Combined Tandem Mass Spectrometry and Ion Mobility Spectrometry in Proteome Analyses

A thesis submitted to the University of Manchester for the

degree of Doctor of Philosophy (PhD) in the Faculty of Engineering and Physical Sciences.

2012

Ross Chawner

School of Chemistry

Table of Contents

List of Figures

greater extent and with appropriate selection of voltages compensation for the energy difference is achieved meaning the ions reach the detector at the same time.

Figure 11. Nomenclature for product ions resulting from peptide backbone fragmentation. Those incorporating the *N*-terminus are labelled b-, c- and a-ions and *C*-terminal species are termed y-, z- and x-ions.

Figure 12. Illustration of a SRIG showing the orientation of electrodes and applied RF voltage.

Figure 13. Movement of the 'travelling wave' created by the transient DC voltage within a T-Wave mobility device. Ions are propelled along the z-axis as a function of time.

Figure 14. Schematic of the Synapt HDMS TW-IMS Mass Spectrometer produced by Waters Corporation.

Figure 15. Diagram illustrating the application of an asymmetric waveform to a flat electrode FAIMS instrument, the subsequent ion motion along the z-axis is also shown. The waveform is applied to the upper plate while the lower plate is held at ground potential.

Figure 16. Plot of drift time vs *m/z* for peptides observed from analysis of trypsin digestion mixtures from a range of 'standard' proteins. From this library, candidate peptides were selected for inclusion in QCAL-IM.

Figure 17. Drift times for the singly, doubly and triply protonated peptides selected for inclusion in QCAL-IM, following trypsin proteolysis of a set of 'standard' proteins.

Figure 18. Peptide map of doubly protonated ions resulting from proteolytic cleavage of QCAL-IM using either trypsin or Lys-C; shown are peptide number and *m/z* value of each species.

Figure 19. QCAL-IM peptide ion Ω values versus *m/z*; cross section values were obtained on the modified Synapt instrument with a Linear field mobility drift tube in the presence of helium drift gas.

Figure 20. Linear regression comparing experimentally determined Ω of doubly charged peptide ions on both the modified Synapt and QTOF Linear field mobility instruments.

Figure 21. Plot of drift time versus *m/z* for QCAL-IM tryptic and Lys-C peptides following T-Wave mobility separation in nitrogen drift gas.

Figure 22. Drift time versus Ω charge dependent calibration lines for QCAL-IM peptides. Both vertical and horizontal error bars are smaller than any given data point.

Figure 23. Plot of T-Wave drift time versus Ω' (charge and reduce mass corrected Ω) for QCAL-IM tryptic and Lys-C peptides.

Figure 24. Linear regression of published Ω against average calculated Ω (n=3) following calibration of the T-Wave instrument with QCAL-IM for a selection of $[M+H]^+$ and $[M+2H]^{2+}$ peptides. Peptides without replicate measurement (n=1) in the literature values were excluded from this plot.

Figure 25. Mobility separation on the first generation Synapt HDMS instrument (A) and the higher resolution Synapt G2 (B) of the $[M+4H]^{4+}$ peptide T19/20 at 1266.64 *m/z* produced from Lys-C proteolysis of QCAL-IM. Three conformers are observed, of which conformers 1 and 3 are sufficiently separated to be used for the resolving power calculation.

Figure 26. QIT product ion spectra generated by (A) CID of doubly charged tryptic peptide MPC*TEDYLSLILNR at 862.91 *m/z* and (B) CID of doubly charged CBPB treated analogue MPC*TEDYLSLILN at 784.87 *m/z*. In each instance C* represents a carbamidomethylated cysteine residue.

Figure 27. Comparison of QIT MS/MS spectra generated by (A) ETD of the doubly charged tryptic peptide TVMENFVAFVDK at 700.35 *m/z* and (B) ETD of the doubly charged CBPB treated analogue TVMENFVAFVD at 636.29 *m/z*.

Figure 28. Comparison of QTOF MS/MS spectra generated by (A) CID of doubly charged tryptic peptide LKPDPNTLC*DEFK at 788.88 *m/z* and (B) CID of doubly charged CBPB treated analogue LKPDPNTLC*DEF at 724.84 *m/z*. In each instance C* represents a carbamidomethylated cysteine residue.

Figure 29. Comparison of product ion apparent isotope distributions resulting from ETD of the Lys-C derived peptide EC*C* HGDLLEC*ADDRADLAK and the cognate CBPB treated analogue (where C* is carbamidomethylated C). (A) and (B) show the apparent isotope distributions of the c_9 product ion from the Lys-C and CBPB peptides respectively, with the CBPB treated analogue exhibiting increased c-1 ion formation. (C) and (D) show the corresponding apparent isotope distributions of the Lys-C peptide z_{12} product ion and CBPB

peptide z_{11} product ion (relating to the same bond cleavage) with the Lys-C peptide exhibiting evidence of z+1 ion formation.

Figure 30. (A) The b_5 ion is generated by CID resulting in formation of the oxazolone ring through attack of the oxygen lone pair on the newly formed terminal carbonyl group. (B) *N*-terminal nucleophilic attack of the oxazolone ring to form the macrocycle. (C) The resultant macrocyclic intermediate undergoes ring opening at various positions to give sequence scrambled b_5 ions.

Figure 31. Comparison of QTOF MS/MS spectra resulting from (A) CID of $[M+H]^+$ peptide KGVLHAV at 723.45 *m/z* and (B) CID of $[M+H]^+$ peptide K(guanidinyl)GVLHAV at 765.47 *m/z*.

Figure 32. Comparison of QTOF MS/MS spectra resulting from (A) CID of $[M+H]^+$ KNVPLY at 733.42 *m/z* and (B) CID of $[M+H]^+$ derivatised peptide K(guanidinyl)NVPLY at 775.45 *m/z*.

Figure 33. Comparison of QTOF MS/MS spectra resulting from (A) CID of $[M+H]^+$ KLVTDLT at 789.47 *m/z* and (B) CID of $[M+H]^+$ derivatised peptide K(guanidinyl)LVTDLT at 831.49 *m/z*.

Figure 34. Comparison of QTOF MS/MS spectra resulting from (A) CID of $[M+H]^+$ RGVLHAV at 751.45 *m/z* and (B) CID of $[M+H]^+$ RNVPLY at 761.43 *m/z* and (C) CID of $[M+H]^+$ RLVTDLT at 817.47 *m/z*.

Figure 35. (A) QTOF MS/MS spectrum generated by CID of the $[M+2H]^{2+}$ peptide FGERALK. The $[b_6 3]b_5$ fragment ion observed at 545.34 *m/z* is attributable to scrambling of the original peptide sequence via ring opening of a macrocycle intermediate. (B) QIT MS/MS spectrum generated by CID of the $[M+2H]^{2+}$ peptide FGERALK.

Figure 36. CID of $[M+2H]^{2+}$ FGERALK with the collision energy set at either (A) 19.5 V or (B) 26 V using a QTOF instrument results in formation of a y_6 ion at 673.39 *m/z*. Distortion of the isotope distribution in (A) suggests a contribution from the b_6 ion at 674.36 *m/z*. Shown in the inset is the theoretical isotope distribution of the y_6 ion which was modelled using the MS-Isotope program within Protein Prospector. (C) Equivalent CID analysis of $[M+2H]^{2+}$ FGERALK using a QIT instrument results in formation of an intense b_6 ion at 674.36 *m/z*.

Figure 37. T-Wave Ion Mobility analysis of (A) peptides generated by trypsin digestion of α-lactalbumin (B) peptides generated by sequential trypsin and CBPB digestion of α-Lactalbumin.

Figure 38. Comparison of drift time versus *m/z* for (A) 20 arginine terminating peptides and (B) 45 lysine terminating peptides both prior to and following removal of the *C*-terminal basic residue.

Figure 39. T-Wave ion mobility separation of b-ions (blue diamonds) and y-ions (red squares) generated by CID of a selection of peptides produced by Lys-C proteolysis.

Figure 40. T-Wave ion mobility separation of b-ions (blue diamonds) and y-ions (red squares) generated by CID of a selection of peptides produced by Lys-N proteolysis.

Figure 41. T-Wave ion mobility separation of c/z-ions generated from ETD of the $[M+3H]^{3+}$ peptide RHPEYAVSVLLR.

Figure 42. T-Wave ion mobility separation of c/z-ions generated from ETD of $[M+3H]^{3+}$ Substance P (RPKPQQFFGLM).

Figure 43. T-Wave ion mobility separation of c/z-ions generated from ETD of $[M+3H]^{3+}$ peptide SLHTLFGDELC*K, where C* represents carbamidomethylated cysteine.

Figure 44. T-Wave ion mobility separation of c-ions (blue diamonds) and z-ions (red squares) generated by ETD of a selection of peptides produced by trypsin/Lys-C proteolysis.

List of Tables

Table 9. Collision cross section of singly protonated peptides produced by Lys-C and Lys-N proteolysis of a range of 'standard' proteins. Cross section values were determined following calibration of the T-Wave instrument using the QCAL-IM ion mobility standard and are an average of 3 replicate measurements.

Table 10. Collision cross section of doubly protonated peptides produced by Lys-C and Lys-N proteolysis of a range of 'standard' proteins. Cross section values were determined following calibration of the T-Wave instrument using the QCAL-IM ion mobility standard and are an average of 3 replicate measurements. C* represents a carbamidomethylated cysteine residue.

List of Abbreviations

AC	Alternating current
BSA	bovine serum albumin
CBPB	carboxypeptidase B
CID	Collision-Induced dissociation
CRM	Charged Residue Model
DC	Direct current
DNA	deoxyribonucleic acid
DTT	dithiothreitol
ECD	Electron capture dissociation
EDTA	ethylenediaminetetraacetic acid
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
FAIMS	Field asymmetric waveform ion mobility spectrometry
HCl	hydrogen chloride
HDMS	High definition mass spectrometry
ICR	Ion cyclotron resonance
IEM	Ion Evaporation Model
IMS	Ion mobility spectrometry
IPTG	isopropyl β-D-1-thiogalactopyranoside
IRMPD	Infrared multiphoton dissociation
LB	lysogeny broth
MALDI	Matrix assisted laser desorption ionisation
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
nLC	nano Liquid chromatograhy
nESI	nano Electrospray ionisation

PBS	phosphate buffered saline
PIC	Pathways in Competition
PTMs	Post translational modifications
QIT	Quadrupole ion trap
QTOF	Quadrupole time of flight
RF	Radio frequency
RPLC	Reversed phased liquid chromatography
SRIG	Stacked ring ion guide
TIC	Total ion current
T-Wave	Travelling Wave
TW-IMS	Travelling wave ion mobility spectrometry
UV	Ultraviolet

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENT

Abstract

The University of Manchester

Ross Chawner

Combined Tandem Mass Spectrometry and Ion Mobility Spectrometry in Proteome Analyses

Doctor of Philosophy

2012

Proteomic studies aim to identify, quantify and characterise the full complement of proteins in a cell or organism under a defined set of conditions, and are important to our understanding of cellular mechanisms. However, such studies represent a major analytical challenge. A typical proteome analysis involves enzyme-mediated digestion of complex protein mixtures to yield an even more complex mixture of peptides. Combined reverse-phase liquid chromatography and tandem mass spectrometry is then traditionally utilised to ascertain sequence information from the characteristic peptide sequences. Analytical data derived for the peptides are employed as search terms in database searching of protein sequences derived from gene sequences. The extreme complexity of the peptide mixtures analysed means that additional novel approaches are required to fully interrogate the vast number of tandem mass spectra generated, assigning peptide identity and thereby helping to address demanding biological questions. The research reported here aims to further our understanding of both gas phase peptide/peptide fragment ion structure and peptide fragmentation behaviour using a combination of tandem mass spectrometry and ion mobility measurement.

To facilitate the determination of peptide ion collision cross section, a novel standard, QCAL-IM, produced using the QconCAT strategy, has been developed to enable calibration of drift time in Travelling Wave Ion Mobility instruments. The standard facilitates empirical determination of the rotationally averaged collision cross section of any peptide/peptide fragment ion that lies within the calibration range encompassed. QCAL-IM was subsequently utilised to determine the collision cross section of a range of peptide ions produced by Lys-C and Lys-N proteolysis of 'standard' proteins. Data produced allowed the effect upon gas phase ion conformation through changing the location of the basic residue lysine within a peptide sequence to be assessed.

The fragmentation behaviour of peptide ions produced by a variety of digestion regimes during both collision-induced dissociation (CID) and electron transfer dissociation (ETD) has also been extensively studied. The proteases trypsin and Lys-C are those typically utilised during proteomic studies and peptides produced by each have either the basic residues arginine or lysine at their carboxy-terminus. Secondary enzymatic treatment with the exoprotease carboxypeptidase B cleaves these basic residues from the *C*-terminus. Tandem mass spectrometric analysis of both tryptic/Lys-C peptides and their CBPB truncated analogue highlights that the dominant fragment ion series observed during both CID and ETD is determined, at least in part, by the location of such basic residues.

Finally, studies were undertaken to investigate the factors which may promote/inhibit scrambling of peptide fragment ion sequence, which has recently been shown to take place during CID. The effect of modifying the gas phase basicity of the *N*-terminal amino acid residue is studied through a combination of derivatisation and synthesis of alternative peptide sequences. Increasing the gas phase basicity is shown to inhibit the observed sequence scrambling while promoting concomitant rearrangement/retention of a carboxyl oxygen at the *C*-terminus to give enhanced formation of b_n+H_2O product ion species.

Acknowledgments

I wish to thank Dr. Claire Eyers and Professor Simon Gaskell for their support and guidance through the PhD programme and for the opportunity to learn from their wealth of knowledge and experience. I've enjoyed every minute studying in the Michael Barber Centre (obviously apart from when the science isn't working) and for that I thank colleagues both past and present for creating a great working environment and for their advice and friendship.

I would also like to thank the many friends that I have made from other research groups within the MIB and the wider University of Manchester who have made my years of study extremely enjoyable.

Thanks are also extended to Dr. Perdita Barran at the University of Edinburgh and Professor Carol Robinson at the University of Oxford for allowing access to their Ion Mobility Instrumentation. I'd also like to acknowledge Waters Corporation for funding and Kevin Giles for guidance regarding Ion Mobility.

Last but not least I thank my family for their continued support, encouragement and love, without them I wouldn't have dared to start the PhD programme in the first place. In particular I thank you mum, you are a truly wonderful person. Finally, I dedicate this thesis to my sister who although no longer with us is a constant source of inspiration and without whom I would never have started along the scientific path in the first place.

Chapter 1

Introduction

1.1. Proteomics

Proteomics is the study of the proteome, which is the complete array of proteins expressed by the genome. The field has previously aimed to identify, characterise and quantify the range of proteins expressed by a cell during its lifetime. In recent times however, proteomics has expanded and now encompasses a much broader range of analyses. Post-translational modifications (PTMs), protein isoforms, protein-protein interactions and protein complexes are now also investigated both in a qualitative and quantitative manner [1].

Mass spectrometry (MS) has become a pivotal analytical technique for proteomic studies; individual proteins within a complex mixture can be identified and/or quantified through matching of peptide fragmentation spectra, traditionally produced by collision-induced dissociation (CID), to known protein sequences [2]. In general, bottom-up studies are conducted where peptides rather than proteins are subjected to tandem mass spectrometry (MS/MS) for sequence elucidation, as greater sensitivity can be achieved and complications arising from protein modification are minimised [3]. Such modifications cause a combinatorial 'explosion' giving a wide variety of possible isoforms, hence making data analysis highly complicated.

Protein identification, particularly large-scale automated analysis, requires sequence information and historically a peptide of up to approximately 20 amino acid residues in length is optimal for such MS analyses [4]. Proteins are therefore digested by action of a protease, commonly trypsin, to give a set of polypeptides for analysis [5]. Enzymes used for cleavage of a protein tend to be sequence specific and create peptides in the desired mass range (~500 to 2000 Da). Action of trypsin is at the carboxy-terminal of arginine and lysine residues, the enzyme will cleave at these sites except when followed by the imino acid, proline. Peptide chains generated from trypsin digestion provide MS/MS spectra that are easily interpretable. Enzymes that are less specific in their action (e.g. pepsin) tend to be avoided as spectra obtained from such digests can become confused due to many overlapping analyte responses. While proteolysis overcomes problems associated with the analysis of large, differentially modified proteins by generating peptides for easier analysis of primary structure, the drawback of protein digestion is that sequence coverage is compromised; only a

small proportion of the total protein sequence is evaluated. Identification of the protein of interest is possible but often at the expense of complete characterisation and many features of the protein sequence, such as PTMs remain undefined [6].

More recently, the increasing availability of electron driven fragmentation techniques such as electron capture dissociation (ECD) [7] and electron transfer dissociation (ETD) [8] has enabled larger peptides, such as those produced by the proteases Lys-C and Lys-N (cleaving *C*-terminal and *N*-terminal to lysine residues respectively) and even intact proteins to be readily studied by MS. Top-down proteomics [9, 10] is a promising research area where analysis of intact proteins can be conducted with sequence information generated by these electron mediated dissociation techniques. This approach, when successful, enables identification of all protein modifications while giving enhanced sequence coverage when compared to bottom-up proteomic approaches. An abundance of data is obtained from a single analysis, meaning time consuming digestion protocols and sample preparation can be avoided.

1.2. Data Analysis and Related Problems

Protein identification within a complex mixture is achieved through submission of MS/MS data to a search engine. Examples of these search tools are MASCOT [11], SEQUEST [12] and X!Tandem [13], each of which works in a similar manner albeit with different algorithms; protein/peptide fragmentation data is compared to known sequences and a list of matching candidate species is returned, ranked according to the strength of the match. This includes the number of characteristic fragment ions and the sequence coverage, i.e. the degree to which the observed masses correspond with the theoretical masses. Each MS/MS spectrum is ascribed a score, determined by the extent to which the experimental/theoretical data match and an associated probability factor is also determined. The problem with such search engines is that only a small proportion of the data present within a given mass spectrum is utilised. Fragment intensities are neglected and recent work [14, 15] suggests that sequence scrambled species can be present within certain analyses. Greater understanding of the underlying mechanistic principles of peptide dissociation would significantly enhance interpretation of MS/MS spectra. Knowledge of fragment intensity/peptide rearrangement trends would facilitate more detailed spectral elucidation, enabling data directed database searches to be conducted, and the research discussed in Chapters 4 and 6 of this thesis focuses on this topic.

1.3. Reversed Phase Liquid Chromatography

Reversed phase liquid chromatography (RPLC) is a separation technique that can be employed in a preparative manner for purification, isolation or separation of analytes [16]. However, in the case of the work presented in this thesis, RPLC is used to reduce sample complexity in-line with MS analysis [17]. A hydrophobic stationary phase is used alongside an initial hydrophilic mobile phase; bound analytes are eluted by increasing the hydrophobicity of the mobile phase, typically through the application of a solvent gradient; hence shifting the equilibrium of the analytes from the stationary phase to the mobile phase [18]. Polar analytes are therefore eluted first with non-polar species having greater affinity for the stationary phase. Peptides produced by tryptic or Lys-C proteolysis of a protein contain the amino acids arginine or lysine and the side chains of these residues provide hydrophobic groups which strongly interact with the stationary phase and assist chromatographic separation in bottom-up proteomic approaches. The stationary phase to which they bind is composed of inert beads whose surfaces are covalently bonded to aliphatic chains of differing length; these chains are denoted C2-C18 with the number detailing the number of CH_2 groups in the chain.

Solvents typically used for elution of analytes in RPLC are methanol and acetonitrile as they are less polar than, but perfectly miscible with water. Elution of analytes can be performed in an isocratic manner (where the water:solvent ratio is maintained for the duration of the run) or more commonly through application of a gradient (where the ratio of solvent:water is increased at a defined rate during the separation). The pH of the mobile phase is also critical and can influence the extent to which a species binds to the column and hence the observed elution time/selectivity of separation [19]. Addition of an acidic ion-pairing agent such as formic acid is often used to promote binding and maintain the pH, enhancing protonation and consequently increasing sensitivity in positive ion electrospray ionisation (ESI). Formic acid is ideal for use with ESI-MS as it is volatile and does not interfere with the ionisation of analytes [20].

1.4. Ionisation Techniques

The two most commonly used ionisation techniques for protein and peptide analysis are electrospray ionisation (ESI) and Matrix Assisted Laser Desorption Ionisation (MALDI).

1.4.1. Electrospray ionisation

ESI is a 'soft ionisation' technique and energy is not imparted to analyte species during the process meaning that fragmentation is not observed [21]. The approach, developed for use with macroions in 1968 by Dole *et al.* [22] operates at atmospheric pressure and can give singly or multiply charged ions. The Nobel Prize for chemistry was awarded to John Fenn in 2002 for his work on ESI of biological macromolecules [23]. Development of the technique has enabled non-destructive gas phase ion production of analytes with high molecular weight and low vaporisability [24, 25]. A capillary tube held at a high voltage is used to deliver sample solutions at low flow rates (typically μL min^{-1}) into a spray chamber. Flow rates lower than this are employed for nESI where solvent is delivered to the capillary at nL min^{-1}. A counter electrode is positioned at the back of the spray chamber creating an electric field between the capillary and the electrode. As a result, the liquid experiences this field and when the capillary is at a positive potential, ions build at its tip. Both positive and negatively charged gas phase ions can be produced by ESI and analysis of either cationic or anionic species is facilitated by reversing the polarity of voltages applied to the mass spectrometer. In general the ESI process is viewed a vehicle for the delivery of pre-formed ions in solution to a mass spectrometer and the process does not initiate ion formation [26]. Experimental evidence detailing that this is not always true exists [27, 28] although such ion production is not considered to be a major process by which ions are observed following ESI.
Upon exiting the capillary, the sample solution produces a 'Taylor Cone' (Figure 1) [29]; conical in shape, the cone forms when an electrically conducting liquid is exposed to an electric field. Once a threshold voltage is reached, inversion of the cone results and a jet forms. Increasing electrostatic force is present within the liquid jet and when this force exceeds that of the surface tension droplet formation occurs via a process known as budding [26].The droplets formed are highly charged so repel one another while being simultaneously attracted to the counter electrode. During this period a curtain of heated nitrogen is directed into the ion transit region assisting solvent evaporation and thereby reducing droplet size. As the droplet shrinks, positive charges accumulate at the surface and Coulombic repulsion between them increases due to their close proximity. Eventually, the 'Rayleigh Limit' is

reached and the electrostatic repulsion between charges overcomes the force of the surface tension. Droplet fission results and is accompanied by production of either solvated ions or smaller droplets which repeat the surface ionisation process (Figure 1).

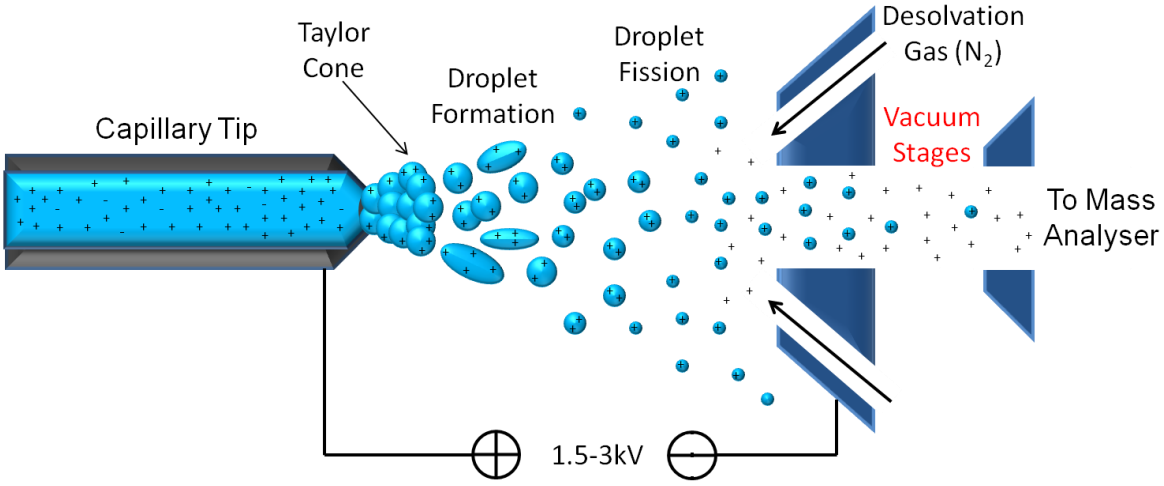

Figure 1. Schematic of a typical ESI source detailing positive ion formation, if the polarity of the voltages is switched then analysis of negative ions can be performed.

Ions present in solution are thus transferred to the gas phase; two mechanisms have been proposed by which this liquid phase to gas phase transfer takes place (Figure 2), the Charged Residue Model (CRM) and the Ion Evaporation Model (IEM). The CRM was proposed by Dole and co-workers [22, 30] and details repeated solvent evaporation to give a charged species complexed with a few molecules of solvent. Following evaporation of the remaining solvent and concurrent retention of charge by the analyte, an isolated gas phase ion results [25]. In contrast, the IEM, proposed by Iribarne and Thomson [31, 32], describes direct desorption of a charged analyte from the solvent droplet at the point at which the electrostatic force is sufficient to overcome solvation. There is much debate as to which model typically applies; however it is accepted that the nature of the analyte may be the determining factor. In particular, highly charged macromolecules are thought to disfavour the IEM due to a large activation barrier preventing direct desorption from the solvent droplet [33]. In reality, a combination of the models is likely to be applicable toward the formation of gas phase ions [34].

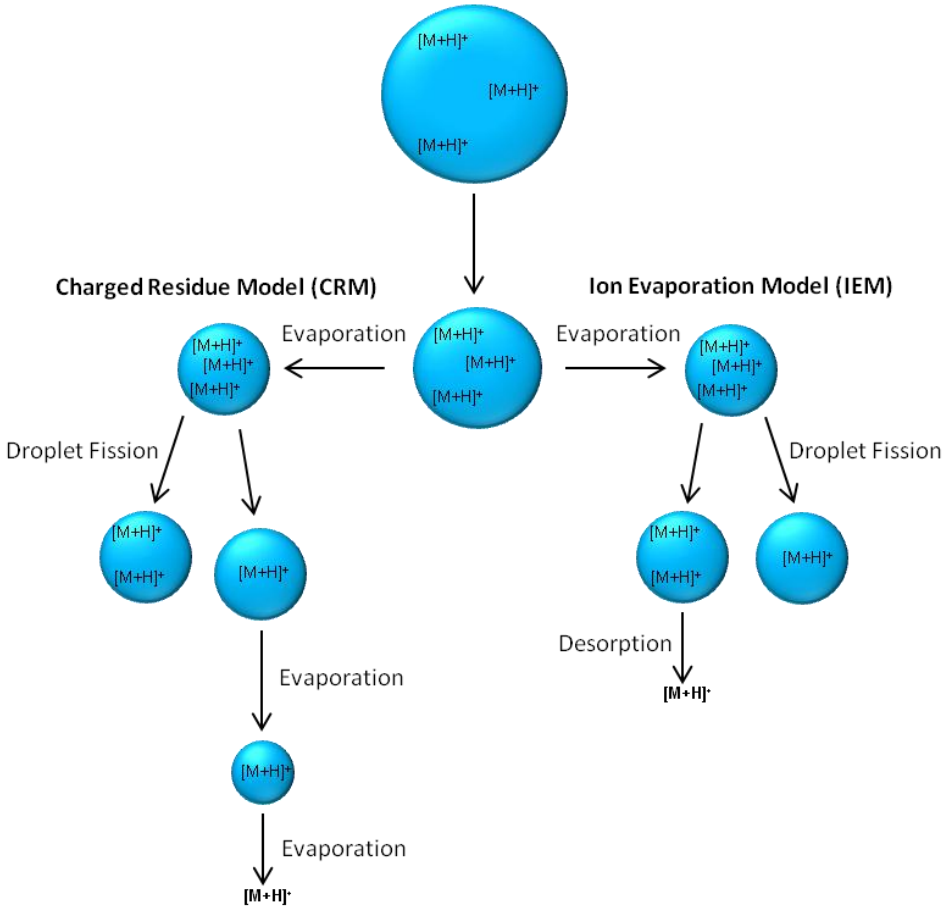

Figure 2. Formation of positively charged ions by either the CRM or IEM models for ESI.

1.4.2. Matrix Assisted Laser Desorption Ionisation (MALDI)

MALDI (Figure 3) for use with proteins above 10 kDa was pioneered by Karas and Hillenkamp [35]. Samples are co-crystallised onto a target with a large excess of UV-absorbing matrix which is usually an organic acid. This crystallised mixture is then subjected to a short irradiating LASER pulse causing sublimation of the matrix and by association the analyte, with each entering the gaseous phase. The process is facilitated by the presence of a strongly absorbing chromophore in the acid and this feature means the majority of the LASER energy supplied is efficiently transferred to the matrix and not the analyte [36]. MALDI is therefore a 'soft ionisation' technique; minimal energy is transferred to the analyte and fragmentation of the species of interest is (largely) avoided. The ionisation mechanism involved in MALDI is not fully understood, however, the generally accepted theory is protonation/deprotonation of an analyte molecule via transfer of a mobile proton during collisions between the analyte and photoionised matrix [37].

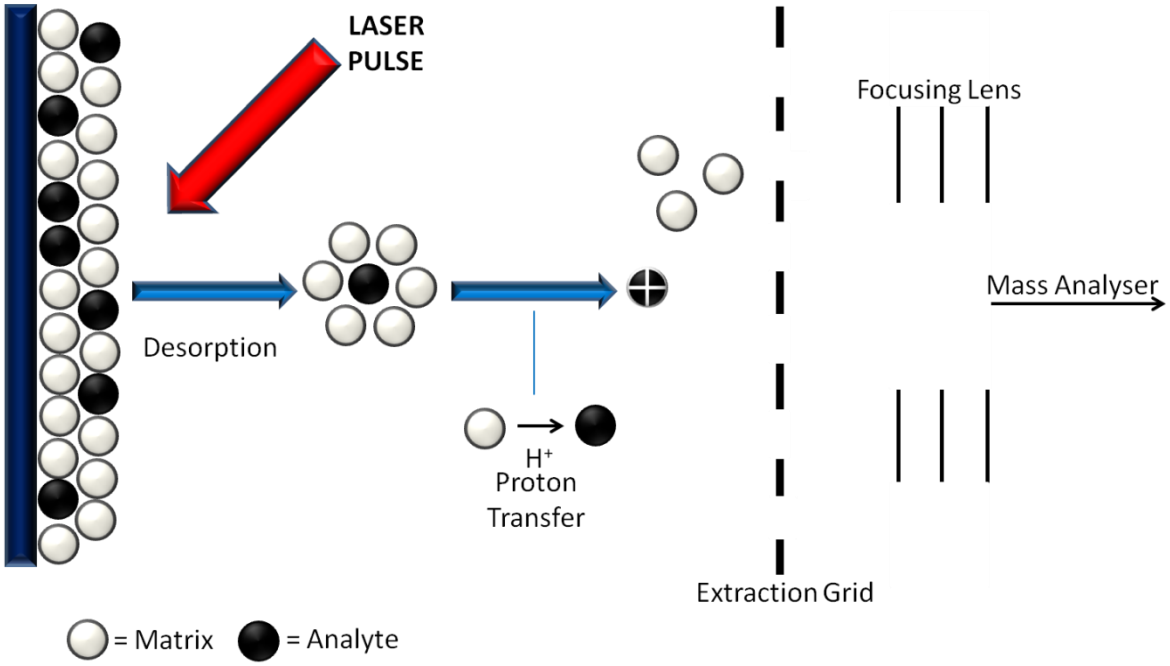

Figure 3. Diagram to illustrate the production of gas phase ions by MALDI.

