(38:4)	799.542	-7.7	M+H			
SM(d41:2)	799.661	-9.6	M+H			
PS(37:3)	800.546	3.4	M+H			
SM(d41:1(OH))	801.683	-1.8	M+H	314	0.4%	313
PI(O-33:1)	807.543	4.6	M-H			
PI(P-33:0)	807.543	4.6	M-H			
PI(O-33:0)	809.553	-2.3	M-H			
PA(44:3)	811.614	-9.4	M+H			
SM(d42:1)	815.698	-2.7	M+H	317	0.4%	316
PG(39:1)	817.593	-4.3	M-H			
PI(O-34:0)	823.569	-1.4	M-H			
PS(O-40:4)	826.593	-3.6	M+H			
PI(O-35:2)	833.560	6.4	M-H			
PI(O-16:0/19:1(9Z))	835.570	-1.1	M-H			
PI(P-35:0)	835.570	-1.1	M-H			
PI(O-35:0)	837.584	-2.6	M-H			
PG(41:1)	845.626	-1.9	M-H			
PG(42:6)	849.573	9.1	M-H			
PI(O-36:0)	851.595	-8.0	M-H			

PI(P-37:2)	859.569	-1.4	M-H			
PI (36:2)	861.556	7.6	M-H	300	0.2%	300
PI(O-37:2)	861.587	1.3	M-H			
PI(P-37:1)	861.587	1.3	M-H			
PI (36:1)	863.570	5.5	M-H	306	0.6%	304
PI(O-37:1)	863.602	0.5	M-H			
PI(P-37:0)	863.602	0.5	M-H			
PS (40:3)	864.577	1.3	M-H	302	-1.1%	305
PG(44:7)	875.575	-7.0	M-H			
PG(44:6)	877.591	-6.7	M-H			
PI(P-39:1)	889.625	8.7	M-H			

[a]Lipid nomenclature adapted from the LIPID MAPS Classification System (E. Fahy, S. Subramaniam, R. Murphy, M. Nishijima, C. Raetz, T. Shimizu, F. Spener, G. van Meer, M. Wakelam, and E. Dennis, *Journal of Lipid Research,* **2009**, *50*, S9-S14). [b]Experimental collisional cross section (CCS) as measured by ion mobility. [c]Percentage difference of experimental CCS compared to the reference literature CCS. [d]Reference literature value of CCS obtained from Waters Metabolic Profiling CCS Library (Waters Corporation) or K. M. Hines, J. C. May, J. A. McLean, L. B. Xu, *Anal. Chem.* **2016**, *88*, 7329. Negative ions are highlighted in grey.

6.3 Appendix 3: MS/MS peak tables for Figure 72, Figure 73 and Figure 74

Table 17: Main sequence ions for CID MS/MS [M+5H]5+ m/z 869

Ion	Experimental *m/z* (mono)	Theoretical *m/z* (mono)	*m/z* Accuracy (ppm)
y_2^+	246.1968	246.1812	63
b_6^{2+}	365.2418	365.2296	33
y_6^{2+}	354.2092	354.2080	3
y_3^+	375.2379	375.2238	38
y_7^{2+}	403.7391	403.7422	-8
b_{13}^{3+}	506.6420	506.6319	20
y_9^{2+}	516.7934	516.7898	7
b_{14}^{3+}	549.6494	549.6461	6
y_5^+	560.3416	560.3402	2
b_9^{2+}	561.8424	561.8358	12
b_{15}^{3+}	582.6757	582.6689	12
b_{10}^{2+}	590.3508	590.3465	7
y_{11}^{2+}	638.8580	638.8504	12
b_{11}^{2+}	646.8944	646.8886	9
b_{22}^{4+}	650.3874	650.3739	21
b_{17}^{3+}	658.3884	658.3779	16
y_{12}^{2+}	674.3672	674.3690	-3
y_6^+	707.4135	707.4087	7
y_{13}^{2+}	723.9108	723.9032	11
b_{20}^{3+}	777.1219	777.1011	27
y_7^+	806.4832	806.4771	8
b_{14}^{2+}	823.9641	823.9655	-2
y_{15}^{2+}	870.9816	870.9716	12
b_{23}^{3+}	915.8631	915.8523	12
b_{16}^{2+}	930.0436	930.0418	2
y_8^+	935.5259	935.5197	7
b_{24}^{3+}	964.8926	964.8751	18
b_{25}^{3+}	997.9117	997.8979	14
b_{26}^{3+}	1021.5961	1021.5769	19
y_9^+	1032.5814	1032.5724	9
b_{36}^{4+}	1056.0659	1056.0864	-19
b_{28}^{3+}	1102.9733	1102.9507	21
y_{10}^+	1179.6567	1179.6408	14
b_{10}^+	1179.6592	1179.6858	-23
y_{21}^{2+}	1240.6566	1240.6705	-11
y_{11}^+	1276.7162	1276.6936	18
b_{11}^+	1292.7584	1292.7698	-9
y_{22}^{2+}	1297.2396	1297.2125	21
y_{12}^+	1347.7568	1347.7307	19
b_{25}^{2+}	1496.3866	1496.3432	29
b_{26}^{2+}	1531.9108	1531.8618	32