1.5. Mass Analysers

1.5.1. Quadrupole Mass Analyser

Quadrupole mass analysers utilise the stability of an ion under the influence of an oscillating electric field to discriminate ions on the basis of *m/z* ratio. The theory of a quadrupole device was introduced by Paul and Steinwedel in 1953 [38]. The analyser consists of four parallel circular (or hyperbolic) rods spaced around a central axis (Figure 4). The rods are electrically connected in opposing pairs and both sets have an applied direct current (DC) and an alternating current (AC). The AC voltage applied to one set is 180° out of phase with that applied to the other set, therefore one set is positive with respect to ground potential and the other is negative. The applied AC constantly switches, meaning that each pair of rods is successively positive and negative and this focuses the ions into the centre of the instrument. Ions are periodically attracted to and then repelled by the two pairs of rods as a consequence of the AC, with a zero net effect in the x- and y-planes, thus permitting transmission of all ions through the rods. The DC then provides mass selection, with application of the appropriate amplitude resulting in a particular species being destabilised in its motion,

causing collision with the rods [39]. Ions entering along the z-axis are subject to a total electric field:

$$\phi_o = + (U\text{-}V \cos \omega t) \quad (1)$$

$$-\phi_o = - (U\text{-}V \cos \omega t) \quad (2)$$

Where ϕ_o is the potential applied to the rods

U is the DC potential

ω is the angular frequency of the AC voltage (rad s^{-1})= $2\pi\upsilon$

υ is the frequency of the AC voltage

V is the peak-to-peak amplitude of the AC voltage

t is time.

An ion entering the quadrupole maintains its velocity along the z-axis but is under the influence of accelerating forces in the x- and y-directions. As long as the force applied in each direction is not substantial enough to propel the ion a distance of r_o (the distance to collision with a rod) then it will maintain a stable trajectory and transmission through the quadrupole results (Figure 4).

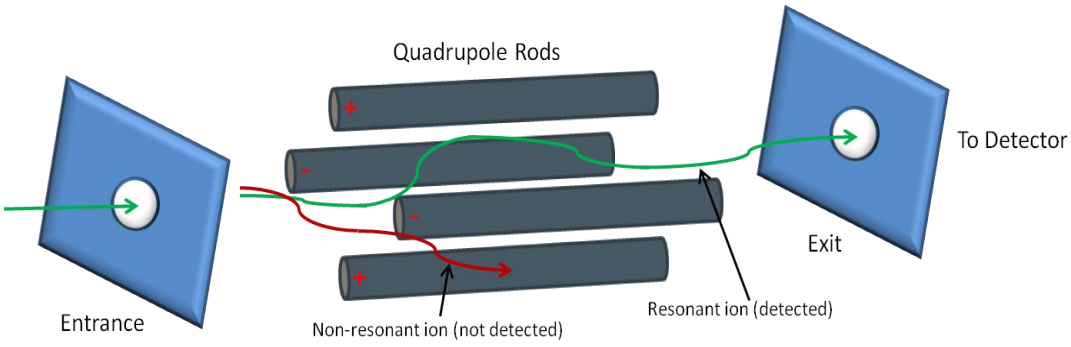

Figure 4. Schematic illustrating the arrangement of the four cylindrical rods within a quadrupole mass analyser. The motions of resonant/non resonant ions through the instrument are also shown.

For any quadrupole, r_o and ω are fixed and transmission across the entire mass range is facilitated by ramping the applied voltages according to a linear relationship called a Scan function. The slope (i.e. the magnitude of U against V) of the Scan function is termed the quadrupole Gain and the initial applied DC is called the Offset. Figure 5 illustrates these features alongside the stability regions for three ions of differing *m/z*. A stability region represents the range of U/V amplitudes that can be applied where transmission of the ion in question will result. The stability region of a given ion may overlap with that of another and to ensure that a particular ion is transmitted the line representing U/V must lay close to the apex of the region. If the line misses the apex then transmission will not result and consequently higher values of U and V allow ions of larger mass to be observed. Increasing the quadrupole Gain gives improved spectral resolution but decreased sensitivity. This phenomenon is of greater consequence toward larger ions due to the larger area of the associated stability region. An increase in the DC offset will also give greater resolution at the expense of instrument sensitivity, however in this instance all masses are affected to the same extent.

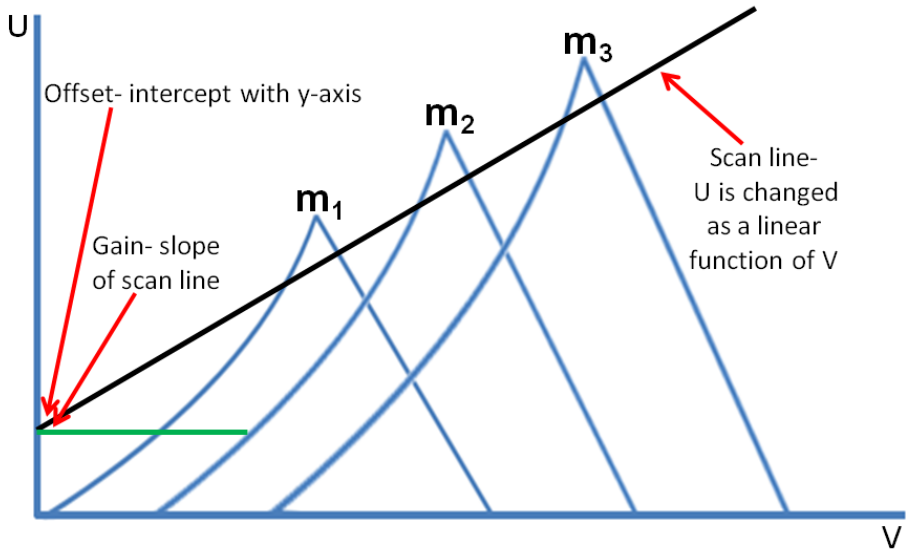

Figure 5. Stability diagram of consecutive *m/z* values ($m_1<m_2<m_3$) in a quadrupole mass analyser, changing U as a function of V enables consecutive *m/z* ions to be observed. The steeper the slope of the Scan line, the greater the resolution. However, if the line does not pass through the stability region of a particular *m/z* it is not transmitted.

1.5.2. Quadrupole Ion Trap

A QIT consists of a hyperbolic ring electrode with two capping electrodes, within which can be applied a quadrupolar electric field (Figure 6). Ions of varying masses are held within the trap using an AC potential and can be ejected from the trap according to their *m/z* value by varying the amplitude of this potential. A background gas, typically helium, is also present; this is used to prevent ion losses upon expansion of the trajectory of motion. Collisions with the helium help to remove excess ion energy, causing a dampening effect toward the centre of the trap which in turn helps to maintain a stable ion trajectory and increases instrument sensitivity [40].

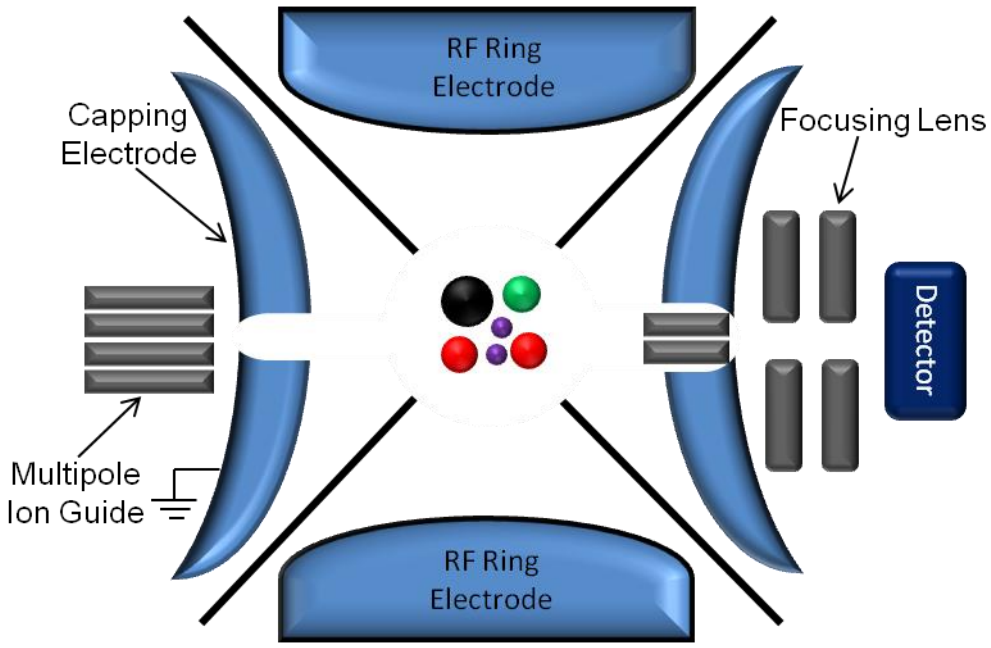

Figure 6. Schematic of a typical QIT, helium background gas is present helping to maintain a stable trajectory of the ion cloud thus increasing instrument sensitivity and mass resolution. Diagram adapted on the 15th August 2012 from that found at www.ctu.edu.vn/~dvxe/Bioinformatic%20course/mod2/mod2_0.html

The trap operates through application of an AC voltage to the ring electrode, no DC component is present. The end cap electrodes are held at or near ground potential. Ions inside the trap are confined by the AC voltage and take on a trajectory in a 'figure of eight'[41]. In the absence of a DC voltage the stability area operates along the AC axis only. Typically, the applied voltage has a set frequency but variable amplitude; this means that if V is increased then ions can be ejected from the trap by Instability Ejection [42]. Stability diagrams at initial

(V1) and increased (V2) amplitude are shown in Figure 7, where the ion at lowest *m/z* is ejected from the trap first.

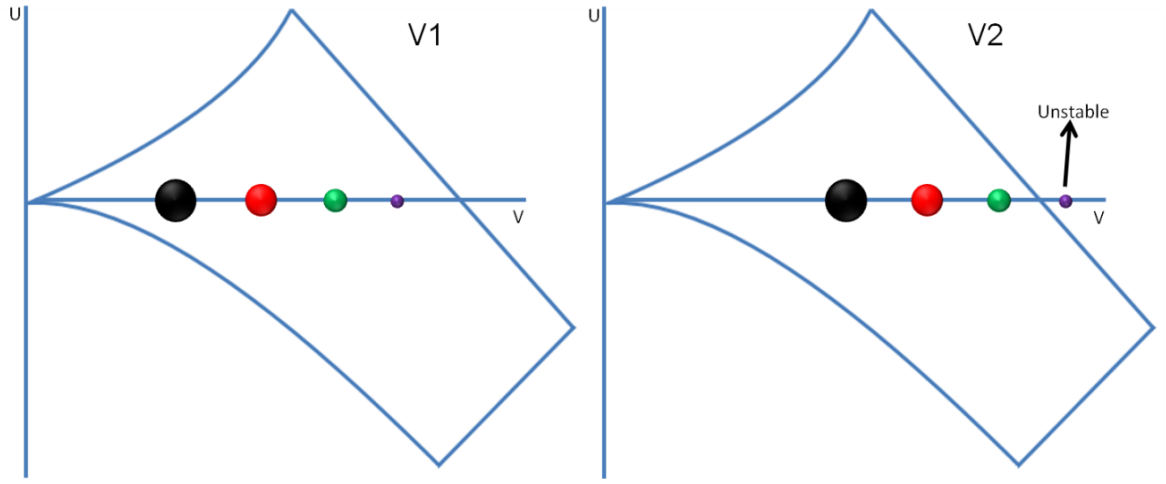

Figure 7. Stability diagrams at initial applied AC voltage (V1) enabling all ions to remain confined within the QIT mass analyser. An increase in amplitude (V2) causes all ions to move along the x-axis and when this value exceeds the stability limit of a given *m/z* the ion is expelled from the trap in the z-direction.

Ions can also be ejected from the trap by Resonant Ejection; here the secular frequency at which an ion oscillates within the trap is utilised for mass analysis. This frequency can be controlled through application of an appropriate value of V at the ring electrode. If the same AC conditions are applied to the capping electrodes it is possible to generate ion motion that is in resonance; causing an increase in the magnitude of its oscillations. When the applied amplitude reaches a required level the ion becomes destabilised and is ejected from the trap (Figure 8).

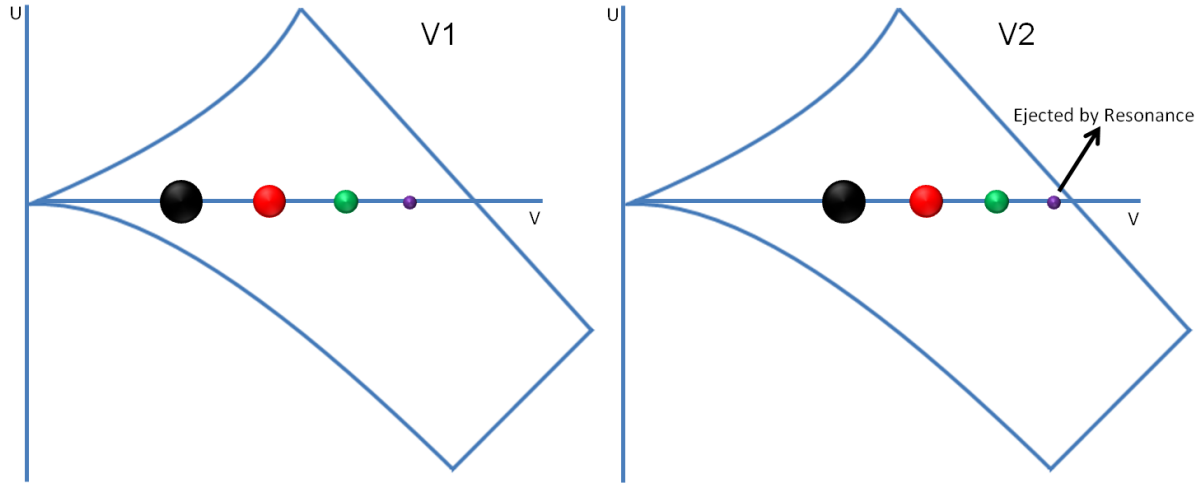

Figure 8. Stability diagrams demonstrating Resonant Ejection from a QIT. At the initially applied voltage (V1) ions are stored within the QIT; no AC voltage is applied to the capping electrodes and all ions are stable within the mass analyser. An AC voltage is then applied to the capping electrodes and V is increased moving the ions along the x-axis, to a value (V2) which corresponds to that of the capping voltage. Ejection takes place at a lower amplitude of V than is required for Instability Ejection.

The total number of ions allowed to enter a QIT is controlled to help minimise space charge effects [43]. Such effects can influence both Instability and Resonant Ejection; the phenomenon occurs when the number of ions present within the QIT is too great. Species positioned at the extremities of the ion cloud begin to shield those toward the centre. Consequently, the effective electric field acting upon these ions is reduced and the shape of the stability diagram is modified. Higher values of V are required to reach the ejection point and this can lead to errors in mass analysis and loss of resolution.

1.5.3. Time of Flight

TOF mass analysers are ideally suited for use with the pulsed production of ions associated with MALDI instrumentation. They analyse the flight time of packets of ions through a field free region that enter the instrument at a specific time point. There are two types of TOF mass analyser; Linear and Reflectron.

1.5.3.1. Linear TOF

Ions are delivered to the mass analyser in discrete packets; this can be achieved by pulsed ion production or due to transient application of a voltage to a focussing lens. They enter the analysis region and are accelerated by a potential V (Figure 9). The ion packet enters the field free region, where it is not subject to any accelerating or steering forces and progresses toward the detector, travelling a distance, *d*.

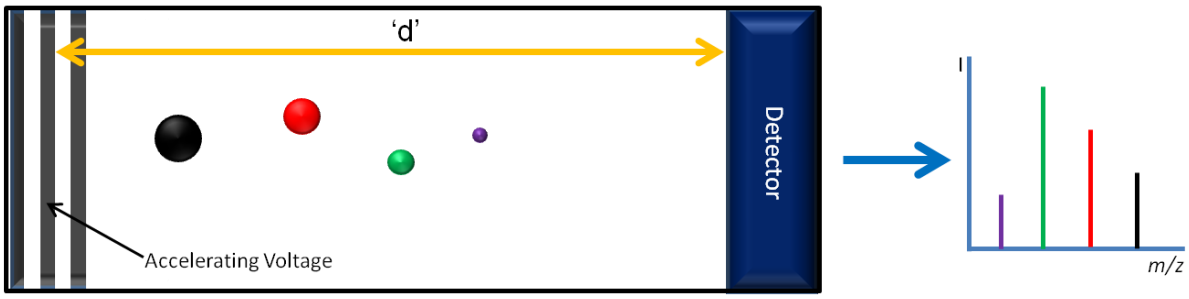

Figure 9. Schematic of a Linear TOF mass analyser.

Time taken to traverse the field free region is unique to a particular *m/z* species and ions consequently arrive at the detector in rapid sequence. The analyte mass to charge ratio is calculated through measurement of the time taken to traverse the field free region between source and detector. Time taken to travel the distance *d* can be calculated using the simple equation [39]:

$$t = d/\mathbf{v} \qquad (3)$$

where t is the time taken to reach the detector

and **v** is the velocity

As the distance travelled and accelerating voltage are constant for each species, the factor that influences flight time is solely *m/z*, consequently this value can be calculated using the following formula:

$$t^2 = \frac{m}{z} \frac{d^2}{2Ve} \qquad (4)$$

where *m/z* is mass to charge ratio

V is the accelerating voltage

and e is the elementary charge

The advantage of TOF mass analysers over other instrumentation is that the entire mass spectrum is recorded for each analysis. This is achieved without having to scan voltages and consequently analysis time is reduced, in turn enhancing duty cycle. In addition, ions do not pass through complicated optic lens systems and efficient ion transmission from source to detector results [44]. A major problem with Linear TOF analyses is that the achievable mass resolution is poor; uniform flight times for ions of the same *m/z* ratio are not observed. Time distribution caused by the length of ionisation pulse generates an extended ion packet, a problem which is further compounded by spatial distribution within the ionisation source [45]. Each ion enters this region with the same initial kinetic energy but is accelerated from a different location within the extraction grid. Ions toward the back of the source experience a larger potential gradient and are accelerated to a higher kinetic energy than those at the front. As a consequence there is a broadening of the signal observed in Linear TOF instruments and a reduction in the ability to distinguish between species of similar mass. The achievable resolution can be improved through use of delayed pulsed extraction rather than the continuous ion extraction discussed thus far. Pulsed extraction minimises kinetic energy spread between ions of the same *m/z*. A time delay between ionisation and extraction is incorporated into the analytical sequence and during this delay species travel into a field free region according to their differing kinetic energies. Ions of greater energy (but same *m/z*) travel closer to the detector during this time and therefore ions of initial lower kinetic energy receive a greater level of acceleration due to their longer occupation of the source. The ions travel at their new respective velocities with the previously lower energy species passing the initial higher energy population. The point at which the two populations converge is known

as the ‘space focus’ and optimisation of the pulse amplitude and time delay employed is necessary to ensure that this occurs as close to the detector as possible [46].

1.5.3.2. Reflectron TOF

Reflectron TOF is a sophisticated technique for correction of kinetic energy spread; it was developed by Mamyrin *et al.* [47] and uses an ion mirror to compensate for these differences in kinetic energy, deflecting the ions and causing them to traverse the flight tube a second time. The total kinetic energy of the ion is maintained, while the direction of the motion is reversed. This switch in direction is controlled by voltages applied to electrodes located between the pusher and detector (Figure 10). The back of the reflectron is at highly positive potential with respect to the entrance (when operating in positive ion mode); ion motion is slowed to a zero value before acceleration back to the initial velocity toward the detector. Greater penetration of the reflectron field occurs for ions of higher kinetic energy and consequently such species have a longer flight path and reach the detector at a time equal to that of the less energetic species of the same *m/z* ratio [48, 49].

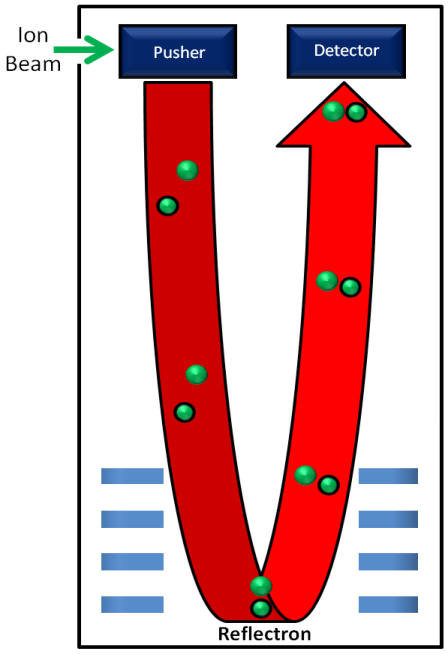

Figure 10. Illustration of the flight path of an ion during analysis using a Reflectron TOF mass analyser. Shown are two ions of the same *m/z* but with initial higher (black outline) and lower (green outline) kinetic energy. The higher energy ion penetrates the reflectron field to a greater extent and with appropriate selection of voltages compensation for the energy difference is achieved meaning the ions reach the detector at the same time.

1.6. Tandem Mass Spectrometry

MS/MS experiments are those incorporating two (or more) sequential stages of mass analysis. Typically, the initial stage of an MS/MS experiment is selection and isolation of a precursor ion of interest, this ion is then activated causing decomposition into diagnostic fragmentation products, prior to a second stage of mass analysis. The number of stages involved in such an MS/MS experiment can be increased when using an ion trapping mass analyser enabling MS^n experiments, where the power (n) denotes the number of sequential steps of isolation, fragmentation and mass analysis. MS^n experiments allow highly detailed structural information to be obtained.

MS/MS experiments can be conducted in either the dimension of space or time [50]. Those conducted 'in space' require multiple analysers connected in series, each performing a stage of mass analysis. A QTOF mass spectrometer is an example of such an approach, whereby a quadrupole mass analyser is utilised for precursor ion selection, a fragmentation step follows and detection of the product ions generated is performed by application of a second stage of mass analysis using a TOF analyser. Such instruments typically incorporate ion bridges and complex steering lenses to enable efficient transmission of ions from one section of the instrument to the next. MS/MS analysis performed 'in time' can be undertaken using QIT or ion cyclotron resonance (ICR) instruments and here analysis events occur within the same volume but separated by a defined period of time. MS^n experiments of higher order are best suited to such instruments as multiple mass analysers are not required.

A tandem quadrupole mass spectrometer is a commonly used 'in space' instrument that is a particularly powerful platform upon which MS/MS can be performed. The instrument incorporates two quadrupole mass analysers separated by a Radio Frequency (RF)-only collision cell (formerly a quadrupole, now usually a hexapole). Such instruments are extremely versatile as a range of scan modes can be employed, each enabling acquisition of informative MS/MS data; the four main analysis regimes are as follows:

1. Product ion scan: a precursor at a specific *m/z* ratio is selected and all the product ions produced by fragmentation recorded.

2. Precursor ion scan: a product ion is chosen and the precursor from which it is produced is determined. The first mass analyser scans all precursors while the second mass analyser selects solely the product ion of interest. This acquisition mode cannot be performed in the time domain.

3. Neutral loss scanning: decomposition resulting in the loss of a neutral fragment ion of interest is monitored. Both mass analysers are scanned; however there is a mass offset between the two. Detection only occurs when a fragment ion is present corresponding to loss of the defined neutral fragment from the precursor ion.

4. Selected reaction monitoring: neither mass analyser is in scan mode, instead the mass analysers are set to stably transmit specific precursor and product ions of defined *m/z* value. This analytical strategy can be very specific and as such the associated duty cycle is increased over other modes of operation meaning the limit of detection is significantly improved.

MS/MS is used in the analysis of peptides to produce sequence specific data; cleavage of covalent bonds across the peptide backbone is induced according to mechanistic principles of the fragmentation technique employed. A characteristic product ion series results and when combined can be used to elucidate the original amino acid sequence. Roepstorff and Fohlman [51] originally proposed a nomenclature for the naming of these fragmentation products and this was later modified by Biemann [52]. Figure 11 illustrates the peptide bond cleavages that result in generation of the fragment ions addressed by the nomenclature. Charge retention by *N*-terminal species results in observation of b-, c- and a-ions whereas charged fragments originating from the *C*-terminal portion of the backbone are denoted as y-, z- and x-ions [53].

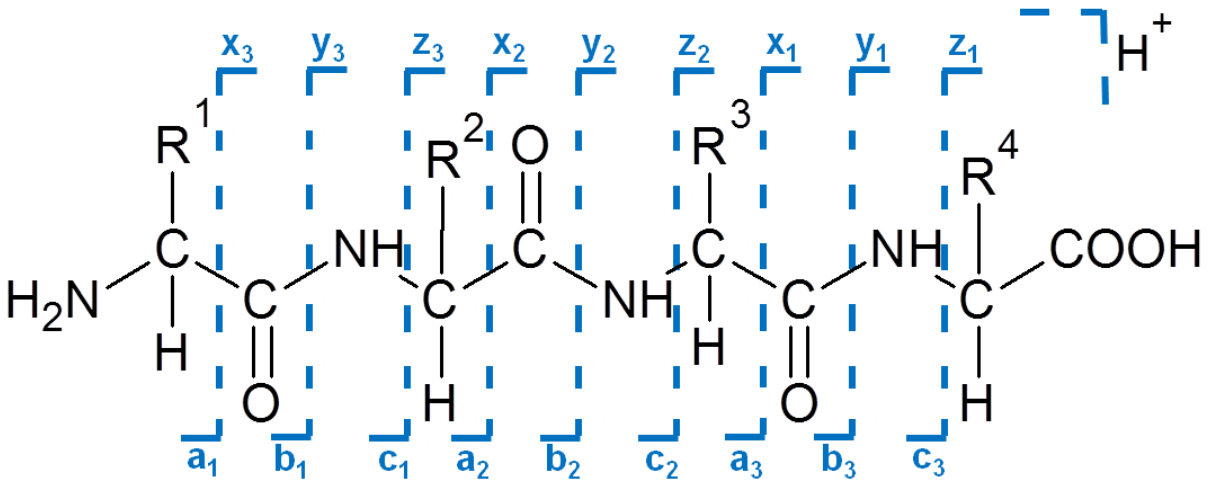

Figure 11. Nomenclature for product ions resulting from peptide backbone fragmentation. Those incorporating the *N*-terminus are labelled b-, c- and a-ions and *C*-terminal species are termed y-, z- and x-ions.

High quality MS/MS spectra enable 'ladder sequencing' to be performed, here each b- and y- (or c- and z-) ion can be sequentially identified, with adjacent fragments separated by the mass of one amino acid. This technique enables determination of amino acid sequence, however, in practise, the entire range of ions is rarely observed. A particular ion may not be generated at a detectable level or two fragmentation events may occur on the same ion to give internal cleavage ions [54]. Extensive fragmentation of the central peptide backbone can generate immonium ions and these correspond to an individual amino acid, providing confirmation of the presence of a particular residue within the sequence [55]. In addition, certain fragmentation pathways are favoured over others, for example, the b_1 ion is never observed as dissociation between the two *N*-terminal residues is not facile by the mechanism of b-ion formation [56].

1.7. Fragmentation Techniques

1.7.1. Collision-Induced Dissociation

Ions produced by 'soft ionisation' are typically of three types; ions which have a long lifetime and are detected prior to fragmentation (decomposition rate k of $<10^5$ s^{-1}), ions with a short lifetime ($k >10^6$ s^{-1}) which fragment prior to detection and metastable ions which have an intermediate lifetime (10^5 $s^{-1} < k < 10^6$ s^{-1}); such species can be mass selected by the first mass analyser while possessing sufficient energy to undergo decomposition prior to arrival at the detector [57]. Metastable ions are ideal for MS/MS experiments; however are present in low abundance making analysis difficult. Collision-Induced dissociation (CID) is a method used to generate metastable character in ions that would otherwise have an extended lifetime. The approach was pioneered by K.R. Jennings [58] and F.W. McLafferty [59] and sees a greatly increased probability of fragmentation when compared to unimolecular metastable decomposition.

During the process of CID, analyte ions are forced to undergo inelastic collisions with a target (often an inert gas) [60] resulting in conversion of translational energy into internal energy [61]. The ion is consequently promoted to an excited state and when the bond energy is exceeded by the internal vibrational energy decomposition results. Fragmentation by CID can be classified as either low energy [62] or high energy [63], with differing product ions expected according to the kinetic and thermodynamic characteristics of the process. Typically, low energy experiments are performed on triple quadrupole, QIT, and ICR instruments; complicated fragmentation pathways are common as product ions often undergo rearrangement [64, 65]. High energy CID is traditionally performed with TOF and electromagnetic mass analysers although has more recently been implemented in the hybrid Orbitrap series of instruments [66]; fragmentation follows less complicated pathways and simpler spectra, rich in sequence information, typically result [67].

Decomposition of peptide ions by low energy CID is known to result in the observation of differing fragmentation products according the instrument upon which analysis is being performed [68]. A particular fragment ion will only be observed if i) decomposition to generate the species is energetically facile and ii) the product ion is stable to the conditions which it encounters post-dissociation. Linear multipole collision cells such as those found in tandem quadrupole instruments are modified quadrupole devices which have an applied RF

voltage but no DC component. The RF voltage assists the transmission of activated precursor ions and fragments by compensating for the deviation in flight path resulting from analyte-target collisions. Application of this voltage also provides directional focussing throughout the length of the cell ensuring that multiple ion-neutral collisions occur and the efficiency of CID is high. The caveat to this is that although extensive y-ion coverage is frequently seen, typically only the lower members of the b-ion series are detected [69]. *N*-terminal b-ion fragments resulting from CID of a tryptic precursor ion are less stable to secondary fragmentation than their y-ion counterparts. The y-ion series incorporates a highly basic *C*-terminal arginine/lysine residue which for singly protonated species sequesters the proton accommodated within the structure. Consequently, this proton is not 'mobile' reducing the observation of secondary fragmentation. Conversely, the *N*-terminal fragment ion series does not typically incorporate a basic residue and secondary fragmentation is initiated by the presence of a 'mobile' proton, leading to depletion of the higher members of the series [70, 71]. This phenomenon is less pronounced when considering CID promoted by resonant excitation such as that performed on a QIT instrument. Once initial fragmentation has occurred the product ions are no longer resonant with the excitation voltage and do not experience multiple ion-neutral collisions [68].

The first proposal for prediction of peptide fragmentation was the mobile proton model [72, 73]. This framework considers proton delocalisation from a stable site of high gas phase basicity, for example a lysine or arginine residue (or the *N*-terminus) to an energetically less favourable position along the peptide backbone. The process is triggered by excitation, which in the case of CID is collision with the neutral background gas. Delocalisation of charge leads to protonation of a backbone amide nitrogen, causing significant weakening of the peptide bond making it more susceptible to fragmentation [74]. The mobile proton model is however simplistic in its approach and cannot accurately predict the observed fragments produced from a peptide. More recently, the more comprehensive Pathways in Competition (PIC) model [75] was proposed. This model incorporates pre-cleavage, amide bond cleavage and post cleavage events in the prediction of peptide fragmentation. The pre-cleavage stage, unlike in the mobile proton model, does not solely incorporate proton transfer processes and internal rotations required to facilitate fragmentation are also considered. Gas phase basicity of the products of fragmentation is of great significance to the post-cleavage phase. Such values are used to predict which species will retain the available proton and consequently be amenable to mass spectrometric analysis as the dominant charged fragment. In spite of the

more extensive approach of the PIC model, a dearth of understanding pertaining to peptide fragmentation/rearrangement processes remains.

1.7.1.1. Peptide Rearrangement during CID

The mobile proton model assumes peptide fragmentation will result in the production of b- and y-ions and these ions are typically used for sequence identification [51, 52]. Studies by Mueller *et al.* [76] and Cordero *et al.* [77] have confirmed the structures of *C*-terminal y-ions to be that of an amino acid (y_1) or a protonated truncated peptide (y_n). However, CID spectra often contain product ions that are not generated following standard fragmentation pathways. Yalcin *et al*. [78] showed b-ion structure to be that of a linear peptide backbone terminating in a 5-member cyclic oxazolone ring at the carboxy-terminus. Neucleophilic attack by the *N*-terminal amine on the electrophilic carbonyl carbon within the oxazolone ring can result in the formation of a fully macrocyclic intermediate from the partially cyclic b-ion. A study by Yague *et al.* [79] demonstrates this process and how consequent ring opening of the macrocyclic intermediate at the amide bonds causes scrambling of the original peptide sequence whereby amino acids that are internal in the starting sequence become exposed at the carboxyl-terminus of the rearranged species (Figure 30). These sequence scrambled species can then be subject to secondary fragmentation via conventional CID pathways, resulting in the observation of product ion series that are not characteristic of the original precursor ion.

Harrison showed that the abundance of product ions generated from these non-native sequences increases with elevated CID collision energies [14]. The work also demonstrates that sequence scrambling can be prevented through acetylation of the peptide *N*-terminus, a finding which also supports the proposed mechanism of macrocycle intermediate formation. Of concern when considering the propensity for such complex fragmentation pathways to occur is that facile sequence rearrangement and subsequent fragmentation could result in the production of intense scrambled product ions leading to mis-assignment of the original peptide sequence. Research into the frequency with which scrambled sequences are observed and potential consequences for large scale MS/MS analyses has been conducted [80-82]. However, there remains a dearth of understanding regarding the extent to which such sequence scrambling occurs or indeed the factors which promote/inhibit the process and these are discussed in more detail in Chapter 6.

1.7.3. Electron Transfer Dissociation

ETD was developed in 2004 by Syka *et al.* [8] as an alternative to the novel electron capture dissociation (ECD) of peptide/protein ions that had been earlier introduced by McLafferty and co workers [7]. Application of ECD is essentially confined to ICR mass spectrometers due to the associated difficulties of storing thermal electrons in the RF field of alternative trapping instruments. ECD has been performed on such instruments although with reduced fragmentation efficiency and mass spectral quality [83-85]. The advent of ETD enabled electron driven dissociation to be routinely employed on widely available, less expensive instrumentation such as the QIT. ETD uses anions with low electron affinity as a vehicle for the delivery of a single electron to a peptide cation of interest. Various anion reagents have been used for this transfer process with fluoranthene and anthracene being those used most widely [86]. Both ECD and ETD are believed to be non-ergodic processes [87, 88] whereby there is no associated redistribution of internal vibrational energy, however work by Tureček suggests this may not be entirely the case [89]. A major benefit of electron driven peptide fragmentation is that preservation of labile post-translational modifications is possible. Phosphoproteomics is a research area of great interest [90-93] and ETD fragmentation enables cleavage of the peptide backbone without loss of the phosphate group from the amino acid side-chain that often results under CID conditions [8]. Retention of additional modifications such as sulfonation, methylation, glycosylation and acetylation has also been shown by ETD [94].

The mechanism by which c/z-ion fragments are generated following electron capture (and electron transfer) dissociation is still debated. However, the Cornell mechanism [7] and Utah-Washington mechanism [95, 96] are favoured. In each case, cleavage of the N-C_α bond results, typically generating c^+ and $z^{+\cdot}$ ions; however $c^{+\cdot}$ and z^+ species can also be produced via a hydrogen atom transfer process [97]. Fragmentation of peptide/protein cations by ETD appears to proceed in a non-specific manner; size of the cation species and amino acid sequence in general do not dictate the site of fragmentation. One exception to this is that cleavage *N*-terminal to the imino acid proline is disfavoured due to the presence of the tertiary nitrogen. However, fragmentation efficiency by ETD is influenced by the charge state of the precursor ion, with highly charged species ($z \geq 3$) being more amenable to analysis by electron driven methods [98, 99]. The likelihood of non-dissociative electron transfer increases with decreasing precursor charge state. Newly formed c/z-ion species resulting from ETD are thought to be held together by non-covalent interactions in a dissociation

complex; this complex has a greater stability when formed from a peptide of lower charge state as repulsion between the separating fragments is reduced. Application of a low energy collisional activation, termed 'supplemental activation' can be used to break this complex, generating diagnostic c/z-ions [100]. The magnitude of the collisional activation required is determined by the charge state of the precursor ion and must be tuned to provide efficient production of ETD fragment ions while minimising unwanted CID of the complex.

1.8. Ion Mobility

Ion mobility spectrometry (IMS) enables separation of analytes on a basis of their gas phase mobility. The technique is electrophoretic in nature and when coupled with MS allows characterisation of species by size and shape in addition to mass and charge. Gas phase rotationally averaged collision cross section (Ω) values can be empirically measured for protein and peptide ions. Such values can then be compared to X-ray and NMR solution structures acquired under native conditions [101, 102]. A variety of systems have been studied by Ion Mobility Mass Spectrometry (IMS-MS) including protein complexes [103], peptide oligomers [104-106], covalently modified peptides [107, 108], chiral enantiomers [109] and stereoisomers [110], highlighting the wide applicability of the technique to many research areas.

1.8.1. Linear Ion Mobility

The principles which govern the mobility of an ion through a drift tube with an applied uniform weak electric field are well established [111, 112]. Each ion has a mobility constant, K ($cm^2\ s^{-1}\ V^{-1}$) that is determined in the presence of an electric field gradient, **E** ($V\ cm^{-1}$) and it is this mobility which determines the velocity, **v** ($cm^2\ s^{-1}$) of an ion:

$$\mathbf{v} = \mathrm{KE} \qquad (5)$$

Mobility constants are typically given in terms of a reduced mobility (K_o) and such values have been corrected for standard temperature and pressure [111]:

$$K_0 = \mathrm{K}\,\frac{273}{\mathrm{T}}\,\frac{\mathrm{P}}{760} \qquad (6)$$

where temperature (T) is in Kelvin and pressure (P) is in Torr.

In a given system the energy of an ion is determined by the electric field strength divided by the number density of the drift gas, E/N. It is this energy that determines the behaviour of an

ion moving through a drift gas. When operating in the low field region (where the energy imparted by the electrical field is far less than the thermal energy of the system) mobility is independent of the applied electric field and mobility is therefore inversely proportional to Ω. Consequently, measurement of the drift time in a particular buffer gas at a given temperature and pressure, combined with knowledge of the applied electric field strength permits direct calculation of Ω through application of the Mason-Schamp equation (Equation 7).

$$\Omega = \frac{3q}{16N}\sqrt{\frac{2\pi}{\mu k_B T}}\frac{1}{K} \quad (7)$$

where q is ion charge

N is the number density of the drift gas

μ is the reduced mass of the ion

k_B is the Boltzmann constant

T is the temperature of the drift gas

Ω is the collisional cross section of the ion

At higher E/N the relationship breaks down: ions can align with the applied field and the mobility of a given species varies with electric field [111].

1.8.2. Travelling Wave Ion Mobility

Travelling Wave Ion Mobility spectrometry (TW-IMS) has emerged as a versatile analytical technique for the study of biological systems [113-115]. The TW-IMS instrument was developed by Waters Corporation and combines the capability to perform ion mobility separations with QTOF mass spectrometry. Mobility separation is performed within a stacked ring ion guide (SRIG); this is a series of ring electrodes, where consecutive electrodes have opposite RF voltage potentials applied (Figure 12) [116].

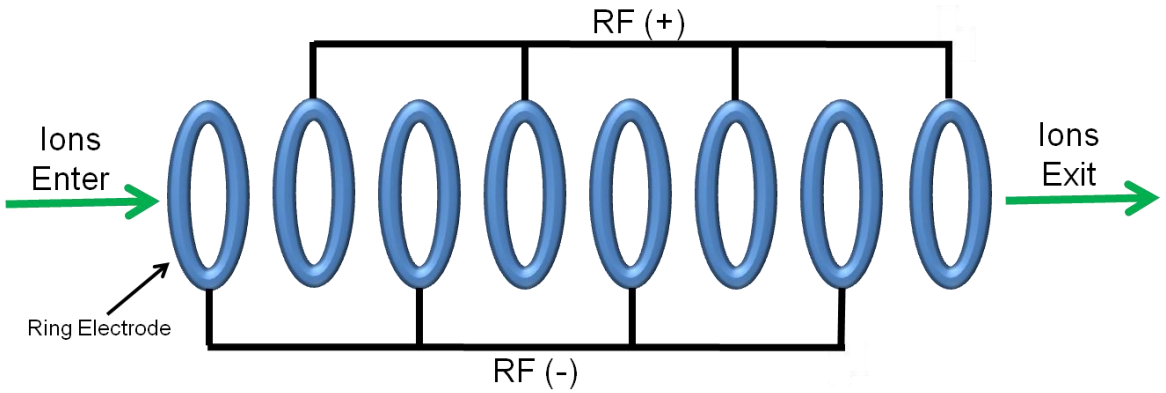

Figure 12. Illustration of a SRIG showing the orientation of electrodes and applied RF voltage. Figure adapted from that given in [117].

Application of the RF voltage generates a radially confining trap which has the ability to slow or even prevent axial movement of an ion through the SRIG when in the presence of a background gas. Propulsion of ions through the guide is achieved by superimposing a transient DC voltage onto adjacent electrode pairs. This generates a 'travelling wave' of potential carrying ions along the z-axis (Figure 13) [117].

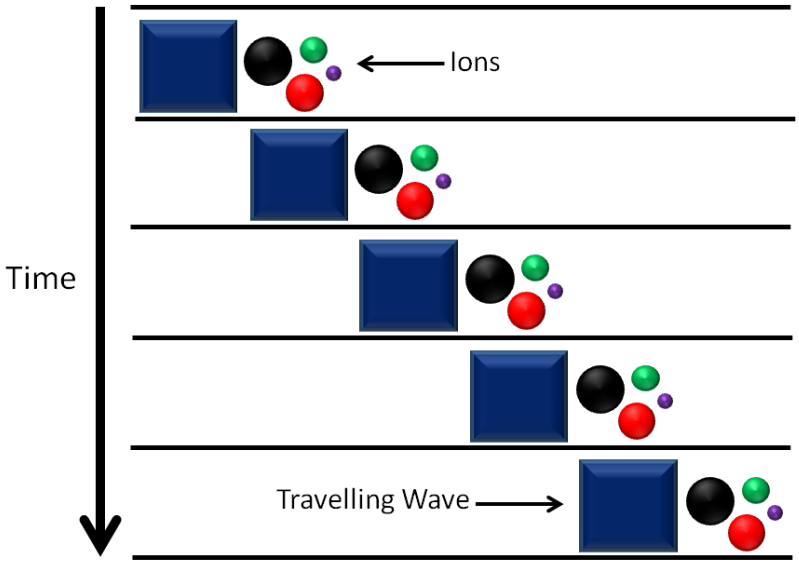

Figure 13. Movement of the 'travelling wave' created by the transient DC voltage within a T-Wave mobility device. Ions are propelled along the z-axis as a function of time. Figure adapted from that given in [117].

The motion of an ion under the influence of the 'travelling wave' is governed by the wave velocity, wave height and gas pressure as well as the mass, charge and crucially the mobility of the species. If the pulse height is set to a value insufficient for permanent ion propulsion, the ion can 'roll over' the wave. In such a case the ion becomes subject to the force of the subsequent wave and its progress along the z-axis is slower than that of an ion which experiences consistent propulsion.

Figure 14 is a schematic of the first generation Synapt HDMS Travelling Wave QTOF instrument, which has been used for a number of the experiments described in this thesis. The trap and transfer compartments operate at pressures similar to that of a standard collision cell; consequently, CID can be performed in these regions. Three operation modes are possible; i) fragmentation of mobility separated precursor ions, ii) mobility separation of product ions and iii) fragmentation of mobility separated product ions. A quadrupole mass analyser is positioned before the trap region, meaning that analysis can be performed on the entire range of species present or on ions at a pre-selected *m/z* of interest. This range of operation modes makes the Synapt instrument ideal for the study of conformer rearrangement following collisional activation on the millisecond timescale [118, 119].

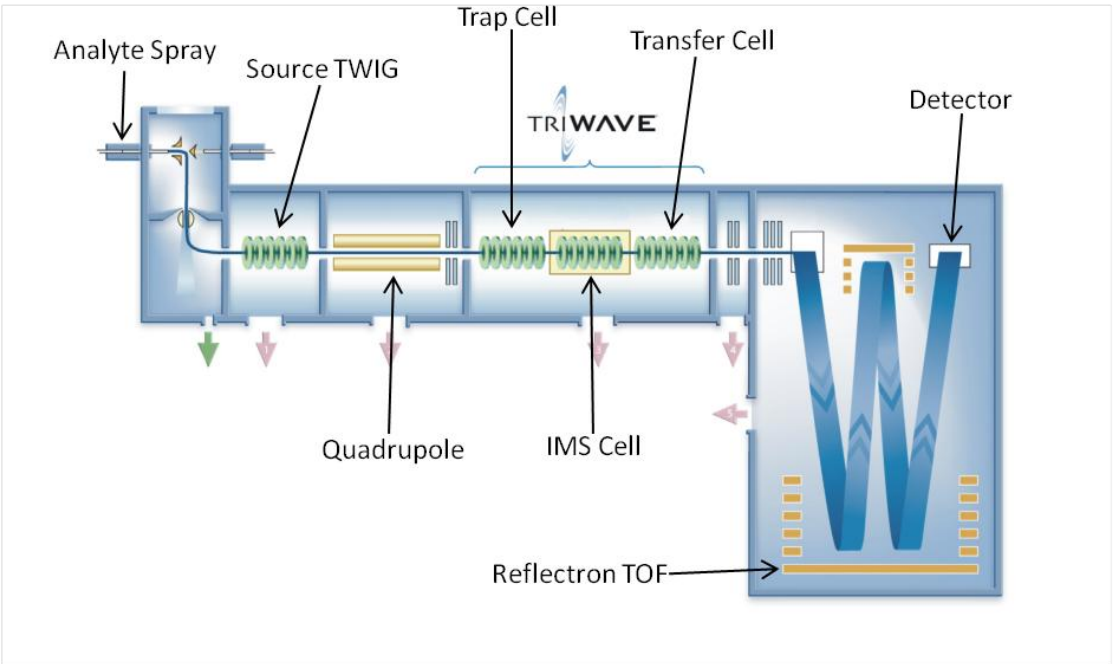

Figure 14. Schematic of the Synapt HDMS TW-IMS Mass Spectrometer produced by Waters Corporation.

Direct measurement of the rotationally averaged collision cross section of gas phase ions is routinely performed by Linear ion mobility; however, such 'first principle' measurements cannot be performed on TW-IMS instruments. The time-varying, non-uniform electric field in the SRIG means that there is no longer a linear relationship between applied electric field and mobility and consequently 'first principle' calculations cannot be performed. Calibration of the Travelling Wave instrument must be performed if empirical determination of collision cross section is required; analytes of known cross section are analysed under identical experimental conditions as those used to study the sample of interest. Several approaches toward such calibration have been proposed [101, 102, 107, 120, 121] and the majority of such approaches utilise the multiplicity of empirical peptide/protein cross sections published by Clemmer and co-workers [122, 123]. Calibration of TW-IMS instruments is discussed in detail in Chapter 3 of this thesis.

The combination of 'travelling wave' mobility separation and TOF MS not only provides additional dimensionality over standard MS analyses but can be used to improve data quality. Mobility separation of complex peptide mixtures prior to MS detection enables multiply charged peptidic species to be discriminated from singly charged background and other interferences such as salts or detergents, thus enabling detection of species present at low levels [124]. Crucially, through ion accumulation and radial confinement, transmission of ions through the TW-IMS instrument is high, meaning that sensitivity is not compromised due to the incorporation of the mobility cell. Consequently, analytically relevant sample quantities are readily studied by TW-IMS making the technique ideal for characterisation of biological systems [125].

1.8.3. Field Asymmetric Waveform Mass Spectrometry

Field Asymmetric Waveform Mass Spectrometry (FAIMS) separates ions on a basis of their differing mobility in a high electric field relative to a low electric field [126]. As with Linear ion mobility, drift mobility (K_o) is proportional to field strength when in the low field regime. However at stronger field strengths this relationship breaks down. The mobility takes on a non constant high field character (K_h) and it is this dependence on the applied electric field that enables FAIMS separations. The value of K_h is analyte specific; some species exhibit an increase in mobility under the influence of a high electric field whereas others have a decreased mobility. A typical FAIMS instrument operates by carrying an ion in a gas stream

between two electrode plates, one held at ground potential and the other with an applied asymmetric waveform. This waveform is modified in the manner shown in Figure 15. A high voltage is applied for a short period of time followed by a voltage of opposite polarity, at lower amplitude, for a longer time period.

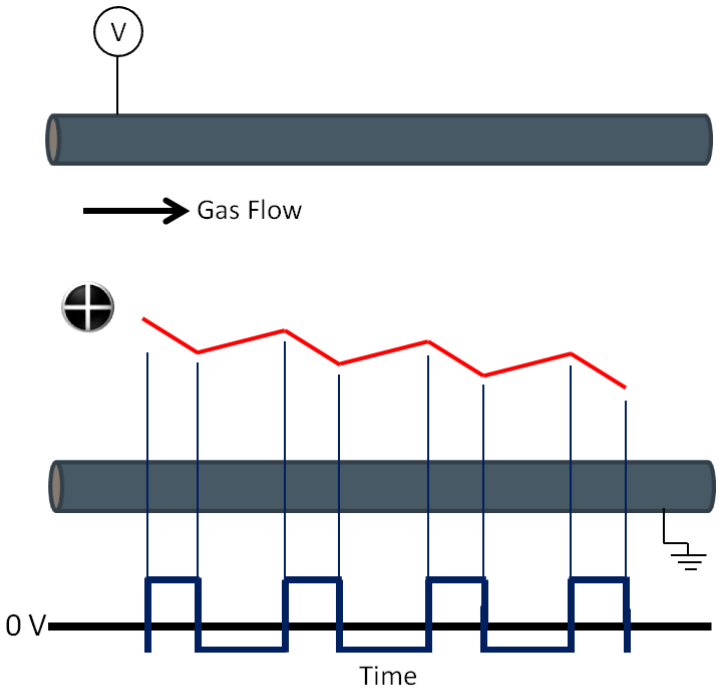

Figure 15. Diagram illustrating the application of an asymmetric waveform to a flat electrode FAIMS instrument, the subsequent ion motion along the z-axis is also shown. The waveform is applied to the upper plate while the lower plate is held at ground potential.

The total voltage-time product across the entire waveform is zero ($V_1t_1 + V_2t_2 = 0$). In the high field portion of the waveform the ion will travel at a velocity of $\mathbf{v}= K_hE_{high}$ and conversely it will travel at a velocity of $\mathbf{v}=K_oE_{low}$ in the low field portion of the cycle. This corresponds to distances travelled of $d= K_hE_{high}t_{high}$ and $d=K_oE_{low}t_{low}$. As the voltage-time product is zero, if the high and low field mobility each had the same value an ion would return to its original radial position upon completion of each waveform. Application of the high electric field portion of the waveform means that the ion undergoes radial displacement. Separation of species is therefore possible with each analyte travelling a unique distance in the y-direction and hence taking varying times to traverse the drift region [127].

Chapter 2

Experimental

The following experimental procedures are those which were used to generate the data presented in this thesis.

2.1. Enzymatic Digestion of 'Standard'/QCAL-IM Proteins

The protein to be digested was diluted to 10 pmol/µl in 50 mM ammonium bicarbonate, reduced with dithiothreitol (DTT) (4 mM, 60 °C, 45 min) and alkylated (iodoacetamide, 14 mM, room temperature, dark, 1 hr). The DTT concentration was then increased to 7 mM and the protein digested overnight with trypsin, Asp-N, Lys-C or Lys-N proteases (2 % [w/w], 37 °C).

2.2. Secondary Digestion with carboxypeptidase B

A 30 µl aliquot was removed from the initial tryptic or Lys-C digestions and then incubated with CBPB (Worthington Biochemical, New Jersey, USA) at an enzyme ratio of 12 units per microgram of protein. Secondary digestion was allowed to proceed for 2 hrs at room temperature prior to being quenched through addition of 10 µl H_2O containing 0.1 % formic acid (FA).

2.3. Desalting of Peptide Samples

Resultant peptide mixtures from protein digestion were desalted using C_{18} ZipTips (Millipore, Watford, UK) prior to MS analysis to help maximise sensitivity. The tips were initially hydrated with 5 cycles of 10 µl H_2O (1 cycle constitutes aspiration and dispensing) prior to being equilibrated with a further 5 cycles of 10 µl (H_2O/MeCN/FA 50/50/0.1). Sample (approximately 100 pmol total) was then loaded onto the tip through application of 15 cycles of 10 µl and salts eluted into 20 µl H_2O through application of a further 15 cycles. Sample was then eluted from the tip into an Eppendorf vial containing 100 µl (H_2O/MeCN/FA 50/50/0.1) through application of 15 repeated cycles ready for analysis by ESI.

2.4. Expression of QconCAT Protein

The expression plasmid encoding the QconCAT protein was received from Entelechon GmbH in a pET21a vector; this was dissolved in 50 µl of Tris-EDTA buffer and introduced into competent BL21(DE3)pLysS cells by heat shock at 42 °C. The transformed *E. coli* cells were then spread on an lysogeny broth (LB) agar plate supplemented with 0.1 mg/ml ampicillin and allowed to grow overnight at 37 °C.

A single colony was then picked using a sterile pipette tip and transferred into 15 ml of LB containing 0.1 mg/ml ampicillin before incubation at 37 °C for 16 hrs (overnight) with shaking. The saturated overnight culture was then diluted in LB (1 in 100) and grown as above for approximately 3 hrs until Abs_{600} = 0.6-0.8. At this point protein expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and allowed to proceed for 3 hrs at 37 °C.

The bacterial culture was then centrifuged at 8,000 × g for 10 min (4 °C), the supernatant discarded and the pellet stored at -80 °C until required.

2.5. QconCAT Extraction

The cell pellet was resuspended in 4.5 ml of ice cold phosphate buffered saline (PBS) containing 6 M guanidine HCl to which 50 U/ml benzonase was added and the mixture shaken for 20 min at room temperature. The cells were then centrifuged at 18000 × g, 20 min, 4 °C and the supernatant (soluble fraction) removed.

The pellet was suspended in 4.5 ml PBS/6 M guanidine HCl and 80 µl of a 10 mg/ml lysozyme solution added, prior to being vortexed at room temperature, 5 min. A further 27 ml of PBS/6 M guanidine HCl was added and then centrifuged as above; the supernatant was again removed and stored on ice.

Finally the pellet was washed with 36 ml of PBS/6 M Guanidine HCl prior to centrifugation; this process was repeated 3 times and the supernatant retained on each occasion.

2.6. QconCAT Purification

The post extraction cell pellet was resuspended in 8 ml PBS/6 M guanidine HCl and incubated for 40 min, 4 °C with 1 ml TALON Affinity Beads (Clontech, California, USA) pre-equilibrated in the same solution. The pellet and beads where then loaded into a gravity elution column (Econopac, BioRad, Germany) and washed with 6 ml PBS buffer/6 M guanidine HCl to remove major impurities. Finally, sample elution was performed using the same buffer with addition of 150 mM imidazole and the eluent collected as 10 × 0.5 ml fractions.

2.7. Estimation of Protein Concentration (Bradford Assay [128])

A 1 mg/ml standard solution of BSA (Sigma Aldrich) was diluted as required using 50 mM ammonium bicarbonate to create a concentration calibration range between 2,000 µg/ml and 25 µg/ml. Following this, 1.5 ml of Bradford reagent was added to 30 µl of each solution and the mixture vortexed at room temperature. All samples were left for a consistent time period of 5 mins and then the absorbance of each measured in a plastic cuvette at 595 nm against a blank of 1.5 ml Bradford reagent and 30 µl of dilution buffer. The calibration curve produced was then used to estimate the protein concentration of the QconCAT sample prepared in the same manner.

2.8. Guanidination of Lysine residues

Peptide digestion mixtures (100 pmol total material) were dried by speed-vac and reconstituted in 10 µl H_2O. To this, ammonium hydroxide (7 M, 10 µl) and *O*-methylisourea (0.5 M in H_2O, 5 µl) were added and the reaction mixtures left overnight at room temperature. Following drying-down of derivatised samples by speed vacuum centrifugation each was reconstituted in 10 µl H_2O and desalted as previously described in Section 2.3.

2.9. Linear Field Mobility Analysis using a Modified Synapt HDMS Instrument

Tryptic and Lys-C digestions of the QCAL-IM standard were prepared at a concentration of 2 pmol/µl in H_2O/MeCN/FA 50/50/0.3 and infused using nESI tips prepared in-house into a modified Synapt HDMS instrument which incorporates a Linear drift cell (Waters, Manchester, UK) and is based at the University of Oxford. Analyses were conducted at a helium drift gas pressure of ~2.1 Torr and a temperature of ~300 K at a range of drift voltages

between 50 V and 200 V across the 18 cm long drift tube. The TOF pusher was operated at a 70 µs interval and data collected between 100 and 2700 *m/z* for each analysis.

2.10. Linear Field Mobility Analysis using a Modified Commercial QTOF Instrument

Further Linear mobility analysis was undertaken on a modified commercial QTOF instrument which incorporates a drift cell and ancillary optics post-source that enables measurement of Linear field ion mobility in addition to *m/z* ratio data. In this instance the proteolysed QCAL-IM samples were prepared at a concentration of ~35 pmol/µl in H_2O/MeCN/FA 50/50/0.3 and infused using glass capillary nESI tips prepared in-house. The helium drift gas pressure was ~2.7 Torr and temperature ~300 K and analysis conducted at a range of drift voltages between 15 V and 60 V across the 5.1 cm drift cell.

2.11. T-Wave Mobility Analysis using a Synapt HDMS Instrument

QCAL-IM samples proteolysed with trypsin or Lys-C were diluted to 1 pmol/µl using 50 % (v/v) acetonitrile in H_2O, containing 0.3 % (v/v) FA. The samples were then infused into a Synapt HDMS instrument (Waters, Manchester, UK) using gold coated nanospray emitter tips (Proxeon, Odense, Denmark). The capillary voltage, cone voltage and source temperature were typically set at 1.9 kV, 40 V and 80 °C respectively. The travelling wave speed was set to 300 m/s and wave height ramped from 4.2 V to 16.0 V over the full drift time. The mass spectrometer was set to collect data for species between 100 and 4000 *m/z* with the TOF pusher operating at a 64 µs interval. The nitrogen gas pressure in the T-Wave cell was approximately 4.4×10^{-1} Torr. Further analyses to elucidate the collision cross section of both Lys-C and Lys-N derived peptides following T-Wave calibration with the QCAL-IM standard and that to perform separation of fragment ions produced by CID in the Trap collision cell were also conducted with the instrument conditions described here.

Analysis of tryptic and CBPB truncated peptide digestion mixtures produced from digests of 'standard' proteins (1 pmol/µl in H_2O/MeCN/FA 50/50/0.1) were also analysed on the Synapt HDMS instrument using gold coated nanospray emitter tips. No mass selection was performed and the entire array of peptide species transmitted to the T-Wave cell for mobility separation. The nitrogen gas pressure within the T-Wave cell was set to approximately 4.4×10^{-1} Torr and data collected between 50-4000 *m/z*. The sequences were analysed with the travelling wave speed set to 300 m/s and at a constant wave height of 9.7 V.

2.12. T-Wave Mobility Analysis Following ETD using a Synapt G2 Instrument

Tryptic digestion of BSA was performed and the resultant peptide mixture, in addition to the 'standard' peptide Substance P (Sigma Aldrich, Dorset, UK) analysed by direct infusion ESI using a Synapt G2 instrument (Waters, Manchester, UK). Samples were prepared at a concentration of 5 pmol/µl using H_2O/MeCN/FA/m-nitrobenzyl alcohol (1/1/0.1/0.1 v/v/v/v). m-nitrobenzyl alcohol was included as a supercharging reagent to aid the formation of higher charge states. The *m/z* of interest was mass selected and ETD induced through co-trapping alongside nitrobenzene ETD reagent in the Trap collision cell. Mobility separation in the T-Wave cell was then performed with the travelling wave speed set at 300 m/s, a constant wave height of 40 V and at a nitrogen pressure of ~2 Torr.

2.13. QIT Analysis

Digested material (500 fmol) was loaded onto an Acclaim PepMap 100 trapping column (75 µm × 2 cm; 5 µm particle size; 100 Å pore size) for desalting in-line prior to separation by RPLC with an Acclaim PepMap C18 column (75 µm × 15 cm; 2 µm particle size; 100 Å pore size) both purchased from LC Packings Dionex (Surrey, UK). The column temperature was maintained at 35 °C and the flow rate for the RPLC separation set to 300 nL/min using an EASY nLC system (Proxeon, Odense, Denmark) coupled to an amaZon ETD ion trap mass spectrometer (Bruker, Bremen, Germany). The column was equilibrated with 0.2 % (v/v) FA (Solvent A) and developed with 90 % (v/v) acetonitrile containing 0.2 % (v/v) FA (Solvent B); 0 to 55 % over 55 minutes, 55 to 90 % over 35 minutes and 90 to 100 % over 5 minutes.

The 3 most intense precursor ions observed in each survey scan were selected for MS/MS analysis with the CID Fragmentation Amplitude value set at 1.2 V and ramped from 30 % to 300 % of this value. ETD MS/MS analysis was performed with a reaction time of 150 ms and the 'Smart Decomp' option set to auto, thereby tuning the amplitude of supplemental activation according to charge state.