Table 18: Main sequence ions for CID MS/MS [M+4H]4+ m/z 729

Ion	Experimental *m/z* (mono)	Theoretical *m/z* (mono)	*m/z* Accuracy (ppm)
y_1^+	166.0893	166.0863	18
y_2^+	313.1601	313.1547	17
b_6^{2+}	365.2325	365.2296	8
b_3^+	382.2412	382.2561	-39
b_{11}^{3+}	431.6056	431.5948	25
y_3^+	469.2502	469.2558	-12
b_8^{2+}	497.8162	497.8065	20
b_4^+	519.3247	519.3150	19
b_{14}^{3+}	549.6517	549.6461	10
b_9^{2+}	561.8514	561.8358	28
y_4^+	582.3381	582.3398	-3
b_{10}^{2+}	590.3516	590.3465	9
y_{10}^{2+}	632.8926	632.8561	58‡
b_{11}^{2+}	646.8951	646.8886	10
y_5^+	695.4228	695.4239	-2
b_{12}^{2+}	695.4228	695.4149	11
b_6^+	729.4448	729.4519	10
b_{13}^{2+}	759.4464	759.4442	3
y_6^+	809.4839	809.4668	21
b_{14}^{2+}	823.9678	823.9655	3
b_7^+	857.5694	857.5468	26
b_{15}^{2+}	873.5227	873.4997	26
b_{16}^{2+}	930.0602	930.0418	20
y_7^+	938.5136	938.5094	5
y_8^+	1052.5731	1052.5524	20
y_9^+	1165.6617	1165.6364	22
y_{10}^+	1264.7825	1264.7048	61‡
b_{11}^+	1292.7837	1292.7698	11
y_{11}^+	1393.8014	1393.7474	39
y_{13}^+	1618.9507	1618.8588	57‡

‡likely overlap with a-ions

Table 19: Main sequence ions for CID MS/MS [M+7H]7+ m/z 1351

Ion	Experimental *m/z* (mono)	Theoretical *m/z* (mono)	*m/z* Accuracy (ppm)
y_2^+	231.1705	231.1703	1
y_4^+	441.2980	441.3071	-21
y_{10}^{2+}	547.8270	547.8397	-23
y_{12}^{2+}	632.8947	632.8924	4
y_6^+	685.4058	685.4283	-33
y_{14}^{2+}	730.9427	730.9530	-14
y_7^+	742.4437	742.4498	-8
y_{15}^{2+}	795.4787	795.4743	6
b_8^+	886.4579	886.4516	7
y_8^+	898.5356	898.5509	-17
y_{17}^{2+}	941.0295	941.0353	-6
b_{24}^{3+}	943.8433	943.8217	23
b_{26}^{3+}	1013.8719	1013.8673	5
b_9^+	1023.5091	1023.5106	-1
b_{10}^+	1136.5979	1136.5946	3
b_{27}^{3+}	1046.2239	1046.2182	5
y_{10}^+	1094.6735	1094.6721	1
y_{11}^+	1151.7012	1151.6935	7
b_{53}^{5+}	1179.8472	1179.8429	4
b_{31}^{3+}	1188.9603	1188.9547	5
b_{32}^{3+}	1237.9468	1237.9775	-25
y_{12}^+	1264.7770	1264.7776	0
b_{69}^{6+}	1310.2096	1310.2015	6
b_{59}^{5+}	§	1331.1140	n/a
b_{70}^{6+}	1331.6980	1331.7086	-8
b_{73}^{6+}	§	1383.2428	n/a
b_{74}^{6+}	§	1392.7464	n/a
b_{24}^{2+}	1415.2584	1415.2289	21
b_{76}^{6+}	§	1425.4332	n/a
b_{77}^{6+}	1451.4886	1451.4501	27
y_{14}^+	1460.8993	1460.8988	0
b_{80}^{6+}	§	1501.6405	n/a
b_{26}^{2+}	1520.3162	1520.2973	12
y_{41}^{3+}	1537.4891	1537.5411	-34
y_{15}^+	1589.9457	1589.9414	3
$y°_{58}^{4+}$	1597.8291	1597.8665	-23
y_{58}^{4+}	1602.4028	1602.3871	10
$b*_{44}^{3+}$	§	1649.5294	n/a
y_{16}^+	1717.9913	1718.0000	-5
y_{31}^{2+}	1775.9213	1775.9678	-26
b_{31}^{2+}	1782.9479	1782.9283	11
b_{32}^{2+}	1856.4905	1856.4626	15
y_{17}^+	1881.0402	1881.0633	-12
y_{18}^+	1994.1421	1994.1473	-3

§: *identification suggested by isotopic pattern and approx. m/z from low intensity signal*

Table 20: Main sequence ions for CID MS/MS [M+8H]8+ m/z 1478

Ion	Experimental *m/z* (mono)	Theoretical *m/z* (mono)	*m/z* Accuracy (ppm)
y_2^+	231.1701	231.1703	-1
b_3^+	395.2552	395.2037	130
y_4^+	441.2979	441.3071	-21
y_{12}^{2+}	632.8943	632.8924	3
y_{14}^{2+}	730.9381	730.9530	-20
y_7^+	742.4234	742.4498	-36
y_{15}^{2+}	795.4708	795.4743	-4
y_8^+	898.5626	898.5509	13
b_{16}^{2+}	979.4773	979.4795	2
y_{10}^+	1094.6942	1094.6721	20
y_{11}^+	1151.6827	1151.6935	-9
y_{12}^+	1264.8586	1264.7776	64[‡]
b_{11}^+	1354.7030	1354.6976	4
y_{14}^+	1460.9055	1460.8988	5
y_{30}^{2+}	1727.4359	1727.4414	-3
b_{14}^+	1727.8645	1727.8614	2

[‡]likely overlap with a-ions