When analysis was performed by direct infusion ESI a syringe driver (5 µl/min) was utilised for delivery of peptide digestion mixture (1 pm/µl in H_2O/MeCN/FA 50/50/0.1) to the nESI source. Mass selection at 410.7 *m/z* corresponding to that of the peptide of interest (FGERALK) was performed followed by CID with the Fragmentation Amplitude set to 0.35 V with data collected between 150 and 1800 *m/z.*

2.14. QTOF Analysis

RPLC on a nanoAcQuity UPLC (Waters, Manchester, UK) was used to analyse 100 fmol of digested material. The sample was desalted in-line using a Symmetry C18 trapping column (180 µm × 20 mm; 5 µm particle size; 100 Å pore size) and RPLC separation performed with a BEH130 C18 column (75 µm × 100 mm; 1.7 µm particle size; 130 Å pore size). Both columns are manufactured by Waters, Manchester, UK. The flow rate was set to 300 nL/min and the column temperature maintained at 35 °C. Solvent A was H_2O/FA (100/0.1 %) (v/v) and Solvent B 100 % acetonitrile; solvent B was developed from 1 to 40 % over a period of 30 minutes, 40 to 85 % over 5 minutes and then the column re-equilibrated with 1 % Solvent B for 25 minutes. The chromatography system was interfaced to a Synapt G2 HDMS mass spectrometer (Waters, Manchester, UK) set to transmit species between 50 and 2000 *m/z*. The 3 most abundant ions observed in a given scan were subjected to CID with the collision offset potential ramped between 5 and 40 V.

For analyses utilising direct infusion ESI both derivatised/underivatised samples were prepared at a concentration of 1 pmol/µl in H_2O/MeCN/FA 50/50/0.1 and were infused using gold coated ESI nanospray emitter tips (Proxeon, Odense, Denmark) on a Synapt HDMS instrument (Waters, Manchester, UK). The *m/z* of interest was mass selected by the quadrupole analyser prior to being subjected to CID in the Trap collision cell. The Collision energy applied in this region was manually tuned to give optimal peptide fragmentation.

Chapter 3

QCAL-IM: A QconCAT Standard for Calibration of Ion Mobility Systems

3.1. Introduction

Travelling Wave (T-Wave) Ion Mobility combined with mass spectrometry enables the gas phase conformation of multiply charged protein ions to be studied. T-Wave ion mobility is frequently employed to ascertain structural information regarding proteins and protein complexes [103, 105, 129], however, there is growing interest in the gas phase conformation of peptide ions and products derived from peptide fragmentation [108, 118, 119, 130]. The added dimension of information regarding the 'shape' of an analyte is a useful tool in furthering our understanding of the gas phase ion chemistry of these species; however, unlike in Linear ion mobility direct measurement of rotationally averaged collision cross section (Ω) is not possible. The time varying, non-uniform electric field of the travelling wave means that there is no longer a linear relationship between drift time and ion mobility, and hence Ω [131]. The complexity of the relationship between mobility in the T-Wave device and Ω therefore precludes simple calculation of absolute Ω values. Drift time in the T-Wave instrument must therefore be calibrated using species of known collision cross section, under identical experimental conditions to those used for analysis. Current approaches toward calibration of the T-Wave instrument [101, 102, 107, 120, 121] in general rely on the peptide/protein ion cross sections included in the database published by Clemmer and co-workers [122, 123].

Use of a single unverified source for instrument calibration is never ideal and the Clemmer database in particular is not optimal for instrument calibration when study of relatively low molecular weight species is required. Proteolysis of multiple proteins is required to cover the Ω range required; for example, to calibrate for doubly charged peptide ions with Ω ranging from ~173-438 $Å^2$ a minimum of two proteins (horse cytochrome c and sheep haemoglobin) must be digested. In addition, the database only incorporates species that permit calibration for singly and doubly protonated ions. The objective of the work discussed in this chapter was to develop a single, well characterised standard for calibration of the T-Wave instrument to enable empirical determination of the gas phase collision cross section of peptide and peptide fragment ions.

3.2. Results and Discussion

3.2.1. Selection of Peptides for Inclusion in QCAL-IM

To identify candidate peptides for inclusion within the ion mobility calibration standard (QCAL-IM) a range of 'standard' proteins (enolase from *S. cerevisiae*, horse myoglobin, human serum albumin, chicken lysozyme, QCAL [132]) were digested using the endoprotease trypsin. Resultant peptide mixtures were sprayed by direct infusion nESI into a Synapt HDMS instrument and for each observed peptide the drift time, charge state and signal intensity was extracted to generate a library of candidate species for inclusion in the standard (Figure 16).

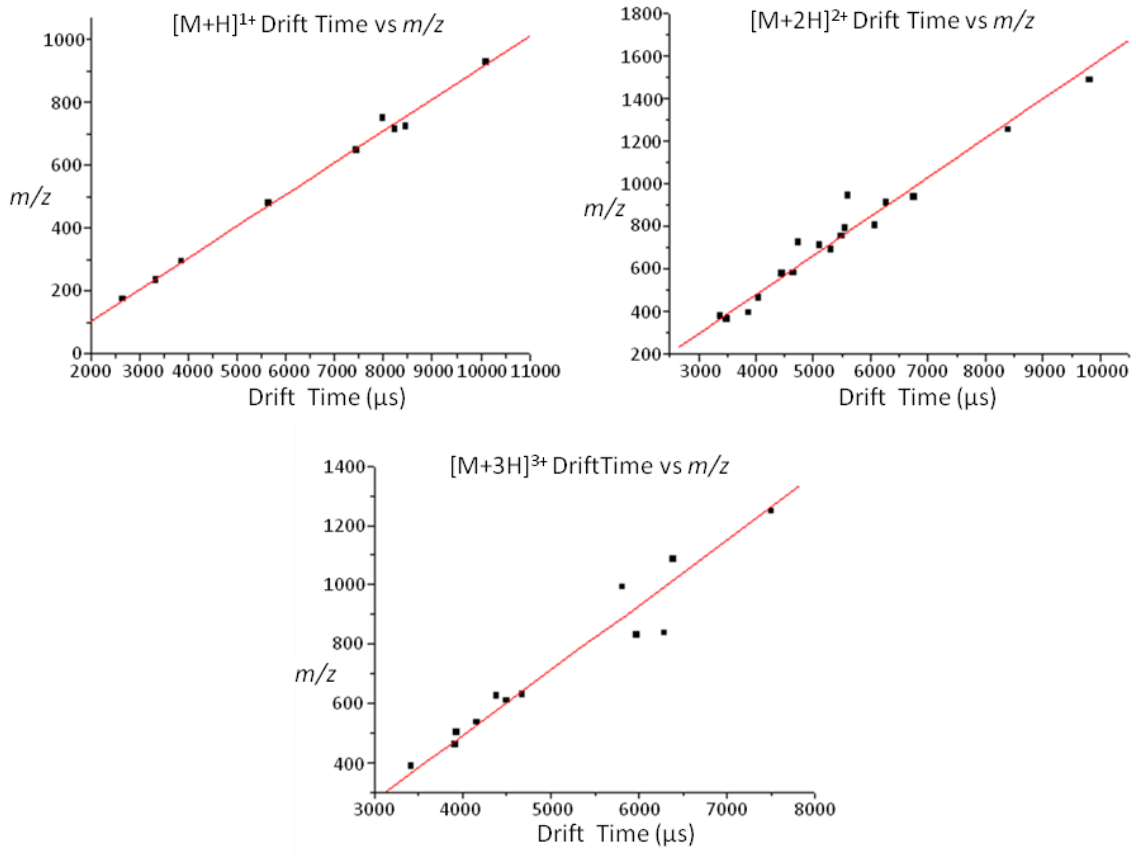

Figure 16. Plot of drift time vs *m/z* for peptides observed from analysis of trypsin digestion mixtures from a range of 'standard' proteins. From this library, candidate peptides were selected for inclusion in QCAL-IM.

Peptides ultimately chosen as calibrant standards from this library were selected to provide the widest possible range of drift times (and hence Ω) and also according to their charge-dependent correlation between *m/z* and drift time (Figure 16). Peptides were also only selected if they were observed with abundant relative signal intensity and therefore likely to provide good gas phase ion yields [133] following proteolysis of the QCAL-IM standard. Those peptides selected cover a mass range of 470-3740 Da and drift time range of 2750-9470 µs (Figure 17) when using the conditions described for T-Wave analysis.

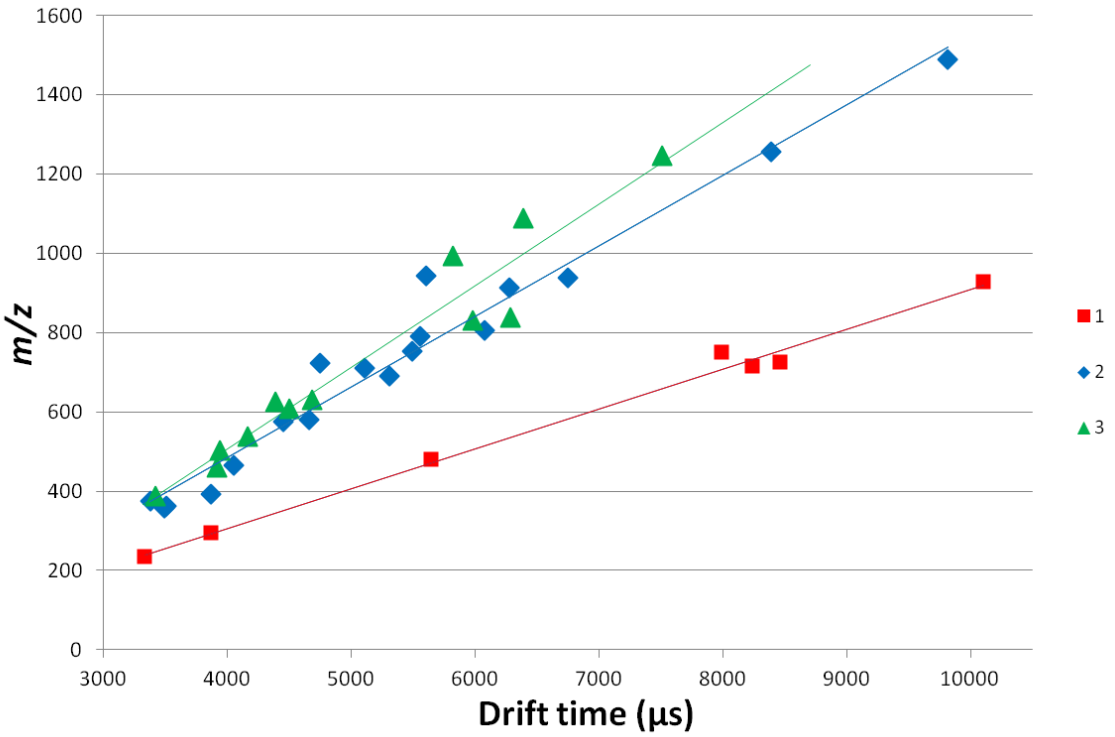

Figure 17. Drift times for the singly, doubly and triply protonated peptides selected for inclusion in QCAL-IM, following trypsin proteolysis of a set of 'standard' proteins.

3.2.2. QCAL-IM Construction and Expression

A consideration when designing the artificial protein construct was to concatenate the tryptic sequences in such a way that the potential range of drift time calibration could be extended through application of alternative proteases such as Lys-C (cleaving *C*-terminal to lysine residues) or Asp-N (cleaving *N*-terminal to aspartic acid residues) to the synthesised artificial protein. The selected peptides were therefore ordered in a manner to allow for this possibility; however outside of this consideration the ordering of peptide sequences was unimportant. Chosen peptide sequences were concatenated *in silico* alongside the necessary methionine (MGALR) translating initiating residue and His_6 tag (ALVALHHHHHH) purification sequences. DNA that encodes for the signature peptides was then synthesised, concatenated, expressed and purified as detailed in the experimental section.

To explore the possibility of extending the achievable calibration range through application of alternative enzymes QCAL-IM was subjected to proteolysis with these enzymes and resultant peptide mixtures analysed using the Synapt HDMS instrument. While trypsin and Lys-C generated many peptides suitable for further analysis (see below), digestion of the standard using Asp-N was however found to be inefficient with several miscleavage sites observed. A total of 7 peptides derived from Asp-N digestion were detected, none of which extended the calibration range beyond that achievable from tryptic digestion of QCAL-IM. Use of Asp-N as an enzyme was therefore not pursued further. Equivalent analysis of the Lys-C digestion mixture proved more successful, with the peptide GVNDNEEGFFSARNLCNIPCSALLSSDITASVNCAK extending the drift time calibration range for triply charged species to 5497.6 µs. Consequently, both tryptic and Lys-C digestions (Figure 18) of the standard were prepared for Linear field ion mobility analysis.

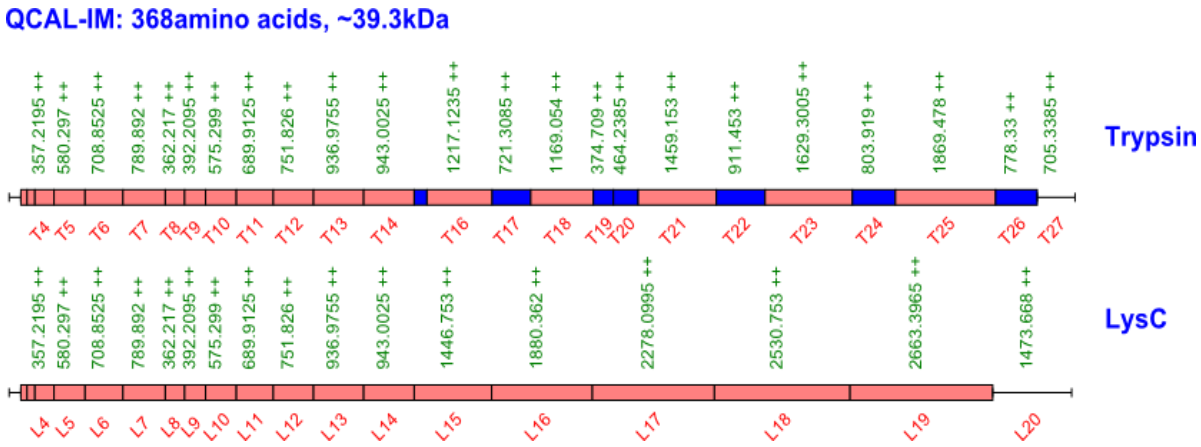

Figure 18. Peptide map of doubly protonated ions resulting from proteolytic cleavage of QCAL-IM using either trypsin or Lys-C; shown are peptide number and *m/z* value of each species.

3.2.3. Linear Field Mobility Analysis

To establish QCAL-IM as an ion mobility calibration standard it was necessary to empirically determine collision cross section values for each of the incorporated peptides. In order to measure such cross section values, proteolysed QCAL-IM was analysed on two separate Linear field ion mobility platforms. Firstly, analysis was performed using a modified Synapt HDMS instrument in the laboratory of Prof. Carol Robinson at the University of Oxford, which has been described in detail elsewhere [129]. In short, the T-Wave mobility cell that normally operates within the instrument has been replaced with a Linear drift tube utilising RF ion confinement thus enabling direct calculation of ion collision cross sections. Secondly, analysis was undertaken using a modified commercial QTOF instrument in the group of Dr. Perdita Barran at the University of Edinburgh [134] which incorporates a post-source drift tube with ancillary optics that also facilitates 'first principle' cross section calculations. A difficulty during analysis of the peptide mixtures on the Linear mobility platforms was that fragmentation of multiply charged species was observed at the higher drift gas pressures typically employed. The onset of fragmentation by CID for multiply charged peptides occurs at lower activation energies than singly charged species and the drift gas pressures employed in the Linear field instruments (> 2 Torr) are far greater than that in the T-Wave instrument (~ 4.4×10^{-1} Torr). Consequently, depletion of the ion current of multiply charged peptides was observed at these higher drift pressures, particularly during analyses at higher drift

voltages. To combat this problem the drift pressures were set as low as possible (approximately 3.2 Torr) while maintaining peak resolution and in each instance, replicate analyses (multiple infusions on different days) performed at a range of drift voltages (50-200 V) across the drift tube. The mobility (and subsequently the Ω) of each observed peptide was then calculated from the straight line gradient obtained from a plot of ion drift time versus reciprocal drift voltage. It was decided that only the Ω values obtained from analysis on the modified Synapt instrument should be ascribed to the QCAL-IM standard. The geometry of this instrument mimics that of the T-Wave Synapt and is therefore optimal for assigning Ω values to the calibrant peptides in QCAL-IM, as potential differences in ion energy and experimental timescales that can contribute toward the calculated cross section are minimised. Such factors can make the use of calculated collision cross sections from drift time data acquired on disparate instrumentation problematic for T-Wave calibration [129]. As such, the calculated absolute cross section values (and associated standard deviations) of the QCAL-IM calibrant peptides on this instrument are detailed in Table 1. A second determination of the rotationally averaged collision cross section of the peptide ions using the modified QTOF instrument provides significant confidence in these values attributed to the QCAL-IM standard.

Table 1. Sequence and *m/z* values of QCAL-IM tryptic/Lys-C peptides; average Ω values obtained from Linear field mobility analysis using the modified Synapt instrument are also shown. The values in parentheses following Ω indicate the standard deviation (or range when only two measurements are available) attributed to each measurement (n = 2-7). When parentheses are absent, n=1. Peptides with no assigned Ω were not observed during analysis.

Peptide	Sequence	$[M+H]^+$	Ω (Å^2)	$[M+2H]^{2+}$	Ω (Å^2)	$[M+3H]^{3+}$	Ω (Å^2)
T1	AATLLPK	713.90	-	357.45	197.0 (1.2)	238.63	-
T2/ L5	IGSEVYHNLK	1160.32	-	580.66	261.8 (2.2)	387.44	300.8 (3.3)
T3	GNPTVEVELTTEK	1417.56	-	709.28	289.0 (1.4)	473.19	-
T4	AVDDFLISLDGTANK	1579.75	-	790.38	312.6 (1.8)	527.25	-
T5/ L8	GVLHAVK	723.89	197.7	362.45	204.9 (1.1)	241.96	-
T6	HLADLSK	783.90	200.3	392.45	212.9 (0.4)	261.96	-
T7/ L10	LVNEVTEFAK	1150.32	261.7	575.66	256.8 (2.1)	384.11	-
T8/ L11	HGTVVLTALGGILK	1379.69	-	690.34	297.6 (2.1)	460.56	353.4 (5.8)
T9	HPGDFGADAQGAMTK	1503.64	-	752.32	-	501.88	-
T10	WLTGPQLADLYHSLMK	1874.23	-	937.62	-	625.41	-
T11	YLEFISDAIIHVLHSK	1886.21	-	943.61	-	629.40	407.1
T12	GVFR	478.57	-	239.79	-	160.19	-
T13	ALVLIAFAQYLQQCPFEDHVK	2491.77	-	1246.39	-	831.26	-
T14	GVNDNEEGFFSAR	1442.49	-	721.75	284.2 (0.6)	481.50	-
T15	NLCNIPCSALLSSDITASVNCAK	2509.40	-	1255.20	435.9 (0.1)	837.13	-
T16	ALELFR	748.91	202.3 (0.1)	374.96	200.2 (1.0)	250.30	-
T17	YLYEIAR	928.08	229.7	464.54	230.6 (1.0)	310.03	-
T18	SHCIAEVENDEMPADLPSLAADFVESK	2976.12	-	1488.56	-	992.71	496.1 (0.4)
T19	SGETEDTFIADLVVGLR	1823.02	-	912.01	349.1 (0.6)	608.34	-
T20	YGASAGNVGDEGGVAPNIQTAEEALDLIVDAIK	3259.57	-	1630.29	-	1087.19	535.7 (1.1)
T21	VEADIAGHGQEVLIR	1607.82	-	804.41	-	536.61	-
T22	TSPYVLPVPFLNVLNGGSHAGGALALQEFMIAPTGAK	3740.39	-	1870.70	-	1247.46	599.0 (0.6)
T23	GGGVNDNEEGFFSAR	1556.60	-	778.80	-	519.53	-
L17	GVNDNEEGFFSARNLCNIPCSALLSSDITASVNCAK	3932.87	-	1966.94	-	1311.62	616.7 (2.3)

Peptides derived from tryptic and Lys-C proteolysis of QCAL-IM cover a cross section range of 197.0-616.7 $Å^2$ as shown in Table 1 and Figure 19. The cross section calibration range for the individual charge states is thus 197.7-261.7 $Å^2$ (singly charged), 197.0-435.9 $Å^2$ (doubly charged) and 300.8-616.7 $Å^2$ (triply charged). An interesting observation from the data is that sequences observed as both singly and doubly charged species exhibit minimal difference in their measured cross section with increasing charge state. However, peptides observed as both doubly and triply charged species exhibit a significant cross section increase in their higher charge state. This phenomenon can presumably be attributed to more significant charge repulsion in the triply charged species as opposed to the doubly charged peptides, resulting in expansion of the gas phase conformation and hence a larger measured collision cross section.

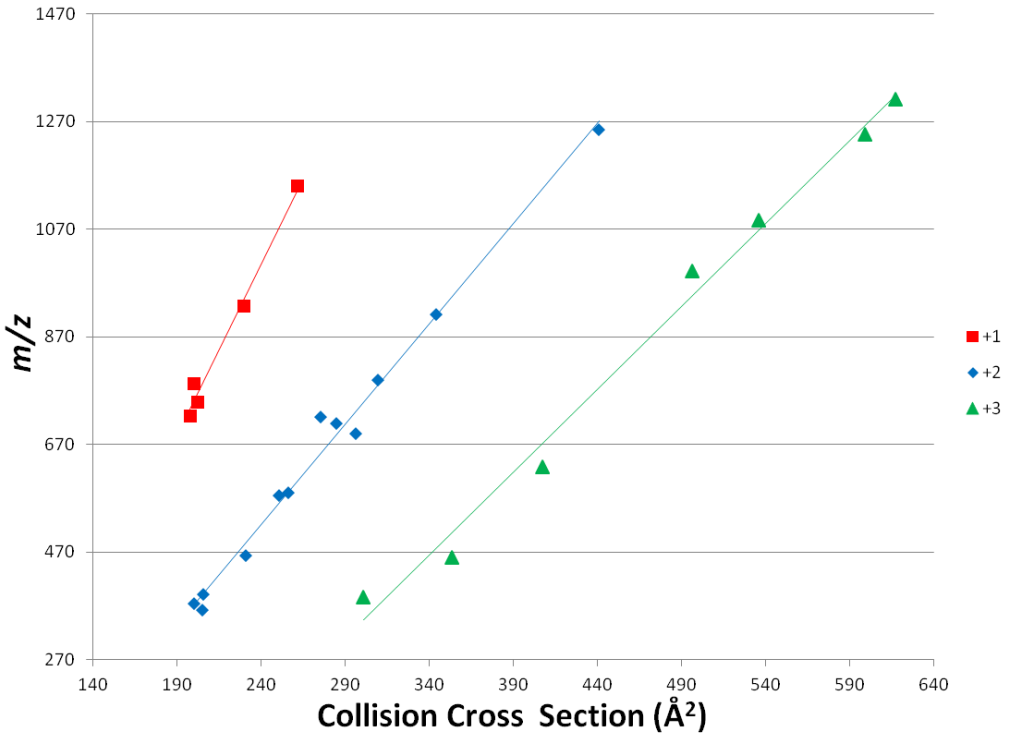

Figure 19. QCAL-IM peptide ion Ω values versus *m/z*; cross section values were obtained on the modified Synapt instrument with a Linear field mobility drift tube in the presence of helium drift gas.

Shown in Table 2 is a comparison of cross section values determined on the modified Synapt and modified QTOF Linear field instruments for a range of doubly protonated species. There is excellent correlation (R^2=0.989) between the two data sets as shown by the linear regression comparing the calculated cross section as a function of the measured arrival time distributions from each platform in Figure 20. Importantly, there is also good agreement, with a mean 4.6 % variation, in the cross section values (maximally 9.2 % for $[M+2H]^{2+}$ of GVNDNEEGFFSAR at 721.74 *m/z*). This variation is within the experimental error as determined by the standard deviations quoted in Table 2, suggesting that potential effects arising from differing instrument geometries (and hence ion energy/experimental timescale) are not of significant consequence toward the measured Linear collision cross section.

Table 2. Comparison of calculated Ω for $[M+2H]^{2+}$ peptide ions determined on the modified Synapt and QTOF Linear field instruments. Values in parentheses indicate the standard deviation (or range when only two measurements are available) associated with each measurement (n=2-4). When parentheses are absent, n=1.

Peptide	**Modified Synapt Ω (Å^2)**	***m/z***	**Modified QTOF Ω (Å^2)**	**Δ Ω (Å^2)**	**% Ω Difference**
T2/ L5	261.8 (2.2)	580.66	248.9 (14.0)	12.9	4.9
T3	289.0 (1.4)	709.28	279.2 (10.6)	9.8	3.4
T4	312.6 (1.8)	790.38	305.3 (15.3)	7.3	2.3
T6	212.9 (0.4)	392.45	198.2 (4.8)	14.6	6.9
T7/ L10	256.8 (2.1)	575.66	242.5 (10.2)	14.3	5.6
T8/ L11	297.6 (2.0)	690.34	292.7 (10.1)	5.0	1.7
T14	284.2 (0.6)	721.75	258.1	26.1	9.2
T15	435.9 (0.1)	1255.20	450.7	-14.8	-3.4
T19	349.1 (0.6)	912.01	334.6	14.4	4.1

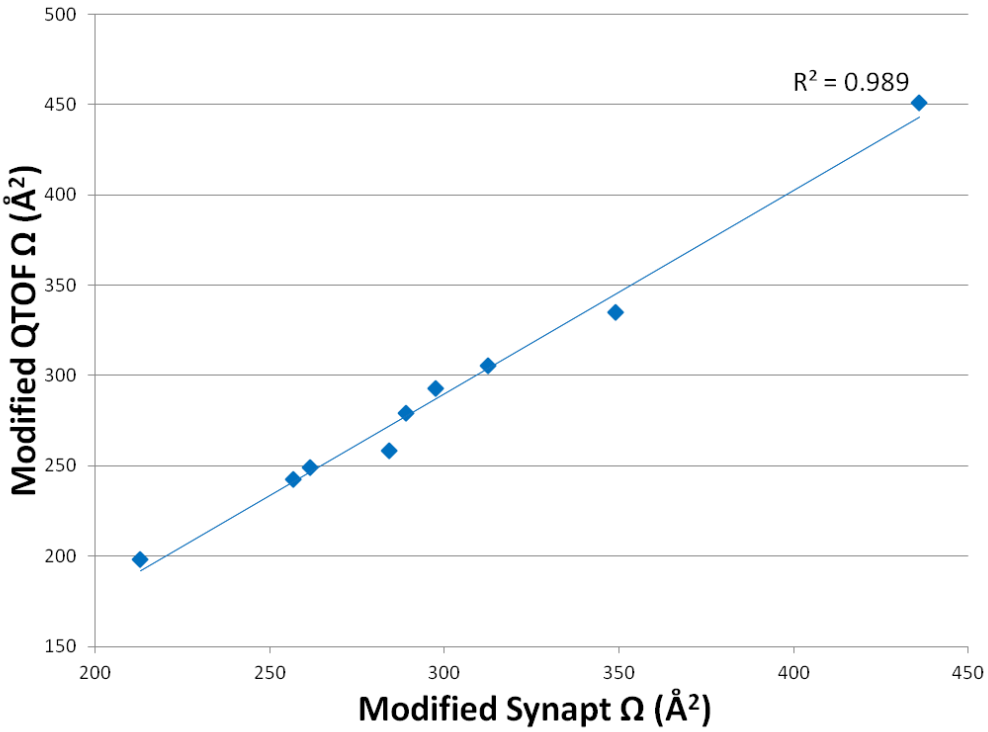

Figure 20. Linear regression comparing experimentally determined Ω of doubly charged peptide ions on both the modified Synapt and QTOF Linear field mobility instruments.

3.2.4. Assessment of QCAL-IM peptides for T-Wave Mobility Calibration

Having defined collision cross section values for the QCAL-IM peptides it was necessary to determine whether the standard could be used for calibration of the T-Wave mobility instrument. As such, the same proteolysed samples were analysed on a standard Synapt HDMS instrument with a T-Wave mobility cell. The drift time of each observed ion was plotted against *m/z* generating the charge state, mobility separated trend lines shown in Figure 21.

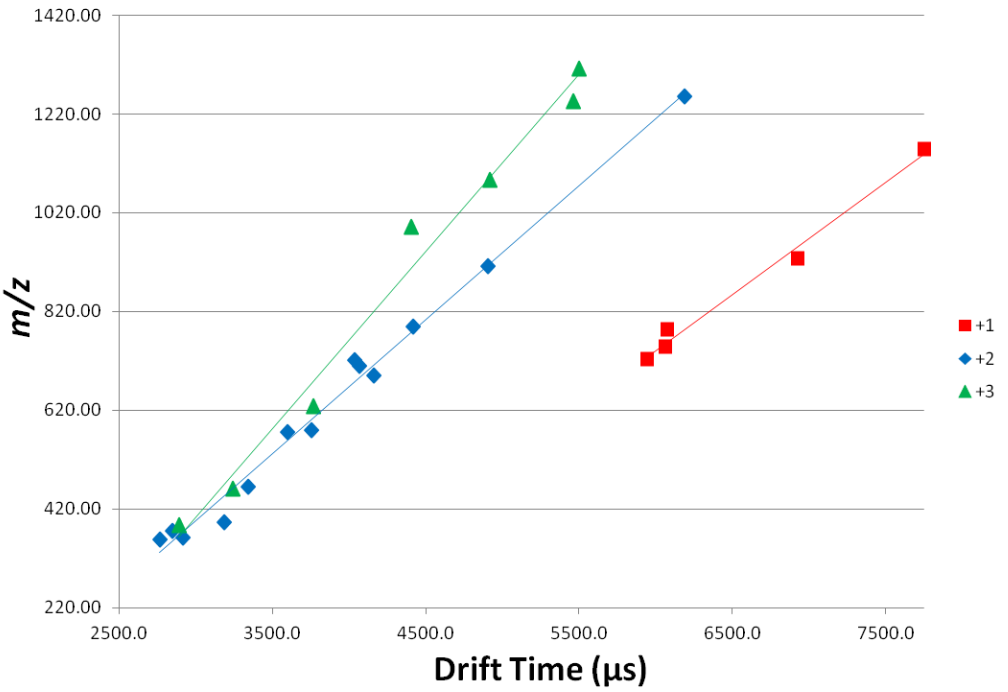

Figure 21. Plot of drift time versus *m/z* for QCAL-IM tryptic and Lys-C peptides following T-Wave mobility separation in nitrogen drift gas.

Using these T-Wave drift times measured in nitrogen drift gas and the empirically determined average cross sections detailed in Table 1 it was possible to construct calibration lines for charge states +1 to +3. Each line yields a linear curve fit with an R^2 value of >0.99 as shown in Figure 22. However, the observed linearity differs from the expected near square root dependence [131, 135]. This is likely to be a result of the T-Wave voltage ramp employed (rather than a fixed wave height) in addition to the range of drift times being analysed; with the QCAL-IM peptides comprising a near-linear sub-section of this square root dependence between drift time and cross section. Figure 23 is a plot of T-Wave drift time versus Ω' (determined from Ω adjusted for the charge state and reduced mass of the ions [120]) which exhibits good linear correlation (R^2=0.982) and provides a more generic calibration curve. Calculation of Ω' enables the correlation between charge states to be assessed, in essence the parameter removes charge state effects toward drift time in addition to any contribution from differences in reduced mass. The plots shown in Figures 22 and 23 therefore enable the rotationally averaged collision cross section of any peptide/peptide fragment ion that falls within the covered calibration range to be empirically determined from the observed T-Wave drift time.

Having constructed calibration lines (Figure 22) which facilitate determination of Ω in helium buffer gas from T-Wave drift time in nitrogen buffer gas, the effect toward the relative drift time by changing the drift gas was explored. It has been reported that alternative gases, in particular carbon dioxide [136], show potential for increasing the achievable resolution of Ion Mobility separation of some species. As such, the proteolysed QCAL-IM samples were again analysed using the Synapt HDMS instrument but with carbon dioxide gas in the T-Wave separation device. Unfortunately, the experiment was unsuccessful, likely because the increased density of carbon dioxide compared to nitrogen resulted in both scattering and activation of the ion cloud upon entry to the T-Wave region. Extensive fragmentation and reduced transmission of the QCAL-IM peptide ions was observed and consequently it was not possible to construct calibration lines from the data obtained. The first generation Synapt HDMS instrument with which analysis was performed does not incorporate a Helium cell prior to the drift tube, a feature that is now incorporated into the second generation instrument [137]. This cell is known to minimise scattering and activation of analyte ions upon entry to the T-Wave mobility cell, allowing both increased gas pressures and gases of higher density to be used successfully.

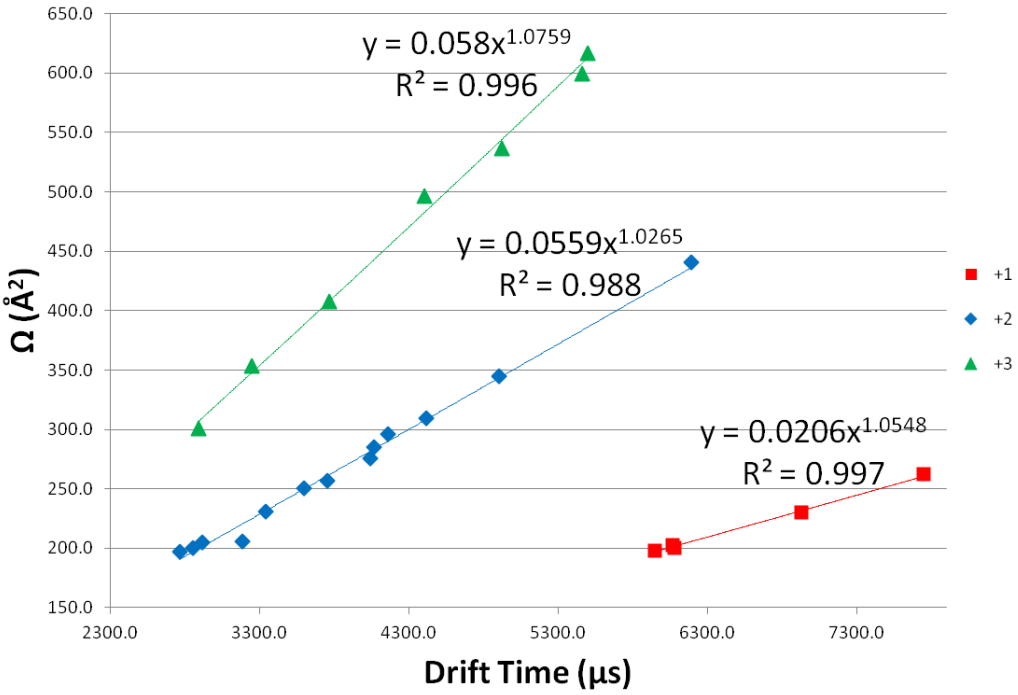

Figure 22. Drift time versus Ω charge dependent calibration lines for QCAL-IM peptides. Both vertical and horizontal error bars are smaller than any given data point.

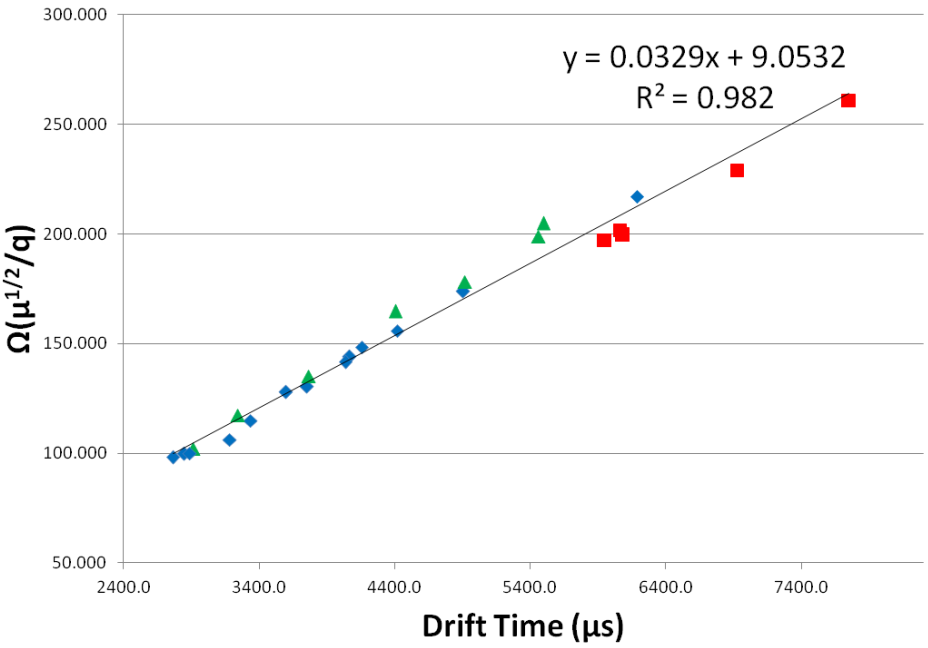

Figure 23. Plot of T-Wave drift time versus Ω' (charge and reduce mass corrected Ω) for QCAL-IM tryptic and Lys-C peptides.

3.2.5. Comparison to Literature Values

Of concern when using peptide cross sections determined on a Linear mobility platform utilising helium drift gas to calibrate the nitrogen T-Wave drift times are buffer gas specific effects upon ion mobility. In order to test the validity of T-Wave instrument calibration using QCAL-IM a selection of peptides whose cross section value have been published in the Clemmer database [122, 123] were analysed on the Synapt HDMS instrument and their cross section calculated by extrapolation of the derived calibration lines (Figure 22). Comparative data for the QCAL-IM derived cross section values of singly and doubly charged peptides and those from the literature are shown in Figure 24 and Tables 3 and 4 respectively. Excellent correlation (R^2=0.997, Figure 24) is observed between the two methods of cross section determination, with a mean difference of 3.9 % across 18 peptides analysed in triplicate. This excellent agreement between nitrogen cross section values derived from calibration of the T-Wave instrument and those measured directly by Linear field ion mobility in helium drift gas provides a high degree of confidence in the use of QCAL-IM toward cross section determination and would suggest that any buffer gas dependant effects can be accounted for in the calibration process.

Table 3. Comprison of published Ω [122, 123] with the average QCAL-IM calculated Ω (n=3) for a selection of $[M+H]^+$ peptides. The values in parentheses following the literature Ω indicate the associated standard deviation.

Peptide	$[M+H]^+$ (*m/z*)	Literature Ω ($Å^2$)	QCAL-IM Ω ($Å^2$)	Δ Ω ($Å^2$)	% Ω Difference
EK	276.16	97.89 (1.14)	92.2	5.7	5.9
LR	288.20	110.03 (1.44)	103.6	6.4	5.8
QR	303.18	104.07 (1.92)	98.4	5.6	5.4
FPK	391.23	123.14 (0.85)	116.1	7.1	5.8
DTHK	500.25	142.74 (1.62)	136.4	6.4	4.5
VASLR	545.34	163.7 (0.92)	154.2	9.5	5.8
AWSVAR	689.37	185.37 (0.81)	180.9	4.5	2.4
ATEEQLK	818.43	206.40 (1.81)	204.0	2.4	1.2
HLK	397.26	133.05 (0.71)	130.1	3.0	2.3
ALELFR	748.44	202.99 (0.33)	198.4	4.6	2.2
ELGFQG	650.31	172.29 (0.14)	168.5	3.8	2.2
NDIAAK	631.34	173.78 (0.84)	170.6	3.2	1.8

Table 4. Comparison of published Ω [122, 123] with the average QCAL-IM calculated Ω (n=3) for a selection of $[M+2H]^{2+}$ peptides. The values in parentheses following the literature Ω indicate the associated standard deviation.

Peptide	$[M+2H]^{2+}$ (*m/z*)	Literature Ω ($Å^2$)	QCAL-IM Ω ($Å^2$)	Δ Ω ($Å^2$)	% Ω Difference
LVTDLTK	395.24	210.68	195.0	15.7	7.5
FKDLGEEHFK	625.31	282.95	261.9	21.1	7.5
LVNELTEFAK	582.32	267.73	249.7	18.0	6.7
HLVDEPQNLIK	653.36	290.36 (0.83)	272.9	17.5	6.0
RHPEYAVSVLLR	720.41	313.44 (2.35)	291.8	21.6	6.9
ASEDLKK	395.72	238.09	197.2	40.9	17.2
LFTGHPETLEK	636.33	286.20 (4.02)	282.6	3.6	1.2
HGTVVLTALGGILK	689.92	312.22 (0.49)	297.1	15.1	4.8
VEADIAGHGQEVLIR	803.93	340.95 (1.0)	333.4	7.5	2.2
GLSDGEWQQVLNVWGK	908.45	352.48 (0.52)	343.6	8.9	2.5

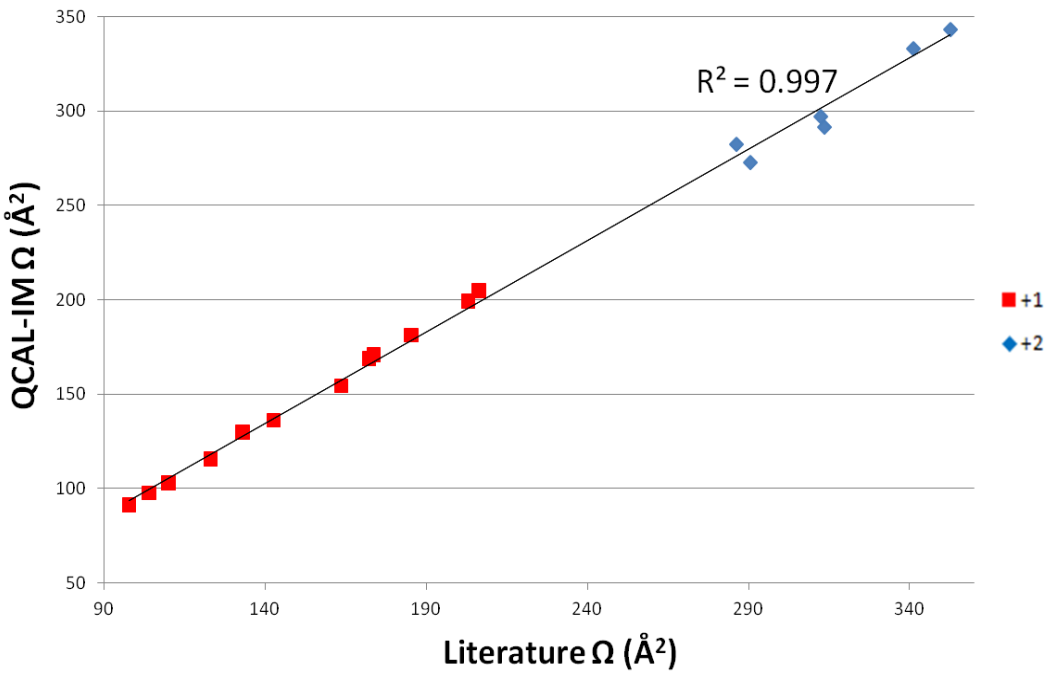

Figure 24. Linear regression of published Ω [122, 123] against average calculated Ω (n=3) following calibration of the T-Wave instrument with QCAL-IM for a selection of $[M+H]^+$ and $[M+2H]^{2+}$ peptides. Peptides without replicate measurement (n=1) in the literature values were excluded from this plot.

3.2.6. Assessment of Mobility Separation Resolution

Peptides chosen for inclusion within the QCAL-IM standard were selected, as previously discussed, using empirical data obtained from T-Wave analysis of a range of tryptic digestion mixtures. It was originally planned to determine the Linear field collision cross section of peptide ions observed in the charge states +1 to +4 so that instrument calibration across the same range would be possible. Sequences selected were concatenated in a manner so that the +4 charge state calibrant peptides would be produced by Lys-C or Asp-N digestion of QCAL-IM, although no preliminary T-Wave analysis of such species was acquired. Upon analysis of the Lys-C peptide mixture it was found that larger peptides, especially those observed in the +4 charge state, exhibit heterogeneous populations with respect to drift time, indicating the presence of more than one conformation. These ions cannot therefore be easily used to create a drift time calibration line since the conformers may not, dependent upon

instrument resolution, always be readily distinguishable. Consequently, the cross sections of the +4 species were not determined by Linear field mobility as the peptides could not be confidently employed for T-Wave calibration. These quadruply charged species can however be utilised, under a given set of experimental conditions, as internal references to assess the relative resolution of the mobility separation or for comparison of mobility resolution between instruments (Figure 25). Initial analysis of these species on the Synapt HDMS instrument (Figure 25A) identified a broad distribution in arrival times. Subsequent analysis of the same ions on the Synapt G2 instrument, which offers significantly enhanced mobility resolving power over the first generation instrument, enabled separation of these multiple conformations to give discrete peaks (Figure 25B).

Such analysis can thus be used as a performance check for the mobility cell, setting a standard for the minimal resolving power that should be achieved. The peak-to-peak resolution (R_{pp}) is calculated using the following equation:

$$R_{pp} = \frac{2\Delta t_d}{(wb_1 + wb_2)} \quad (8)$$

where Δt_d is the difference in drift time between the peaks of interest, and wb_1 and wb_2 are the measured peak widths at half height.

The resolving power value (R_{pp}) obtained from such calculations is presented in Table 5, demonstrating resolving power in excess of 4.3. Analysis of these highly charged peptide ions derived from QCAL-IM thus enables the resolution performance of the instrument to be monitored.

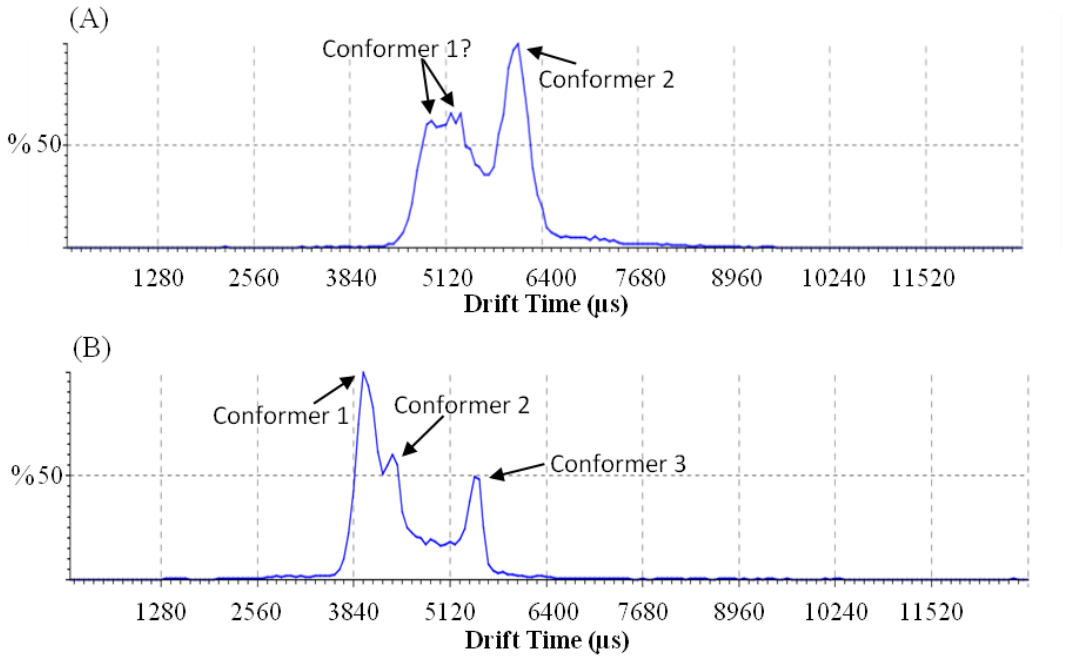

Figure 25. Mobility separation on the first generation Synapt HDMS instrument (A) and the higher resolution Synapt G2 (B) of the $[M+4H]^{4+}$ peptide T19/20 at 1266.64 *m/z* produced from Lys-C proteolysis of QCAL-IM. Three conformers are observed, of which conformers 1 and 3 are sufficiently separated to be used for the resolving power calculation.

Table 5. Resolving power value for the separation of multiple conformers from quadruply charged Lys-C peptide T19/20 at 1266.64 *m/z*. Values shown for the apex of each conformer are an average of 6 repeat analyses.

m/z	Drift time (bins) Conformer 1 Apex	Drift time (bins) Conformer 2 Apex	Resolving Power
1266.64	62.28	85.26	4.36

3.3. Conclusions

QCAL-IM (Figure 18) is a designer protein that upon proteolysis with the enzymes trypsin and Lys-C yields a stoichiometric mixture of specific peptides which enable calibration of collision cross section from travelling wave ion mobility data. Together these peptides enable the observed drift times from analysis on a T-Wave ion mobility instrument to be utilised for empirical determination of peptide and peptide fragment ion collision cross sections from a single protein standard. This provides a robust alternative to the current widely used internal calibration against the database previously published by Clemmer [122, 123]. Data have been obtained for the calibrant peptides included within the standard from both Linear and T-Wave mobility instruments in helium and nitrogen drift gases respectively. Application of QCAL-IM as an ion mobility standard enables direct calculation of cross sectional areas for charge states +1 to +3 under a variety of T-Wave experimental conditions. In addition, the standard can be used to monitor the performance of the IMS cell for separation of species with similar mobilites through calculation of the peak-to-peak resolution achieved from separation of multiple conformers of quadruply charged peptides. Finally, as the recombinant protein standard can be produced in limitless supply, proteolysed stocks of the QCAL-IM standard can be stored ready for analysis alongside the analyte of interest.

Chapter 4

The Influence of a *C*-terminal Basic Residue on Peptide Fragmentation Pathways

4.1. Introduction

MS/MS analysis of peptides produced from tryptic proteolysis of a protein is typically performed during a 'bottom-up' proteomic workflow. Product ion data resulting from decomposition of the observed precursor ions is then utilised to infer sequence information and hence protein identity from whence they are derived. Trypsin cleaves peptide bonds on the carboxy-side of arginine and lysine residues, generating peptides ideal for study by MS/MS as they appear in the optimal *m/z* range for analysis and are commonly observed as doubly protonated species following ESI [4]. The highly basic arginine/lysine residue located at the carboxyl-terminus sequesters one of the available protons, while the proton from the amino-terminal amine group can become 'mobile' upon activation thus facilitating fragmentation by CID [73]. The advent of ETD, which exhibits highly efficient peptide fragmentation for multiply charged species, has led to growing implementation of the protease Lys-C (cleaving *C*-terminal solely to lysine residues). Resultant peptides are larger than those produced by tryptic digestion and may contain internal arginine residues enhancing the formation of species of higher charge state. The exoprotease CBPB selectively cleaves the *C*-terminal basic residues arginine and lysine from a polypeptide chain and such activity may be particularly useful following tryptic/Lys-C digestion of a protein.

Peptide decomposition chemistry in both CID and ETD is determined by the nature of the immediate precursor to fragmentation; an even electron protonated species and charged reduced radical cation respectively. Therefore in CID and ETD regimes, bond cleavage results in the formation (in most cases) of a pair of singly charged ions and an ion/neutral pair respectively. For a given fragment ion to be observed two criteria must be satisfied: formation of the ion must be energetically favourable and the ion must be stable to the conditions to which it is subjected during its residence in the mass spectrometer. It has previously been shown that singly charged *N*-terminal product ions are more susceptible to secondary fragmentation events during CID than their *C*-terminal counterparts [68]. This effect is attributable

to proton mobilisation within the respective fragment ion series, with the presence of a highly basic residue at the *C*-terminal minimising mobilisation within the y-ion series. This leads to an under-representation of *N*-terminal fragments in product ion spectra, particularly those collected on QTOF instruments rather than QIT instruments (due to the promotion of secondary fragmentation events resulting from multiple ion-neutral collisions experienced in the QTOF regime). The fragment ions most likely to be observed in a given CID MS/MS analysis are therefore determined by the amino acid composition (and hence gas phase ion chemistry) of the peptide in question and the instrument upon which analysis is performed.

When considering dissociation of singly charged radical cations formed by electron transfer to doubly charged precursors, the observation of product ions reflects the respective gas phase basicity of the complementary fragments. Dissociation of tryptic/Lys-C peptides by ETD typically results in preferential generation of z-ion fragments over their c-ion counterparts. This is attributable to the presence of the basic residue at the *C*-terminus; meaning the proton affinity of *C*-terminal fragment ions is greater than that of *N*-terminal fragment ions. Consequently, the z-ion series is more likely to retain charge post fragmentation and is therefore more highly represented in ETD MS/MS spectra.

CBPB treatment of such peptides could be expected to alter the observed bias in z-ion formation as the *C*-terminal basic residue is no longer available for charge retention. A similar trend may also be observed in CID MS/MS spectra as y-ion species generated from tryptic/Lys-C peptides are known to be more stable than their b-ion counterparts due to the highly basic *C*-terminal residue sequestering an available proton (as discussed above). Removal of this site of high gas phase basicity will influence charge site distribution in *C*-terminal fragment ions, with a consequent change in the relative stabilities of b- and y-series ions. The products of tryptic/Lys-C digestion of a protein and their CBPB treated analogue may therefore yield data of complementary value when analysed by MS/MS. The objective of the work presented in this chapter was to investigate the influence of a *C*-terminal basic residue upon observed peptide fragmentation pathways by CID and ETD.

4.2. Results and Discussion

BSA was digested overnight with trypsin or Lys-C prior to an aliquot of the peptide mixture being removed and subjected to secondary enzymatic treatment with CBPB. The samples were then analysed by nLC MS/MS on both a QIT and QTOF instrument. Data produced by this nLC MS/MS analysis are shown in Figures 26-28 and these spectra from peptides observed as doubly charged species both prior to and following CBPB treatment have been chosen to demonstrate the general effect upon peptide decomposition following removal of the *C*-terminal basic residue.

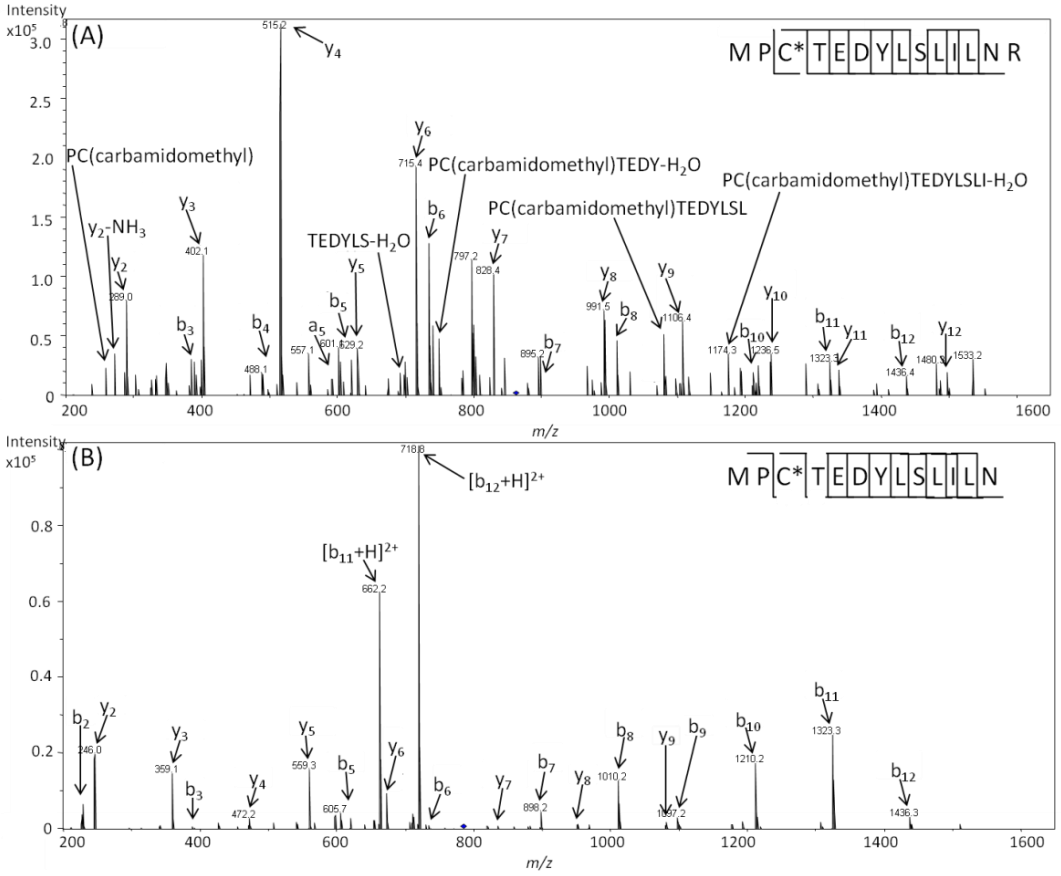

Figure 26. QIT product ion spectra generated by (A) CID of doubly charged tryptic peptide MPC*TEDYLSLILNR at 862.91 *m/z* and (B) CID of doubly charged CBPB treated analogue MPC*TEDYLSLILN at 784.87 *m/z*. In each instance C* represents a carbamidomethylated cysteine residue.

The CID MS/MS spectrum of peptide MPC*TEDYLSLILN (Figure 26B) shows an unusual propensity for the formation of doubly charged b-ions. Equivalent analysis of the tryptic analogue of the same sequence (Figure 26A) does not result in the production of these species; this difference can be attributed to the presence of the arginine residue at the *C*-terminus of the tryptic peptide which sequesters one of the available protons within the y-ion series formed upon CID. Removal of the highly basic arginine residue through CBPB treatment causes both a reduction in proton affinity of the *C*-terminal fragment ions and 'mobilisation' of the proton across the peptide backbone with retention by the *N*-terminal b-ion series now possible. Significant doubly protonated b-ions are only observed for sequences incorporating 10 or more amino acid residues, with the b_{11} and b_{12} (base peak) species being the two most intense ions within the MS/MS spectrum. Associated signal intensity of these doubly protonated species rapidly increases with the number of amino acid residues, suggesting that their stability is dictated by the corresponding decrease in coulombic repulsion between the protons accommodated within the b-ion structure.

Figure 26 shows there are additional differences in the observed CID fragmentation pathways following secondary enzymatic treatment with CBPB. Both the b_2 and b_9 fragment ions are present in the MS/MS spectrum of MPC(carbamidomethyl)TEDYLSLILN (Figure 26B) which are not seen for the tryptic analogue. Formation of the b_2 ion is unexpected as bond cleavage *C*-terminal to a proline residue is required; such activity had previously been thought to be disfavoured due to the steric strain implicit with formation of a bicyclic oxazolone structure incorporating a proline residue [138, 139]. However, a study by Hopkinson and colleagues [140] found this destabilisation to be negligible. Instead, they report that the presence of a prolyl residue at position 2 in the tripeptide Gly-Pro-Gly slightly increases the barrier to b_2 ion formation while simultaneously decreasing the barrier to y_2 ion formation when compared to the tripeptide Gly-Phe-Gly. The published study does however support the observation of the b_2 ion seen in Figure 26B as they report that a range of tripeptides containing a prolyl residue in the central position readily fragment to form both the b_2 and y_2 products. Observation of this b_2 ion is attributed to the relatively higher basicity of the secondary amide incorporating the prolyl nitrogen. Protonation at the *N*-terminal bond is competitive, but the $C\text{-}N^+$ bond of the secondary amide is sterically crowded when compared to

that of a primary amide; the bond is therefore weaker and cleavage *N*-terminal to the prolyl residue results. Extrapolation of these observations in relatively small tripeptide systems to the larger sequences being discussed here must however be performed with caution. A study conducted in the Wysocki laboratory [141] investigated the corresponding bond cleavage *C*-terminal to proline in larger tryptic peptides and found evidence suggesting that proline containing b_2 ions may be formed via the diketopiperazine pathway rather than the more commonly expected oxazolone pathway. It is therefore possible that either an alternative mechanism for b_2 ion formation from the CBPB treated analogue is followed or simply that the energy barrier toward b_2 ion formation is reduced. In this instance the normally favoured bond cleavage *N*-terminal to proline is also not expected as a b_1 ion would result and this is not predicted by the oxazolone pathway.

A further notable difference between the spectra shown in Figures 26A and 26B is that intense internal ions are generated during CID of the tryptic analogue which are not observed as major product ions from the CBPB analogue. It has previously been reported that *C*-terminal y-ion fragments have an increased tendency to undergo secondary fragmentation to generate internal ions when compared to *N*-terminal b- or a-ions [54]. Such behaviour is primarily expected during CID performed in the multiple collision regime of a QTOF collision cell but can also be observed during QIT analysis. The reduced propensity for internal ion formation from the CBPB analogue can therefore be rationalised by considering the less favourable formation of *C*-terminal y-ion fragments following removal of the highly basic *C*-terminal arginine residue. As a consequence the number and intensity of y-ion fragments which can be viewed as intermediates leading toward facile internal ion formation are reduced and consequently such species are less frequently observed.

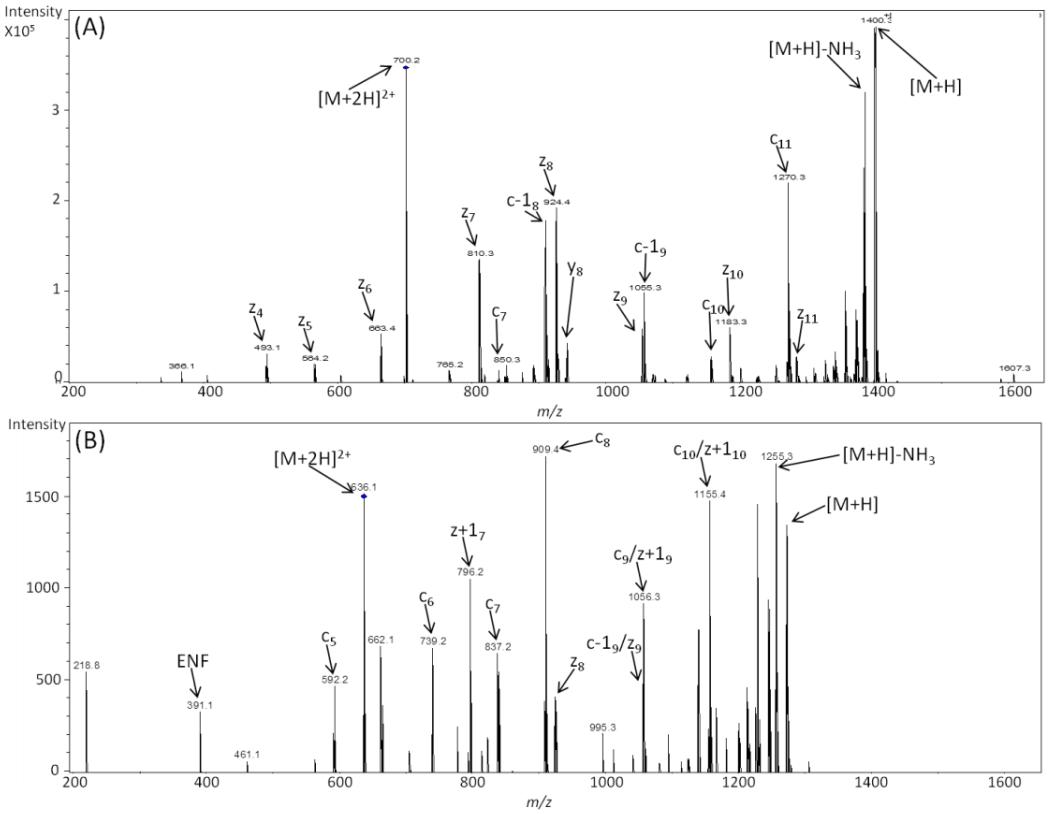

Figure 27. Comparison of QIT MS/MS spectra generated by (A) ETD of the doubly charged tryptic peptide TVMENFVAFVDK at 700.35 *m/z* and (B) ETD of the doubly charged CBPB treated analogue TVMENFVAFVD at 636.29 *m/z*.

The observed fragmentation products resulting from CID and ETD of the CBPB treated peptide analogues (Figures 26B and 27B respectively) exhibit an expected phenomenon. Removal of the *C*-terminal basic residue results in increased prominence of the *N*-terminal b- and c-ion series in each case. Prior to CBPB treatment the corresponding y- and z-ion series dominate the respective MS/MS spectra both in terms of frequency of observation and in the majority of cases, signal intensity. Following removal of the *C*-terminal basic residue the *N*-terminal ion series are typically seen with greater frequency and in several cases with higher relative signal intensity when compared to the product ion spectra of the initial sequence. This effect is particularly prevalent when considering ETD spectra of doubly charged tryptic precursor ions. The observed domination of ETD MS/MS spectra of tryptic peptides (Figure 27A) by *C*-terminal product ions can be attributed to the high proton affinity of the arginine/lysine residue located within the z-ion

series. Following acceptance of an electron during the ETD process, competition for retention of the available proton during dissociation takes place and the z-ion series is that which is most likely to remain charged. Removal of the *C*-terminal basic residue would appear, dependent upon amino acid sequence, to result in at least a partial switch in the dominant ion series, with the peptide *N*-terminal region now being the site of highest gas phase proton affinity.

This switch in the dominant product ion series is also observed following CID of certain peptide sequences. The effect is especially prominent in the QTOF MS/MS spectra produced by CID of the doubly charged peptide LKPDPNTLC(carbamidomethylated)DEFK, derived from Lys-C digestion and the cognate CBPB treated analogue shown in Figure 28.

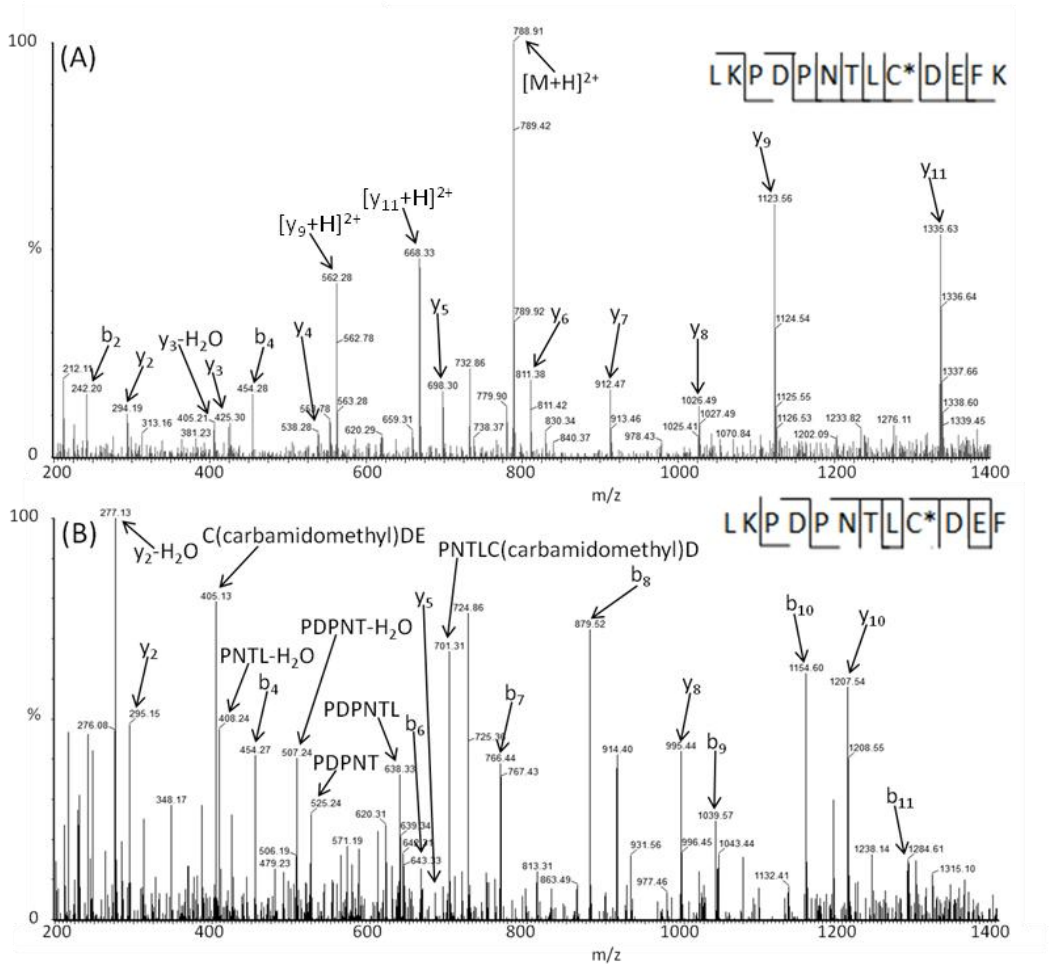

Figure 28. Comparison of QTOF MS/MS spectra generated by (A) CID of doubly charged tryptic peptide LKPDPNTLC*DEFK at 788.88 *m/z* and (B) CID of doubly charged CBPB treated analogue LKPDPNTLC*DEF at 724.84 *m/z*. In each instance C* represents a carbamidomethylated cysteine residue.

The peptide incorporates an internal lysine residue close to the *N*-terminus (arising from a KP sequence where proteolysis by serine proteases is sterically hindered) and the presence of this highly basic amino acid appears to assist the apparent switch in dominant product ion series following removal of the *C*-terminal residue. The lysine terminating product ion spectrum shown in Figure 28A contains two *N*-terminal b-ion fragments and nine *C*-terminal y-ion species whereas that of the CBPB truncated analogue (Figure 28B) contains seven b-ions and just four y-ion fragments, exemplifying the switch in bias of observed product ions resulting from removal of the highly basic residue. In addition, the doubly charged y-ions prominent in the MS/MS spectrum of the Lys-C precursor are not observed in that of the CBPB analogue. The absence of these species can be attributed to the converse effect which resulted in the generation of the doubly charged b-ions seen in Figure 26B. The y-ion series now deficient in a highly basic residue can no longer accommodate two protons due to reduced proton affinity and increasing coulombic repulsion between charges. As such, the MS/MS spectrum is heavily populated by *N*-terminal fragment ions and doubly charged *C*-terminal fragment ions are no longer stable and are therefore not observed.

To enable the validity of these generalisations toward peptide fragmentation behaviour to be assessed, MS/MS data from multiple tryptic and Lys-C peptides in addition to their CBPB truncated analogues were collected. Peaklists from both the CID and ETD MS/MS spectra on the QIT platform were generated using Data Analysis software (Bruker) and the resultant .mgf files searched using Batch-Tag-Web in Protein Prospector [142] version 5.9.4. The precursor charge state range was set to consider +1 to +4 species with a precursor mass tolerance of 0.4 Da and a fragment mass tolerance of 0.8 Da. A script written by R. Chalkley at the University of California that has been described in detail elsewhere [143] was used to identify all product ions which match the expected theoretical fragment masses from the observed precursor ions. The script considers the sequence ions b, a and y in CID spectra and potential neutral losses of $-H_2O$, $-NH_3$ and $-SOCH_4$ from these ions and the precursor from which they derive. In the case of ETD spectra the script considers the potential sequence ions b, $c\text{-}1^{\cdot}$, c, y, $z^{\cdot}$ and z+1. Those product ions acknowledged as originating from the observed precursors were then collated, combined and

analysed using excel. The ion statistics produced are shown in Tables 6 and 7 and these represent the percentage contribution of each fragment ion type toward the total ion current (TIC) from nLC QIT MS/MS analysis of the proteolytic cleavage products of BSA.

Table 6. Summary of CID fragmentation behaviour following different digestion regimes, with analysis by ESI-QIT MS/MS. The values shown are the percentage of total product ion current from all precursors observed both prior to and following CBPB treatment.

Ion Type	Lys-C	Lys-C CBPB	Trypsin	Trypsin CBPB
MH-$SOCH_4$	0.1	0.3	0.1	0.0
MH-H_2O	0.3	0.3	0.0	0.0
MH-NH_3	0.3	0.3	0.0	0.0
b	19.8	33.0	22.2	32.6
a	6.3	4.9	5.3	6.8
y	37.8	31.6	35.6	29.4
b-H_2O	11.4	11.0	10.5	10.7
b-NH_3	4.3	4.0	5.3	5.7
y-H_2O	10.4	9.9	10.5	11.8
y-NH_3	8.6	4.1	9.9	2.9
y-$SOCH_4$	0.6	0.3	0.4	0.0
a-$SOCH_4$	0.0	0.3	0.0	0.0
Total Precursor Ions	**29**	**14**	**30**	**12**
Precursor Charge States	**n=25 for z=2 n=4 for z=3**	**n=12 for z=2 n=2 for z=3**	**n=21 for z=2 n=9 for z=3**	**n=11 for z=2 n=1 for z=3**
Total Fragment Ions	**780**	**328**	**902**	**279**

N.B. Ion types whose percentage ion current does not exceed 0.2 % in any analysis have been omitted to minimise random matching.

Table 7. Summary of ETD fragmentation behaviour following different digestion regimes, with analysis by ESI-QIT MS/MS. The values shown are the percentage of total product ion current from all precursors observed both prior to and following CBPB treatment.

Ion Type	Lys-C	Lys-C CBPB	Trypsin	Trypsin CBPB
MH+1	0	0	0.1	2.5
MH-17	0.4	0.0	0.0	0.0
c	25.0	32.1	28.5	30.0
z	31.2	21.0	31.9	12.5
z+1	15.4	8.0	13.2	5.0
c-1	15.8	29.0	15.6	42.5
b	4.0	4.9	2.0	0.0
y	8.0	4.9	8.4	7.5
Total Precursor Ions	**25**	**9**	**32**	**3**
Precursor Charge States	**n=8 for z=2 n=10 for z=3 n=7 for z=4**	**n=6 for z=2 n=3 for z=3**	**n=11 for z=2 n=16 for z=3 n=5 for z=4**	**n=3 for z=2**
Total Fragment Ions	**696**	**162**	**813**	**40**

N.B. Ion types whose percentage ion current does not exceed 0.2 % in any analysis have been omitted to minimise random matching.

The CID datasets derived from analysis of tryptic and Lys-C peptide digestions exhibit a similar increase in the percentage of b-ion fragments observed following CBPB treatment with 10.4 % and 13.2 % increases respectively. There is a corresponding decrease of 6.2 % in the occurrence of y-ions observed in both the tryptic and Lys-C analyses. Removal of the *C*-terminal basic residue from Lys-C peptides results in a 7.1 % increase in the number of c-ion fragments generated by ETD. However, a similar increase in the number of c-ions is not demonstrated following ETD of tryptic precursor ions, with a negligible increase of 1.5 % observed. This is likely to be due to the small number of tryptic peptides which are observed as doubly charged (or greater) species following treatment with CBPB, meaning that only a small number of peptides remain amenable to study by ETD. Those CBPB truncated species which are observed as doubly charged species possess an extremely low charge density and thus poor fragmentation efficiency by ETD is typically observed. This is highlighted by the 2.5 % of the ion current from the tryptic CBPB analysis which is accounted for by charge reduced species $[M+2H]^{+\cdot}$. These species result following initial acceptance of an electron but where no consequent dissociation takes place and are an indication that the CBPB treated tryptic peptides have greatly reduced ETD efficiency.

Interestingly, the ion statistics from the ETD analyses show that the propensity for hydrogen atom transfer differs greatly in the tryptic/Lys-C peptides when compared to their CBPB treated analogues. Following removal of the *C*-terminal basic residue, an increase from 15.6 % to 42.5 % of the ion current is attributable to c-1 species for trypsin derived precursor ions and there is a corresponding increase from 15.8 % to 29.0 % for Lys-C derived sequences. The increased observation of c-1 ions is accompanied by a decrease in the formation of z+1 species with the associated ion current falling from 15.4 % to 8.0 % (Lys-C peptides) and 13.2 % to 5.0 % (tryptic peptides), contrary to the expectation of a concomitant increase in z+1 ion formation. This effect is illustrated by the spectra in Figure 29 where the apparent isotope distributions (reflecting the distorting effect of hydrogen migrations accompanying fragmentation) of product ions resulting from ETD of the Lys-C peptide EC*C*HGDLLEC*ADDRADLAK, where C* represents carbamidomethylated cysteine, and its CBPB treated analogue are shown. In this instance the Lys-C derived precursor is observed as a quadruply charged species and the CBPB

truncated precursor is observed as a triply charged ion. Figure 29A shows the apparent isotope distribution of the c_9 product ion (1131.4 monoisotopic *m/z*) and no evidence of hydrogen atom transfer is visible. However, the corresponding apparent isotope distribution of the c_9 ion resulting from ETD of the CBPB treated precursor (Figure 29B) incorporates a significant peak indicating that hydrogen transfer has taken place. This change in propensity for hydrogen atom transfer is further demonstrated by the spectra in Figures 29C and 29D, relating to the z_{12} and z_{11} ions from the Lys-C and CBPB precursors respectively. These ions are both produced from the same bond cleavage (between two leucine residues) and the apparent isotope distribution in Figure 29C indicates significant z+1 ion formation whereas that in Figure 29D does not.

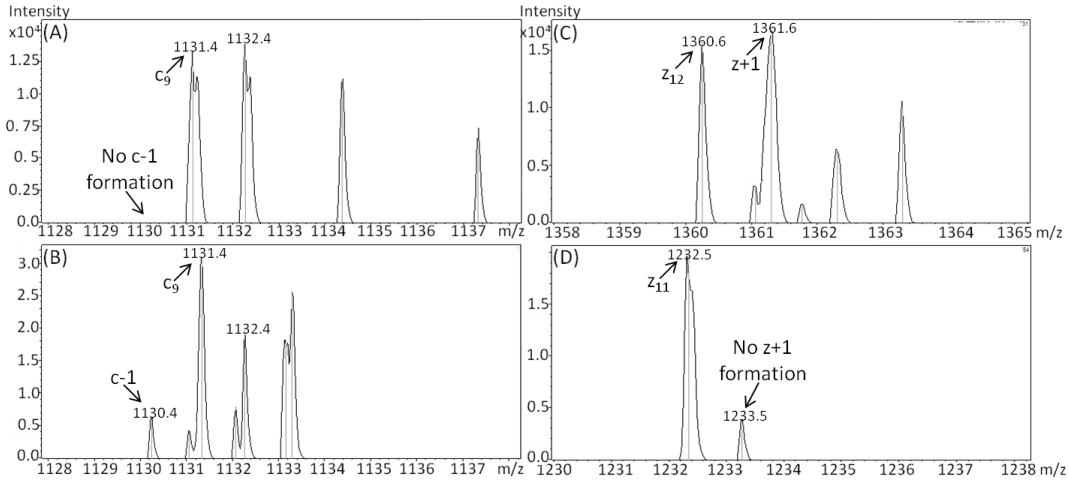

Figure 29. Comparison of product ion apparent isotope distributions resulting from ETD of the Lys-C derived peptide EC*C* HGDLLEC*ADDRADLAK and the cognate CBPB treated analogue (where C* is carbamidomethylated C). (A) and (B) show the apparent isotope distributions of the c_9 product ion from the Lys-C and CBPB peptides respectively, with the CBPB treated analogue exhibiting increased c-1 ion formation. (C) and (D) show the corresponding apparent isotope distributions of the Lys-C peptide z_{12} product ion and CBPB peptide z_{11} product ion (relating to the same bond cleavage) with the Lys-C peptide exhibiting evidence of z+1 ion formation.

This altered propensity for hydrogen atom transfer is initially a surprising effect resulting from CBPB treatment. It is possible that the observed increase in proton abstraction of c-ions may be related to the greater likelihood that doubly protonated CBPB treated peptides will contain internal amino acids with basic side chains (His, Lys, Arg) necessary to yield multiply charged precursor ions for ETD fragmentation. The observed fragmentation trends can be rationalised with reference to a recent study by Nishikaze and Takayama [144] where they report evidence indicating that hydrogen atom transfer is highly dependent upon precursor charge state and amino acid composition, in particular the presence of basic sites. Doubly protonated species are shown to exhibit extensive hydrogen atom transfer to give $c^{+\cdot}$ ions (described in this thesis as c-1 ions) whereas a triply protonated form of the same sequence shows no hydrogen transfer. By extension, this situation is applicable to the example shown in Figure 29, as the quadruply protonated Lys-C precursor does not show evidence of hydrogen atom transfer whereas the CBPB analogue observed as a triply protonated ion does. In their study Nishikaze and Takayama also demonstrate that highly basic doubly charged peptides exhibit prevalent hydrogen atom transfer as a result of an extended lifetime of the dissociation complex formed between separating c- and z-ion fragments post N-C_α bond cleavage. The precursor ion can adopt a fairly compact structure as coulombic repulsion only exists between two protonation sites. In addition, internal hydrogen bonding toward the basic residue assists in stabilising the intermediate complex and hence facilitates prevalent hydrogen atom transfer. The absence of hydrogen transfer for triply charged precursor ions is attributed to the less compact structure of such species due to coulombic repulsion between three protons accommodated along the backbone. Consequently, the dissociation complex has reduced stability due to this repulsion and the fragments separate on a timescale shorter than that required for hydrogen atom transfer to occur. In the context of the data presented here, the reported differences in the behaviour of peptide ions of differing charge state upon ETD accounts for the increased observation of c-1 ions following CBPB treatment of both tryptic and Lys-C peptide sequences. The data shown in Tables 6 and 7 are derived from a mixture of precursor ion charge states as CBPB treatment of certain sequences results in a reduction of the observed charge state whereas for other sequences it does not. The increased propensity for hydrogen atom transfer giving a more frequent occurrence of c-1 ion formation following CBPB treatment can therefore be explained by the reduced charge density of these

species and concomitant switch in the terminal within the peptide which has the highest proton affinity. An extended lifetime of the intermediate dissociation complex results, hydrogen transfer becomes more prevalent and the *N*-terminal fragment ion series is more likely to retain the available charge as the *C*-terminal series is now deficient of a highly basic residue.

4.3. Data Analysis

The data suggest that sequential analyses, using both CID and ETD, of tryptic/Lys-C peptides and their products of CBPB treatment could provide a simple analytical strategy for improved determination of proteins during proteome characterisation. Consequently, a comparison between this approach and the simple duplication of nLC-MS/MS analysis as is commonly employed to reduce the "under-sampling" problem in automated analyses was conducted. The output files (.mgf) produced from nLC-MS/MS analysis of tryptic/Lys-C peptides and their CBPB truncated analogues were merged using a script written in-house. The script combines .mgf files generated from samples treated with consecutive proteases (trypsin or Lys-C) followed by CBPB. Fragment ion data is combined when the mass difference between two precursor ions (one from each analysis) corresponds to the expected mass shift following CBPB treatment (156.1 *m/z* for singly protonated R terminating peptides or 128.1 *m/z* for singly protonated K terminating peptides). Precursor ions are only considered to match if the observed mass difference falls within a given 300 ppm tolerance. Spectra for unmatched (unique) peptides from the CBPB treated analysis are then added, generating a combined .mgf file containing all MS/MS data from both analyses. Original output files from analysis of tryptic and Lys-C digestions of BSA both by CID and ETD were searched using MASCOT [11] alongside the combined files containing additional data from the CBPB treated sample. Files were searched against the SwissProt database, specifying Mammalian Taxonomy and considering carbamidomethylated cysteine residues as a permanent modification and methionine oxidation as a variable modification. Precursor charge states of +1 to +3 were considered and the precursor tolerance was set to 0.4 Da with an MS/MS tolerance of 0.8 Da. A comparison of the results produced from each search is shown in Table 8.

Table 8. Comparison of MASCOT search results produced from nLC-MS/MS analysis of tryptic/Lys-C digests of BSA and those produced by combined analyses incorporating additional data following secondary enzymatic digestion with CBPB.

	Single Protease		**Dual Protease**	
	Score	Peptide Queries Matched	Score	Peptide Queries Matched
CID Trypsin	1645	118	1460	118
ETD Trypsin	1147	90	1197	85
CID Lys-C	1182	71	1158	77
ETD Lys-C	888	69	925	67

The combined files failed to provide any improvement in MASCOT search score for CID analyses and provided only slight improvement in the score returned from ETD analyses. Consequently, a direct comparison between a combined .mgf from repeat trypsin/Lys-C nLC-MS/MS analyses and that generated from dual analysis following CBPB treatment was not conducted. The single protease scores would simply further increase making comparison to the dual protease analysis redundant. Sequential use of CBPB following initial proteolysis with trypsin/Lys-C does not provide sufficient (if any) increase in the level of information obtained and due to the additional experimental and instrumental effort required this makes it difficult to justify the approach as a valid technique in proteomic research. The reason for the reduced MASCOT search scores from CID data is likely due to an increased occurrence of 'random matching' from the combined file. The CID trypsin searches from the single and dual enzyme output files both matched 118 peptide queries and the dual CID Lys-C search matched an additional 6 peptide queries in comparison to the single protease file. However, both dual protease searches returned lower scores, presumably due to a greater number of fragment ions not matching those expected from the observed precursor ions.

4.4. Conclusions

'Bottom-up' proteomic studies utilise MS/MS analysis of proteins which have typically been subjected to proteolysis using the endoproteases trypsin or Lys-C. Resultant peptides in each instance terminate in the basic amino acid residues arginine or lysine and current understanding of peptide fragmentation trends and mechanistic principles are heavily influenced by the presence of these *C*-terminal basic residues. A range of product ion spectra generated by CID and ETD of peptide ions produced by ESI both prior to and following the removal of this *C*-terminal basic residue have been analysed, with notable differences in fragmentation trends observed. The data suggest that the observed dominant ion series from both CID and ETD MS/MS analyses is at least in part determined by the position within the polypeptide backbone of the residue with highest gas phase basicity. Also of consequence toward the degree to which a particular ion series dominates is the presence and position of highly basic internal arginine/lysine (and to a certain extent histidine) residues. Fragmentation at sites not observed in the tryptic/Lys-C precursor was observed for the CBPB treated analogue following removal of the *C*-terminal residue. In addition, removal of the *C*-terminal residue was found to influence the extent of hydrogen atom transfer during ETD analysis. As such, the data substantiate the notion that the simple sequential application of trypsin or Lys-C and CBPB digestion yields data of complementary value to the use of trypsin/Lys-C alone. However, through comparison of MASCOT search scores from combined output files following CBPB treatment to those produced by single enzyme analysis the approach does not appear to have long term applicability toward large-scale proteomic studies. The results discussed in this chapter do however further our understanding of peptide gas phase ion chemistry and assist the development of more sophisticated spectral interpretation algorithms. At present there is a high level of 'redundant' information within MS/MS spectra as complete understanding of complex peptide fragmentation pathways has not yet been achieved.

Chapter 5

Proposal for a Common Nomenclature for Peptide Fragment Ions Generated Following Sequence Scrambling during Collision-Induced Dissociation

As discussed in Chapter 1 b-ion species generated during CID can be observed as fragment ions which are not the product of simple bond cleavage. Consequently, as part of the research presented in this thesis a common nomenclature for assignment of sequence scrambled b-ion species was proposed. This nomenclature has been published [145] and the document is detailed below. My contribution to the work was the initial idea for the required nomenclature and writing the majority of the manuscript. The nomenclature was decided between myself and the other contributing authors Prof. Simon Gaskell and Dr. Claire Eyers.

Tandem mass spectrometry is often used in the analysis of peptides, generating sequence specific fragment ions. The characteristic fragment ions produced by a particular peptide can then be used to elucidate the original amino acid sequence. A nomenclature for the labelling of these fragments was originally devised by Roepstorff and Fohlman [51] and then later modified by Biemann [52]. Collision-Induced dissociation (CID) results in cleavage of the amide bond, producing either b- or y-ions, where the charge is retained on the amino or carboxy terminus respectively [53]. Subsequent loss of carbon monoxide from b-ions can also yield a corresponding a-ion series.

The structure of y-ions has been shown by the work of Mueller *et al.* [76] and Cordero *et al.* [77] to be that of a protonated amino acid (y_1) or a protonated truncated peptide (y_n). Work has been undertaken by various groups to provide a greater understanding of the mechanisms causing production of such species. Early work by Yalcin *et al.* [78] showed b-ions to consist of a linear peptide backbone terminating in a cyclic oxazolone ring. This ring structure is of particular interest as it hinders the formation of b_1 fragment ions under standard conditions. However, b-ion dissociation products are often observed in a form that cannot be explained by simple bond cleavage. Work by Vazquez *et al.* [146] showed formation of a macrocyclic intermediate from the initial oxazolone terminating b-ion, thought to be generated through attack of the *N*-terminal amine on the carbonyl carbon within the oxazolone ring. This macrocycle intermediate can undergo further fragmentation

following ring opening at various positions, to give non-sequence product ions via rearrangement. The generation of these non-native sequence product ions (sequence scrambling) has been supported by the work of other groups [70, 147, 148] who have demonstrated that b_N-ions (where $N \geq 5$) can form fully cyclic structures. Harrison [14] has shown that the abundance of non-sequence ions produced following macrocyclic ring opening increases with elevated CID collision energies. Interestingly, they demonstrated that peptide chain *N*-acetylation prevents such sequence scrambling, supporting the proposed mechanism of macrocycle formation. The presence of the acetyl group at the *N*-terminus is thought to prevent nucleophilic attack of the oxazolone ring, meaning macrocycle formation does not occur and subsequent rearrangement is avoided. Recent work by Maitre *et al.* [149] has demonstrated that a-ions can also undergo scrambling processes, whereas it was previously thought that such fragmentation pathways were exclusive to b-ions.

Elucidation of the mechanisms that govern scrambling of peptide ions [150-153] and the resultant consequences for large scale MS/MS analyses [154-156] are research areas of growing interest. However, description of the origins of these non-native peptide sequence ions is fraught with confusion which is further complicated by the absence of a uniform approach toward naming the scrambled products. Consequently, an extension to the currently used peptide fragmentation nomenclature [51, 52] was proposed, enabling fragment ions deriving from macrocycle formation to be easily assigned. The proposed nomenclature can be applied to the products of scrambled b-ions and enables both the fragment ion in question and the pathway from which it was derived to be assigned (Figure 30). The nomenclature can also be extended to the scrambling of a-ions. Importantly, this notation makes no assumption regarding the precise mechanism of scrambling.

To facilitate use of the proposed nomenclature for scrambled product ions, the peptide bonds in the initial sequence are numbered in accordance with standard nomenclature. A square bracket is used to contain the series of fragmentation processes that result in formation of the initial scrambled product. Inside this bracket, the point at which initial fragmentation has taken place, and hence where attack of the *N*-terminal residue forms the macrocycle is specified, followed by the amide bond number at which macrocycle ring opening occurs. Finally, (outside of the

square bracket) any ion resulting from further fragmentation of the scrambled product ion is assigned, again according to traditional nomenclature.

For example, using the peptide YAGFL that has been shown to exhibit sequence scrambling by Gaskell *et al.* [157], initial fragmentation to generate the b_4 ion (YAGF) with the oxazolone structure, is then followed by *N*-terminal attack of the oxazolone ring and macrocycle formation. If ring opening were subsequently to occur at bond 2 (giving the scrambled sequence GFYA), the process would be described inside the square bracket as [$b_4$2]. In the absence of further fragmentation, this scrambled product is isobaric with the initial b_4 ion and thus the two ions cannot be distinguished. However, subsequent fragmentation of this scrambled product (either during MS/MS in a QTOF where the products of multiple dissociation events are often observed, or at the level of MS^3 in a QIT) will yield a sequence determinable scrambled fragment ion; in the example above, fragmentation of the scrambled product to generate a b_3 ion (GFY) would yield a product ion with the nomenclature [$b_4$2] b_3. Should scrambling arise from an a-ion, this can be described either as [$a_4$2], if the precursor ion has been isolated prior to subsequent fragmentation, or as [$b_4$2]-28 should the mechanism of final product formation be ambiguous.

Further fragmentation/rearrangement events can be easily described by extension of this proposed nomenclature.

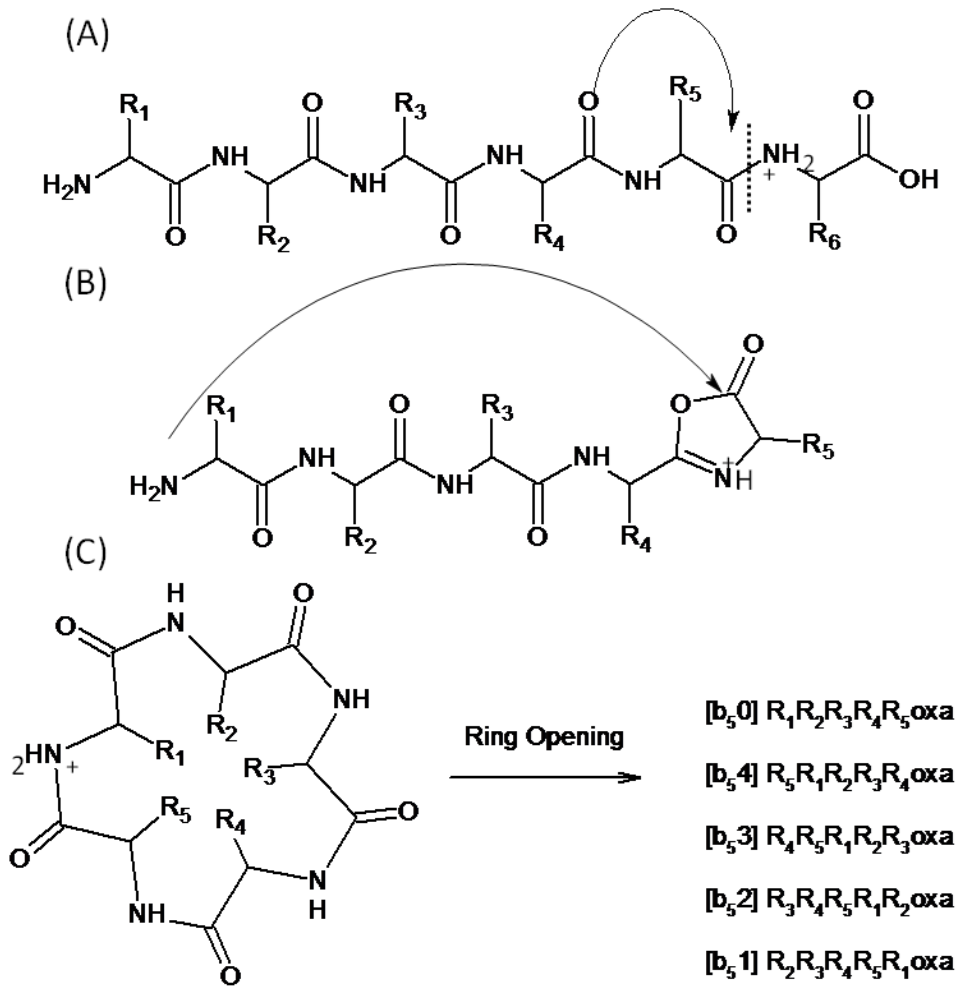

Figure 30. (A) The b_5 ion is generated by CID resulting in formation of the oxazolone ring through attack of the oxygen lone pair on the newly formed terminal carbonyl group. (B) *N*-terminal nucleophilic attack of the oxazolone ring to form the macrocycle. (C) The resultant macrocyclic intermediate undergoes ring opening at various positions to give sequence scrambled b_5 ions.

Chapter 6

The Effect of basic Amino Acid Residues upon b-ion Rearrangement during Collision-Induced Dissociation

6.1. Introduction

Proteomic studies typically employ MS/MS with CID for identification and quantification of large numbers of peptides and proteins [158, 159]. However there remain limitations in the assignment of fragmentation products due to incomplete understanding of peptide dissociation pathways. Low energy CID typically results in charge directed cleavage of amide bonds in accordance with the mobile proton model [72, 160]. The model describes the intramolecular transfer of a proton from a basic side chain to a heteroatom along the peptide backbone. Weakening of the peptide bond results and fragmentation to give diagnostic b-, y- and a-ions occurs. These ions are then used to determine the amino acid sequence of the unknown peptides and hence elucidate identity of the protein from whence they derive.

The structure of *C*-terminal y-ion fragments is recognised to be that of a truncated peptide (y_n) or protonated amino acid (y_1) however the structure of their b-ion counterparts is still debated. Initially b-ion structure was believed to be that of an acylium ion, however more recently the structure has been shown to incorporate a *C*-terminal oxazolone ring [78]. Nucleophilic attack by the *N*-terminal amine on the electrophilic carbonyl carbon within this protonated oxazolone ring results in the formation of a fully macrocyclic intermediate from the partially cyclic b-ion (Figure 30) [79]. It has previously been demonstrated that CID of protonated linear YAGFL-NH_2 results in the formation of a heterogeneous b_5 ion population with respect to drift time during analysis by TW-IMS [118]. This observation is consistent with the formation of a macrocyclic intermediate which is stable on the timescale of the ion mobility separation. Direct spectroscopic evidence for the existence of this macrocyclic intermediate has also been obtained from Infrared multiphoton dissociation (IRMPD) analysis of the b_5 ion produced from G_5R [161]. The macrocyclic intermediate can ring open at various amide bonds, yielding non-native, scrambled peptide sequences, whereby previously internal amino acid residues become exposed at the *C*-terminus. These rearranged species can then undergo secondary fragmentation events, particularly when analysed using a QTOF type

collision cell, producing a product ion series which is not representative of the initial peptide sequence.

Van Stipdonk and co-workers [153] reported that for the tetrapeptide YAFG and permuted isomers, CID of the b_3 ion does not result in the production of non-sequence ions, an observation rationalised by considering extensive steric strain following cyclisation. However, they did observe product ions indicative of sequence scrambling for larger species, ranging from pentapeptides to decapeptide methyl esters. The increased size and associated degrees of freedom of these larger species may be expected to disfavour macrocycle formation but the results are in accordance with the observations of Gaskell and co-workers who observed scrambling of the b_9 ion produced from the peptide neurokinin (DMHDFFVGLM-NH_2) [119]. Yalcin conducted a detailed study into the effect of acidic side chains upon peptide rearrangement pathways and found the position of glutamic acid and aspartic acid residues to have little effect upon the observed sequence scrambling [162]. Conversely, it has been reported by Harrison *et al.*[152] that the position of the basic amino acid histidine within permuted isomers of a hexapeptide sequence containing five alanine residues and a single histidine does influence the observed scrambling of the b_5 ion generated upon CID. When the histidine residue is located near the *N*-terminus of the peptide no experimental evidence for cyclisation was observed, however when histidine was located near the *C*-terminus of the b_5 ion, scrambling of the native sequence was seen. A similar study examining the effect of various amino acid side chains upon the ring opening of a putative macrocyclic b-ion structure was conducted by Molesworth *et al.* [151]. It has previously been shown that the presence of arginine and lysine residues can heavily influence observed decomposition pathways due to the formation of salt-bridging interactions [163, 164] or through nucleophilic attacks [164-166]. These factors might therefore be expected to be significant in the formation/consequent ring opening of the b-ion macrocycle. Interestingly, the study found that in contrast to peptides containing the nucleophilic amino acid lysine the presence of an internal arginine residue appears to inhibit the peptide scrambling fragmentation pathway. This observation is strongly related to the results discussed in this chapter and was followed by a more detailed study [150], which reported that the b_5^+ species generated from the sequence YARFLG and its permuted isomers produce only products of 'standard' fragmentation pathways upon

CID and that elimination of internal residues via sequence scrambling is not observed. The exact reason for this finding is undecided but was attributed to the high basicity of the arginine guanidine side chain which sequesters the available proton, preventing facile macrocycle formation.

In order to better characterise the positional effect of basic side-chains on the prevalence of scrambling this chapter assesses the fragment ions generated from a selection of peptides produced by Lys-N (cleaving *N*-terminal to lysine residues) and Lys-C (cleaving *C*-terminal to lysine residues) proteolysis of a range of 'standard' proteins subjected to CID. The effect of amino acid gas phase basicity toward b-ion scrambling was also assessed through modification of the lysine side chain located at the *N*-terminus of the Lys-N peptides. Reaction with *O*-methylisourea converts lysine residues into homoarginine producing a derivatised residue analogous to arginine, increasing the gas phase basicity of the *N*-terminal residue [167].

6.2. Results and Discussion

To study the effect of increasing the gas phase basicity of the amino-terminal residue upon sequence scrambling, BSA, enolase from *S. cerevisiae* and horse myoglobin were proteolysed overnight using Lys-N or Lys-C enzymes and an aliquot of the Lys-N digestion mixture incubated with *O*-methylisourea to convert lysine residues into homoarginine. It was then necessary to identify those peptides produced from Lys-N digestion which show evidence of rearrangement during CID. A selection of native Lys-N peptides and their corresponding derivatised sequence, now with an *N*-terminal homoarginine residue, were isolated and subjected to CID in the Trap collision cell of a Synapt HDMS instrument. As expected, increased Collision Energy was required to fragment the derivatised analogue due to reduced mobility of the proton now sequestered by the homoarginine residue. Figure 31 compares the CID MS/MS spectra produced from analysis of the native and derivatised forms of the peptide KGVLHAV. The spectrum produced from the native peptide (Figure 31A) contains several species attributable to scrambling of the original sequence. However, these rearranged product ions are not observed in the MS/MS spectrum of the derivatised peptide (Figure 31B). CID of the derivatised peptide does not result in the generation of the b_6 ion and hence the precursor to formation of many of the observed rearranged fragment ions is not present. There is however an intense peak

at 666.40 *m/z* which corresponds to the b_6+H_2O ion and such species have previously been shown by Thorne *et al.* to have a structure analogous to that of an intact peptide [64]. Consequently, scrambled species cannot be generated from this ion as the oxazolone moiety is required for the macrocyclic intermediate to form. However, the $[b_5 3]b_4$ scrambled species is also not observed upon CID (for which the b_5 precursor to sequence scrambling is observed) of the derivatised peptide and this may be related to the homoarginine residue now located at the *N*-terminus.

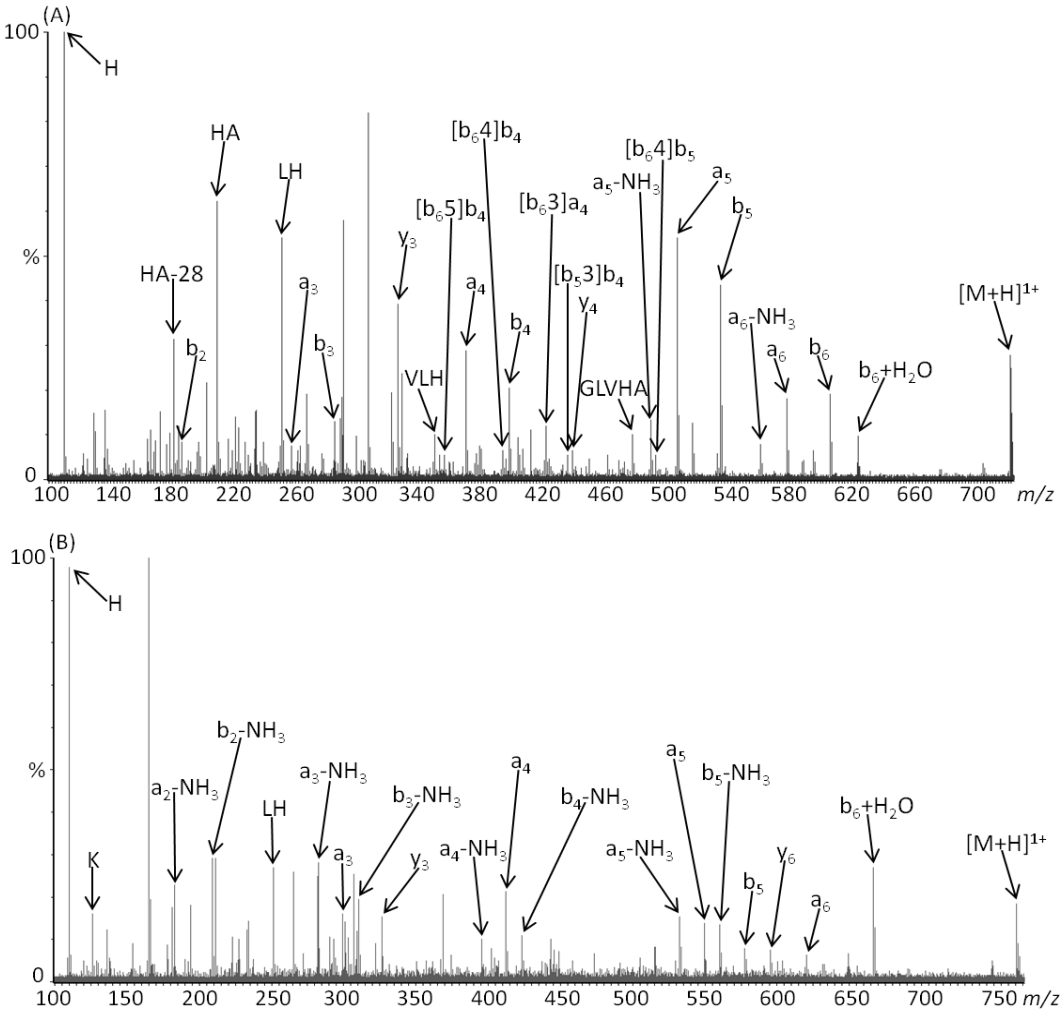

Figure 31. Comparison of QTOF MS/MS spectra resulting from (A) CID of $[M+H]^+$ peptide KGVLHAV at 723.45 *m/z* and (B) CID of $[M+H]^+$ peptide K(guanidinyl)GVLHAV at 765.47 *m/z*.

Further examples of this enhanced formation of b_n+H_2O species following guanidination of the *N*-terminal lysine residue are shown in Figures 32 and 33. The peptides KNVPLY (Figure 32A) and KLVTDLT (Figure 33A) both decompose during CID to produce minor b_n+H_2O species which correspond to loss of the *C*-terminal amino acid [64], in addition to ion series that are indicative of sequence scrambling (only those scrambled species with significant signal intensity are labelled). In each example the derivatised peptide sequence fragments to produce b_n+H_2O species that are of significant intensity and the formation of which appears to be promoted by the *N*-terminal homoarginine residue, however no products of sequence scrambling are observed.

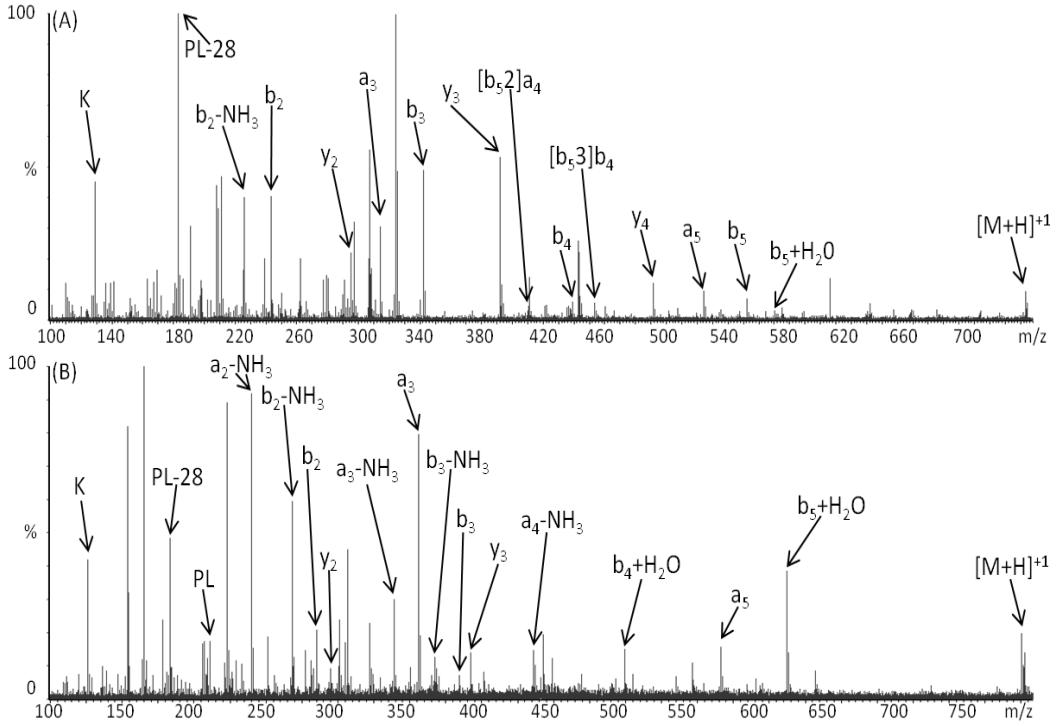

Figure 32. Comparison of QTOF MS/MS spectra resulting from (A) CID of $[M+H]^+$ KNVPLY at 733.42 *m/z* and (B) CID of $[M+H]^+$ derivatised peptide K(guanidinyl)NVPLY at 775.45 *m/z*.

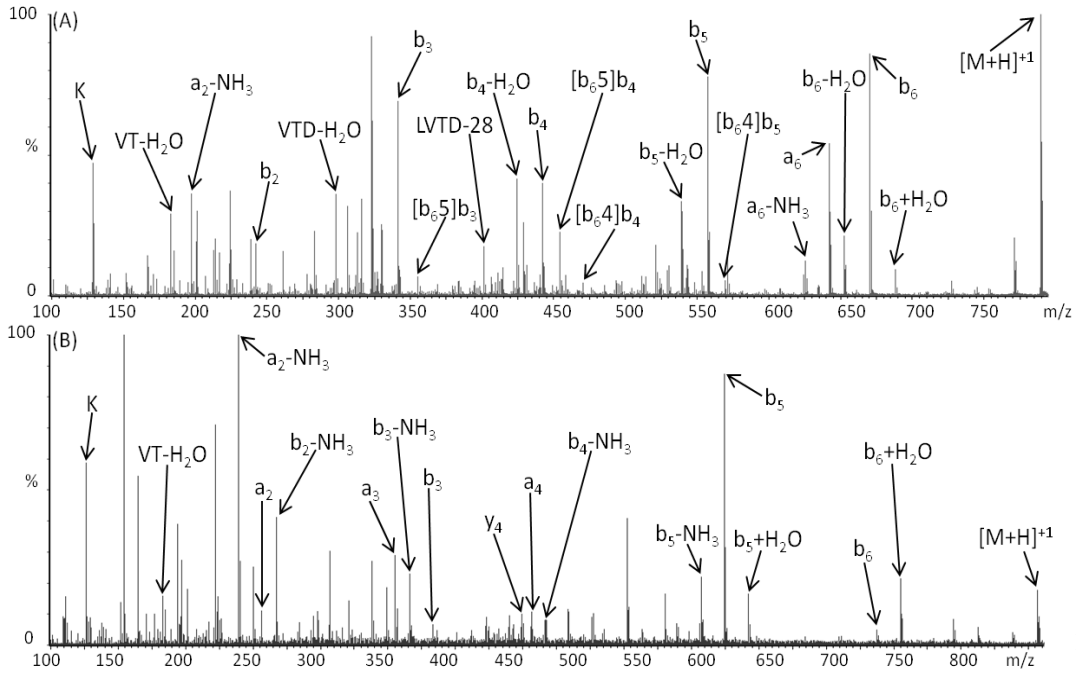

Figure 33. Comparison of QTOF MS/MS spectra resulting from (A) CID of $[M+H]^+$ KLVTDLT at 789.47 *m/z* and (B) CID of $[M+H]^+$ derivatised peptide K(guanidinyl)LVTDLT at 831.49 *m/z*.

The observation of enhanced b_n+H_2O ion formation is consistent with the findings reported by various groups [64, 168, 169], whereby loss of the *C*-terminal residue accompanied by retention of a carboxyl oxygen within the newly formed b-ion structure is promoted by the presence of a basic residue at the *N*-terminus. Derivatisation of the lysine side chain to homoarginine increases the gas phase basicity of the *N*-terminal residue hence giving a greater extent of rearrangement.

As a consequence of this enhanced b_n+H_2O formation, the apparent inhibition of peptide sequence scrambling may simply be a result of ion current being directed away from the formation of b-ion fragments. The intensity of b-ion species is in most cases vastly reduced, meaning that the precursors to sequence scrambled species are much less abundant. Alternatively, the apparent reduced propensity of sequence scrambling during CID may result due to the presence of the homoarginine residue, an observation analogous to that reported by reported by Molesworth *et al.* [150] for arginine containing sequences. The authors concluded that inhibition of sequence scrambling for such peptides may be the result of either a reduced propensity for macrocycle formation via rearrangement of the initially formed oxazolone structure

or the prevention of direct generation of a cyclic b-ion structure upon CID of the precursor ion. It was postulated that the arginine residue may sequester a proton necessary for facile macrocycle formation or alternatively promote formation of an intramolecular hydrogen bonding interaction stabilising a structure which prevents 'head to tail' cyclisation. If such factors are of consequence toward the inhibition of sequence scrambling it may be expected that a similar effect would be observed following modification of the lysine side-chain to homoarginine. The peptide K(guanidinyl)NVPLY does not produce b-ion species with significant signal intensity so it is not possible to conclude that inhibition of sequence scrambling is not simply a result of the absence of a precursor to macrocycle formation. However, the peptide KLVTDLT fragments to give b_6 and b_5 species both prior to and following derivatisation suggesting that the homoarginine residue does indeed play a role in the inhibition of sequence rearrangement.

To further test this hypothesis the peptide sequences RGVLHAV, RNVPLY and RLVTDLT were synthesised (JPT Peptide Technologies, Berlin) and subjected to equivalent analysis by CID. The resultant MS/MS spectra are shown in Figure 34 and again, in each instance, extensive formation of b_n+H_2O species is observed, however no product ions resulting from sequence scrambling are present. Thus it would appear that the presence of an arginine or homoarginine residue precludes sequence scrambling while a highly basic *N*-terminal residue promotes rearrangement at the *C*-terminus to give b_n+H_2O fragment ions.

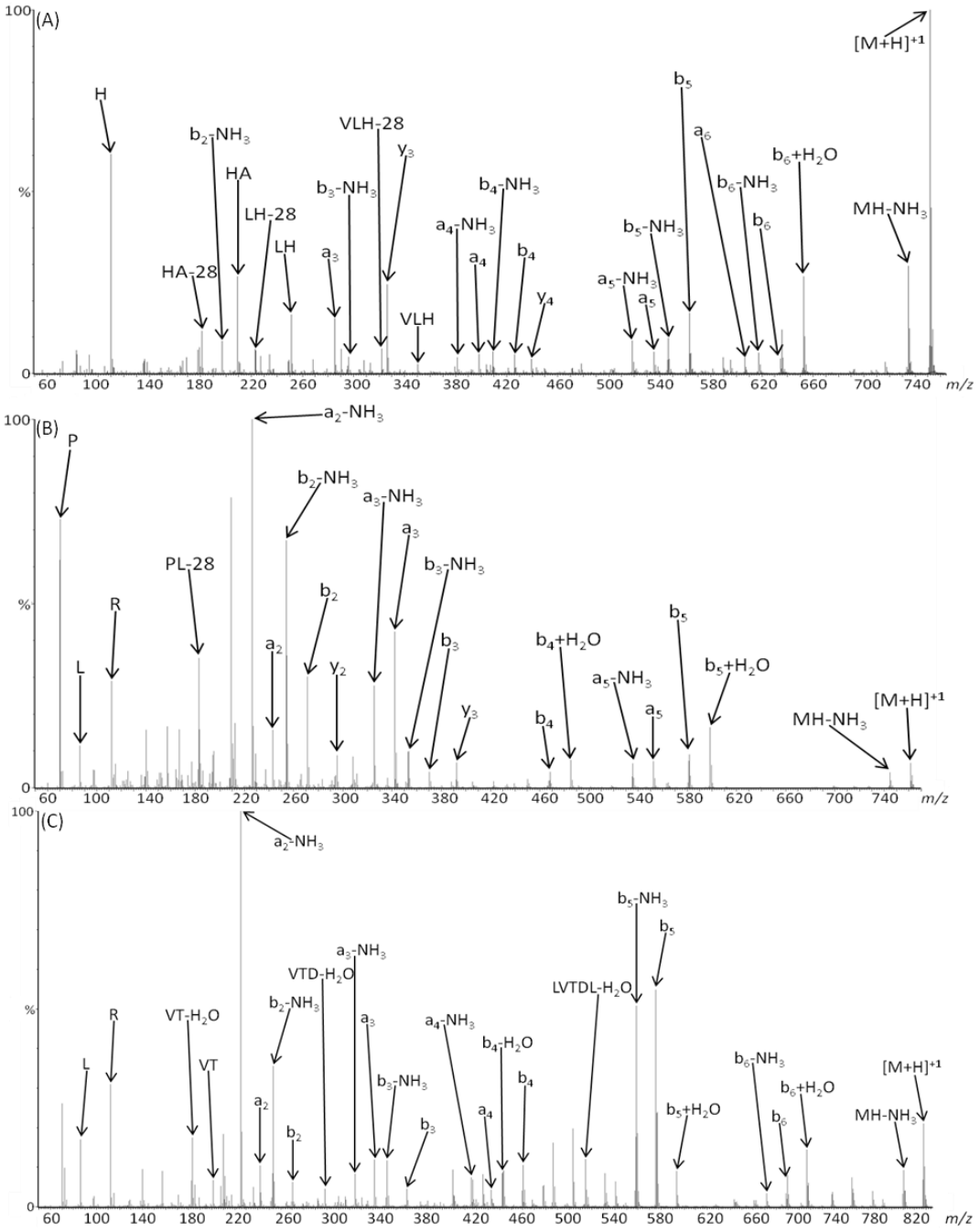

Figure 34. Comparison of QTOF MS/MS spectra resulting from (A) CID of $[M+H]^+$ RGVLHAV at 751.45 *m/z* and (B) CID of $[M+H]^+$ RNVPLY at 761.43 *m/z* and (C) CID of $[M+H]^+$ RLVTDLT at 817.47 *m/z*.

It is not clear why the presence of such residues inhibits sequence scrambling, however as previously discussed it may be due to the lack of a 'mobile' proton accommodated within the structure. Mobilisation of the proton sequestered by the arginine (or homoarginine) side chain would normally be required for efficient CID, Molesworth *et al.* instead proposed mobilisation of an amide proton from a leucine residue in the sequence YARFLG, leading to formation of the b_5^+ oxazolone ion, in accordance with the model previously published by Paizs and co-workers [170]. This proposed fragmentation pathway proceeds via an iminol tautomer to generate a b_5 ion following nucleophilic attack to yield the deprotonated oxazolone ring. Each of the sequences discussed here also incorporates a leucine residue and therefore may potentially fragment via the same pathway once the previously 'mobile' proton is sequestered by the arginine/homoarginine residue.

To help determine whether inhibition of peptide scrambling for arginine (or homoarginine) containing sequences is related to the lack of such a proton that can be mobilised upon activation the doubly protonated peptide FGERALK was studied by CID. An intense ion was observed following CID on the QTOF at 545.34 *m/z* (Figure 35A) whose origin can be attributed to ring opening of an intermediate macrocycle structure, generating $[b_6 3]b_5$. The large intensity with which this fragment ion is observed is somewhat surprising as scrambled species are typically minor components of MS/MS spectra, yet an ion of this *m/z* cannot be accounted for by any of the normal sequence derived fragment ions or their derivatives. The mechanism by which this scrambled product ion is generated will be determined by the immediate precursor to the scrambling process. The doubly charged, even electron precursor ion either fragments to give a singly charged ion pair originating from opposing termini of the peptide (a singly charged b-ion and singly charged y-ion) or a doubly charged a_6 fragment ion resulting from retention of two protons within the *N*-terminal product ion series. Thus, the spectrum suggests two possible mechanisms: firstly the scrambled species may be generated from the doubly charged a-ion observed at 323.69 *m/z*, with ring opening of the macrocycle initiated by the additional 'mobile' proton accommodated within the structure. Secondly, preferential ring opening of the macrocycle may occur *C*-terminal to glutamic acid via a charge remote cleavage as is frequently observed during CID of linear peptide ions [171].

As the $[b_6 3]b_5$ ion is observed with such a large signal intensity it would appear likely that formation is attributable to ring opening of the macrocycle intermediate occurring predominantly between the E and R residues, followed by subsequent loss of the now *C*-terminal glutamic acid residue. Additional scrambling pathways in the QTOF spectrum produced from FGERALK are not present at a detectable level except for a $[b_5 3]b_4$ product ion observed at 432.24 *m/z* corresponding to the same charge remote cleavage of a macrocyclic intermediate arising from the b_5 fragment ion. If ring opening were initiated by a mobile proton, a more heterogeneous population of sequence scrambled product ions would be expected, with ring opening occurring at various sites. However, it must be noted that a corresponding charge remote cleavage *C*-terminal to aspartic acid in the peptide K(guanidinyl)LVTDLT is not observed (Figure 33B).

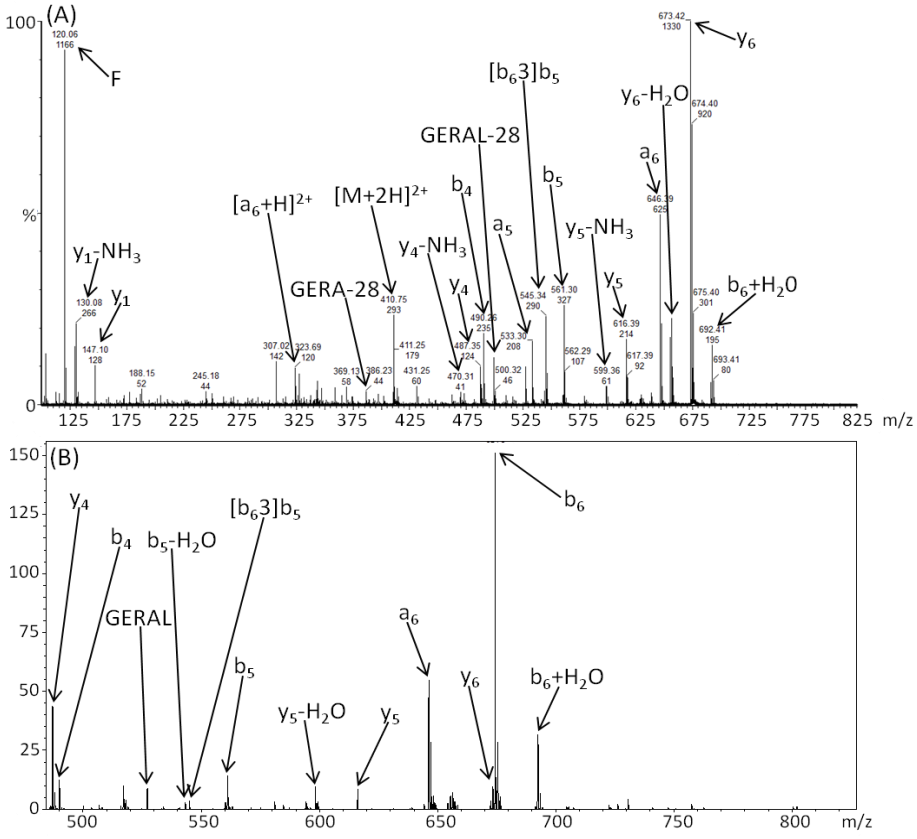

Figure 35. (A) QTOF MS/MS spectrum generated by CID of the $[M+2H]^{2+}$ peptide FGERALK. The $[b_6 3]b_5$ fragment ion observed at 545.34 *m/z* is attributable to scrambling of the original peptide sequence via ring opening of a macrocycle intermediate. (B) QIT MS/MS spectrum generated by CID of the $[M+2H]^{2+}$ peptide FGERALK.

Equivalent CID analysis of the FGERALK peptide using an amaZon QIT (Bruker, Bremen, Germany) provides evidence to support this conclusion (Figure 35B). The intensity of the $[b_6 3]b_5$ fragment in the QIT MS/MS spectrum is much reduced relative to that of the y_4 ion that is observed in each analysis and should be stable to the conditions encountered on each instrument platform. In addition, inspection of the isotope ratio of the b_6 (674.36 *m/z*) and y_6 (673.39 *m/z*) ions in each spectrum suggests that the b_6 population is depleted during analysis in the multiple collision regime of a QTOF type collision cell (Figures 36A and 36B) while the doubly protonated a_6 ion is observed with significant intensity during each analysis (data not shown). It has previously been demonstrated that *N*-terminal fragment ions are less stable toward secondary fragmentation events than their y-ion counterparts and as a consequence are often under-represented in QTOF MS/MS spectra when compared to those obtained using a QIT [68]. During MS/MS on a QTOF instrument, peptide ions are subject to multiple collisions with the neutral buffer gas causing preferential secondary decomposition of b-ion species to lower members of the ion series. In this instance it would appear that during analysis on the QTOF platform once initial decomposition has occurred, the b_6 and b_5 fragment ions undergo further activation, promoting rearrangement and secondary fragmentation leading to formation of the $[b_6 3]b_5$ and $[b_5 3]b_4$ species. This is highlighted by analysis of the y_6 ion isotope ratio observed from analysis on the QTOF instrument at high and low collision energy. At lower collision energy (Figure 36A, collision energy set at 19.5 V) there is a distortion of the distribution suggesting a contribution from the b_6 ion at 674.36 *m/z*. Conversely, at higher collision energy (Figure 36B, collision energy set at 26 V) there is no distortion of the isotope distribution, reflecting depletion of the b_6 ion population through secondary fragmentation, potentially forming the $[b_6 3]b_5$ scrambled fragment ion. This phenomenon is of reduced consequence during QIT analysis as products of initial fragmentation are no longer resonant with the excitation voltage and are therefore not subject to multiple ion-neutral collisions. Higher members of the b-ion series are therefore often observed and this is demonstrated in Figure 36C where, in contrast to the QTOF data, the b_6 ion generated from FGERALK is observed with greater signal intensity relative to the y_6 ion. Generation of the scrambled $[b_6 3]b_5$ species on the QIT instrument may be expected to be enhanced by the longer trapping times associated with such analysis,

however, it appears that as further activation of the b_6/b_5 fragment ions is avoided, so too is rearrangement of the peptide sequence.

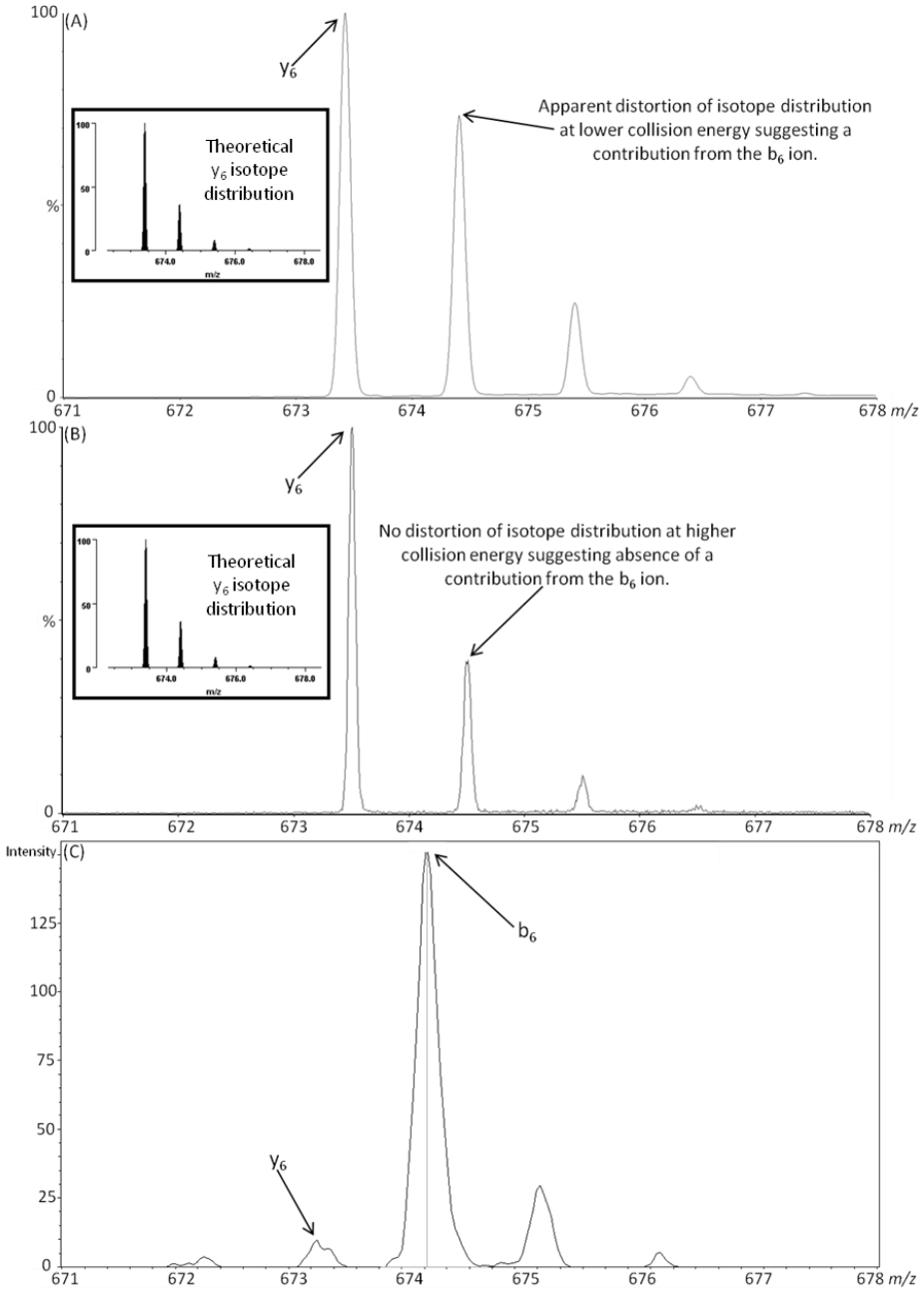

Figure 36. CID of $[M+2H]^{2+}$ FGERALK with the collision energy set at either (A) 19.5 V or (B) 26 V using a QTOF instrument results in formation of a y_6 ion at 673.39 *m/z*. Distortion of the isotope distribution in (A) suggests a contribution from the b_6 ion at 674.36 *m/z*. Shown in the inset is the theoretical isotope distribution of the y_6 ion which was modelled using the MS-Isotope program within Protein Prospector. (C) Equivalent CID analysis of $[M+2H]^{2+}$ FGERALK using a QIT instrument results in formation of an intense b_6 ion at 674.36 *m/z*.

6.3. Conclusions

MS/MS analysis has been used to demonstrate that increasing the gas phase basicity of amino terminal basic groups through derivatisation of an *N*-terminal lysine side-chain to give homoarginine results in enhanced formation of b_n+H_2O species via a *C*-terminal rearrangement. A consequent reduction in the complexity of the products of CID is also observed as scrambling of the initial peptide sequence is avoided. This apparent inhibition of sequence scrambling for singly protonated peptides may simply be the result of ion current being directed away from standard b-ion fragments which are the precursors to scrambling of the initial sequence, hence making detection of resultant scrambled species problematic. However, further analysis of a doubly protonated arginine containing peptide appears to exhibit sequence scrambling predominantly via a charge remote opening of the macrocycle intermediate. This finding suggests that the presence of a guanidine moiety, located within the arginine/homoarginine side chain will in most cases preclude the scrambling fragmentation pathway for singly protonated species, in accordance with previously published studies [150, 151]. The data reported here suggests that formation of the macrocycle intermediate remains facile. However, the reduced mobility of the proton, now sequestered by the guanidino-group, prevents ring opening of the intermediate in the absence of a charge remote cleavage.

Future directions for the work discussed in this chapter would be to conduct T-Wave mobility analysis of b-ion species generated from arginine containing sequences. The data may help to ascertain if the presence of a guanidine side-chain precludes formation of a macrocyclic structure or whether the lack of evidence (presence of sequence scrambled product ions) for macrocycle formation is simply a result of the basic residue discouraging (by sequestering the mobile proton) ring opening of the macrocycle. If a heterogeneous b-ion population was observed with respect to drift time this would suggest that a macrocyclic structure is formed, however in the absence of a mobile proton ring opening does not typically occur.

Chapter 7

Ion Mobility Analysis of Peptide and Peptide Fragment Ion Structure

7.1. Introduction

Gas phase peptide and peptide fragment ion collision cross section is expected to be determined by factors including charge state, mass and amino acid composition. For example, proline containing peptides have been shown to induce multiple resolved features during ion mobility analyses; an observation attributed to the presence of both cis and trans isomers of such species [172]. The secondary structure of protonated peptides has been shown to have a direct influence upon observed fragmentation behaviour during both CID [171] and ETD [144]. Probability of dissociation is determined by the energetic and kinetic accessibility of a reactive configuration and a transition between isomeric forms of a given ion may be required for a fragmentation pathway to be followed. The position of basic amino acids such as arginine, lysine and histidine could play a key role in the secondary structure adopted by a particular ion. The high gas phase basicity of these groups means that they typically sequester any available proton accommodated within a peptide structure and consequently have the capability to form gas phase salt-bridge interactions [173]. Solvation of the charged site through intra-molecular interactions is also common and this is known to influence both the secondary structure and consequent fragmentation behaviour of the peptide. Knowledge of factors which influence the gas phase conformation of a peptide and/or peptide fragment ions can help in the development of more sophisticated bioinformatics based peptide identification tools. For example, if universal separation of peptide fragment ions was possible by ion mobility (i.e. to distinguish between *N*- and *C*-terminal fragments) this would drastically improve the specificity of the searches performed by spectral interpretation algorithms. The work presented in this chapter investigates the effect of proton localisation and intramolecular interactions upon gas phase conformation of peptides generated by a range of digestion regimes and peptide fragment ions generated by CID and ETD.

7.2. Results and Discussion

To investigate the influence of proton localisation on the gas phase conformation of singly charged peptides, a selection of 'standard' proteins were proteolysed overnight using trypsin. An aliquot of the peptide mixture was then subjected to secondary enzymatic treatment using the exoprotease CBPB. Resultant peptide sequences enabled any change in ion mobility behaviour following removal of the *C*-terminal basic residue to be assessed by analysis using a Synapt HDMS instrument (Waters, Manchester, UK). Shown in Figure 37 is an example of the data obtained from T-Wave ion mobility analysis of a tryptic digest of the protein α-lactalbumin, showing separation of peptide ions of different charge states. Removal of the *C*-terminal basic residue following CBPB treatment, results in a bias towards the population of lower charge states. This is reflected by the observed lower intensity of triply charged species following CBPB treatment and a corresponding increase in the intensity of both the doubly and singly charged species (Figure 37). A proportion of the tryptic peptides detected as singly charged species were not observed after CBPB digestion; this can be attributed to the reduction in basicity associated with removal of the *C*-terminal arginine/lysine residue.

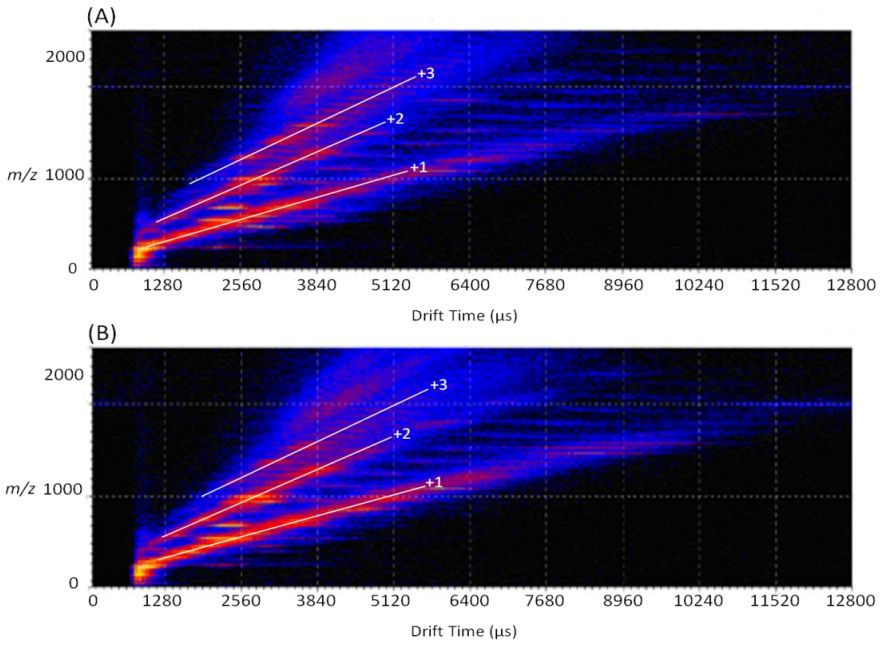

Figure 37. T-Wave Ion Mobility analysis of (A) peptides generated by trypsin digestion of α-lactalbumin (B) peptides generated by sequential trypsin and CBPB digestion of α-Lactalbumin.

The drift time of those peptides observed as singly protonated species both prior to and following CBPB treatment was extracted and then plotted against *m/z* (Figure 38). Surprisingly, there does not appear to be a significant effect on the gas phase conformation of singly protonated tryptic peptide ions following change of proton localisation away from the *C*-terminus. In the main, for both the 20 arginine terminating sequences (Figure 38A) and 45 lysine terminating sequences (Figure 38B), removal of the highly basic residue appears to reduce the observed drift time in accordance with the removal of a single amino acid residue. If the associated relocation of the site of protonation following CBPB treatment impacts upon peptide conformation a more significant change in the *m/z* versus drift time correlation would be expected. The data suggest that the global peptide conformation is maintained while the cross section is reduced through removal of a single amino acid residue.

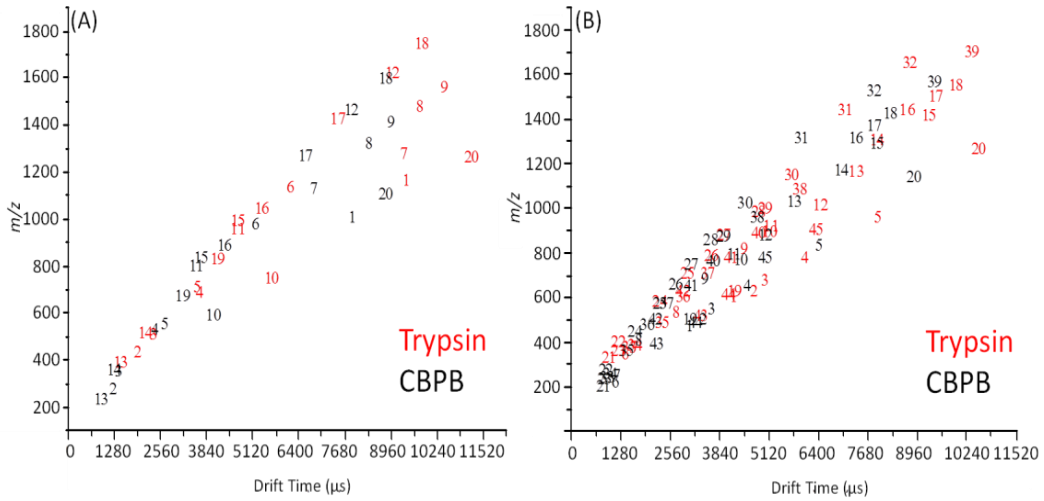

Figure 38. Comparison of drift time versus *m/z* for (A) 20 arginine terminating peptides and (B) 45 lysine terminating peptides both prior to and following removal of the *C*-terminal basic residue.

To further this study, T-Wave analysis of a range of singly and doubly protonated peptides generated by overnight Lys-C and Lys-N proteolysis of a selection of 'standard' proteins (BSA, enolase from *S. cerevisiae*, horse myoglobin) was conducted. The sequences chosen differ solely by the position of a lysine residue,

located at either the *C*-terminus following Lys-C digestion or at the *N*-terminus following Lys-N digestion. Consequently, following calibration of the T-Wave instrument using the QCAL-IM mobility standard discussed in Chapter 3, collision cross section values for each of the permuted isomers were determined (Tables 9 and 10) and any effect resulting from proton localisation at opposing peptide termini assessed.

Table 9. Collision cross section of singly protonated peptides produced by Lys-C and Lys-N proteolysis of a range of 'standard' proteins. Cross section values were determined following calibration of the T-Wave instrument using the QCAL-IM ion mobility standard and are an average of 3 replicate measurements.

m/z	Sequence	Average Ω ($Å^2$)	*m/z*	Sequence	Average Ω ($Å^2$)	Ω Difference ($Å^2$)
789.47	KLVTDLT	204.6	789.47	LVTDLTK	205.7	1.1
818.43	KATEEQL	201.6	818.43	ATEEQLK	207.1	5.5
886.42	KDDSPDLP	209.4	886.42	DDSPDLPK	213.7	4.3
906.47	KIETMRE	216.5	906.47	IETMREK	214.4	-2.1
1163.63	KLVNELTEFA	258.3	1163.63	LVNELTEFAK	265.9	7.6
790.43	KASEDLK	197.4	790.43	ASEDLKK	198.9	1.5
783.44	KHLADLS	201.7	783.44	HLADLSK	202.4	0.7
814.50	KAADALLL	212.3	814.50	AADALLLK	218.7	6.4
901.55	KRIATAIE	215.0	901.55	RIATAIEK	218.3	3.3

For singly charged ions the position of the lysine residue appears to have minimal effect upon the peptide conformation (Table 9) although there is a trend toward a small increase in collision cross section for those peptides with the basic amino acid at the *C*-terminus, produced by Lys-C proteolysis. The single peptide (IETMREK) which has a larger collision cross section when the lysine residue is located at the *N*-terminus incorporates an internal arginine residue. It is therefore likely that the proton accommodated along the peptide backbone is located on the side chain of this residue due to the associated high gas phase basicity. Regardless of the location of the terminal lysine amino acid, the site of protonation is likely to remain unchanged due to this centrally positioned arginine amino acid.

Any increase in gas phase collision cross section arising from the site of protonation is likely to result due to coulombic repulsion between charges accommodated along the peptide backbone. In the absence of zwitterion formation this effect is not of consequence for singly charged ions and it may be expected that a greater difference in cross sectional area would be observed for multiply charged ions. The data shown in Table 10 supports this theory with a number of doubly protonated peptide ions adopting conformations with significantly different cross sectional areas. For example, the calculated cross section of the peptide LGANAILGVSLAASRAAAAEK is 50.9 $Å^2$ larger (7.41 % increase) when the lysine residue is located at the *C*-terminus. This is also true for the peptides TC*VADESHAGC*EK (17.5 $Å^2$, 15.99 % increase) and VPQVSTPTLVEVSRSLGK (44.4 $Å^2$, 8.0 % increase). In each instance there is an internal arginine or histidine residue located more closely to the *C*-terminus than to the *N*-terminus. Considering the likely sites of protonation within these peptides being the internal arginine/histidine residue and terminal lysine residue a greater effect from coulombic repulsion is expected for the Lys-C analogue due to the proximity of these charged amino acids. Such repulsion may inhibit a smaller conformation that is likely to be invoked due to salt-bridge formation for the Lys-N analogue between the *N*-terminal basic amino acid and the *C*-terminus. Three sequences that do not follow this trend are QEPERNEC*FLSHK (0.8 $Å^2$, 0.26 % decrease), SLHTLFGDELC*K (0.9 $Å^2$, 0.31 % decrease) and HGTVVLTALGGILK

(2.1 $Å^2$, 0.71 % decrease). Again, consideration of the likely sites of protonation within the Lys-N analogue of these peptides can be used to explain why this small decrease in cross section is observed. The peptide KQEPERNEC*FLSH has three basic residues, including a *C*-terminal histidine and in order to minimise repulsion between charges this site may in fact be protonated rather than the *N*-terminal lysine. The potential for heterogeneity in the site of protonation can be postulated as the reason for the minimal difference in cross sectional area between the Lys-C and Lys-N analogues. When considering the Lys-N sequences KSLHTLFGDELC* and KHGTVVLTALGGIL both are likely to be protonated at the K and H residues, which for each sequence are in close proximity. Strong repulsion between the positive charges is likely and an expansion in the gas phase conformation will result; hence giving a slightly increased cross section for the Lys-N analogue. This is supported by the greater percentage cross section difference being observed for the KHGTVVLTALGGIL sequence, as the adjacent basic sites will result in more significant columbic repulsion.

Table 10. Collision cross section of doubly protonated peptides produced by Lys-C and Lys-N proteolysis of a range of 'standard' proteins. Cross section values were determined following calibration of the T-Wave instrument using the QCAL-IM ion mobility standard and are an average of 3 replicate measurements. C* represents a carbamidomethylated cysteine residue.

m/z	Sequence	Average Ω (Å^2)	*m/z*	Sequence	Average Ω (Å^2)	Ω Difference (Å^2)
410.73	KFGERAL	203.7	410.73	FGERALK	209.0	5.3
494.27	KSEIAHRF	217.5	494.27	SEIAHRFK	224.9	7.4
573.32	KAWSVARLSQ	244.3	573.32	AWSVARLSQK	251.5	7.2
710.35	KSLHTLFGDELC*	289.8	710.35	SLHTLFGDELC*K	288.9	-0.9
732.29	KTC*VADESHAGC*E	279.9	732.29	TC*VADESHAGC*EK	297.4	17.5
751.81	KEYEATLEECC*C*A	284.9	751.81	EYEATLEEC*C*AK	285.2	0.3
777.83	KDDPHAC*YSTVFD	296.6	777.83	DDPHAC*YSTVFDK	306.4	9.8
837.39	KQEPERNEC*FLSH	309.4	837.39	QEPERNEC*FLSHK	308.6	-0.8
949.04	KVPQVSTPTLVEVSRSLG	355.2	949.04	VPQVSTPTLVEVSRSLGK	399.6	44.4
680.88	KALELFRNDIAA	286.7	680.88	ALELFRNDIAAK	294.8	8.1
689.92	KHGTVVLTALGGIL	294.2	689.92	HGTVVLTALGGILK	292.1	-2.1
751.83	KHPGDFGADAQGAMT	291.6	751.83	HPGDFGADAQGAMTK	301.0	9.4
628.85	KTGAPARSERLA	263.1	628.85	TGAPARSERLAK	265.9	2.8
878.47	KTAGIQIVADDLTVTNP	339.9	878.47	TAGIQIVADDLTVTNPK	345.8	5.9
974.51	KTFAEALRIGSEVYHNL	381.4	974.51	TFAEALRIGSEVYHNLK	400.9	19.5
977.55	KLGANAILGVSLAASRAAAAE	377.2	977.55	LGANAILGVSLAASRAAAAEK	428.1	50.9

The structure of a typical b-ion is thought to be partially cyclic in nature [78], whereas y-ions have a linear structure [76, 77]; thus it is conceivable that separation of these two ion types, in the presence of sufficient ion mobility resolving power, may be achievable on a basis of secondary structure. However, there may be a contribution to this secondary structure resulting from intramolecular interactions such as charge solvation or salt-bridge formation. Differential proton localisation at either terminus could therefore potentially have an effect on the conformation (in addition to the nature) of fragment ions produced by CID (as discussed in Chapter 4). The geometry of the Synapt HDMS instrument means that collisional activation of analyte ions can be performed both prior to and following mobility separation. Peptides generated by Lys-C and Lys-N proteolysis were therefore subjected to CID prior to mobility separation, enabling the effect of switching the highly basic lysine residue from the carboxy- to the amino-terminus (and hence the product ion series in which it is incorporated) upon the conformation of product ions generated by CID to be assessed. The resultant data is shown in Figures 39 and 40 and was generated by tuning the applied collision energy across the Trap collision cell (see Figure 14 in Chapter 1) to give efficient peptide fragmentation resulting in generation of diagnostic b- and y-ions. The drift time of each observed fragment ion was then extracted and plotted against *m/z*.

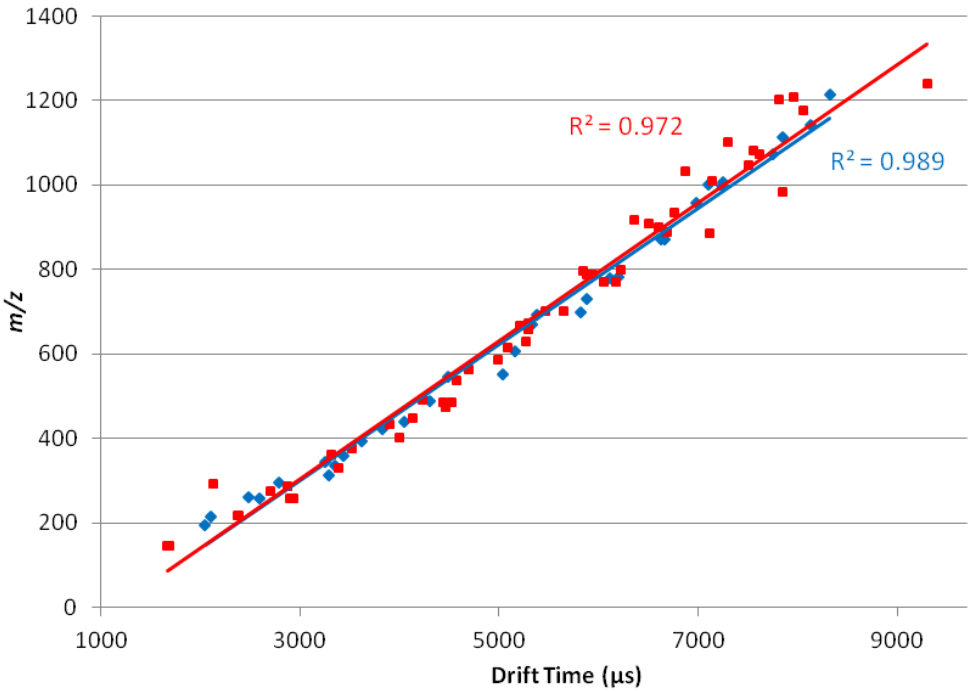

Figure 39. T-Wave ion mobility separation of b-ions (blue diamonds) and y-ions (red squares) generated by CID of a selection of peptides produced by Lys-C proteolysis.

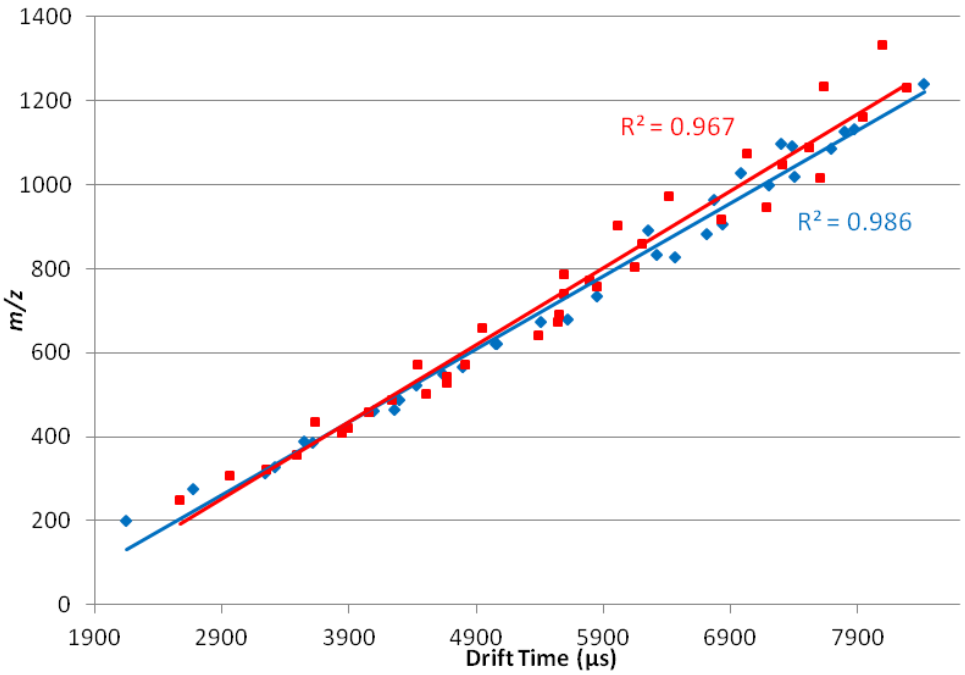

Figure 40. T-Wave ion mobility separation of b-ions (blue diamonds) and y-ions (red squares) generated by CID of a selection of peptides produced by Lys-N proteolysis.

The drift time vs *m/z* correlation for b- and y-ions generated by CID of both Lys-C and Lys-N peptides show that universal separation of the individual product ion series is not possible, at least not with the resolving power achievable using the first generation Synapt HDMS instrument. Neither the differences in primary structure between b- and y-ions nor any contributions toward secondary structure from proton localisation at either terminus result in sufficient conformational change to enable distinction between the product ion series by T-Wave ion mobility separation. A notable observation from the analysis of fragment ions derived from CID of Lys-N precursor ions is that there appears to be an increasingly broad distribution in the recorded drift time of y-ion fragments above ~800 *m/z*. The *C*-terminal y-ion series is typically considered to have a linear structure corresponding to a truncated peptide or protonated amino acid. However, this is based on y-ion species observed in proteomic studies which typically involve analysis of peptides produced by tryptic or more recently Lys-C proteolysis of the protein(s) of interest. In each instance the *C*-terminal amino acid will be a highly basic arginine or lysine residue which, being the site of highest gas phase basicity will sequester an available proton. The presence of this residue and consequent reduced proton mobility within the fragment ion structure is acknowledged as being responsible for the increased stability and homogeneous structure of y-ions when compared to their b-ion counterparts. In the absence of a missed cleavage site or *C*-terminal arginine or histidine residue, peptides produced by Lys-N digestion do not possess a *C*-terminal basic site. The single proton within the y-ion is therefore likely to be mobile under CID conditions resulting in both the potential for enhanced secondary fragmentation and formation of a heterogeneous population of y-ion conformations; determined primarily by sequence, resultant proton location and consequent intramolecular interactions.

ETD is a less well established dissociation technique when compared to CID, having been developed by Syka *et al.* in 2004 [8]. Studies toward understanding the structure of product ions generated by ETD therefore remain in their infancy. It is likely that the primary sequence of c/z-ions will have a direct influence upon the secondary structure adopted by each fragment ion. As such, as with the products of CID discussed earlier, the potential for formation of intramolecular interactions, columbic repulsion between charges and kinks induced by the presence of a proline

residue may enable separation of ETD generated fragment ions by T-Wave ion mobility.

To investigate these hypotheses tryptic and Lys-C digestions of a selection of proteins were prepared and resultant peptide mixtures analysed by direct infusion ESI using a Synapt G2 instrument (Waters, Manchester, UK). Peptides selected were known to be amenable to ETD as they are typically observed in a multiply charged state. All samples were prepared at a concentration of 1 pmol/μl using water/acetonitrile/m-nitrobenzylalcohol (1/1/0.1 v/v/v) containing 0.1 % formic acid. The inclusion of m-nitrobenzylalcohol was as a supercharging reagent to aid the formation of higher charge states. The precursor of interest was *m/z* selected using the quadrupole and ETD analysis performed following co-trapping of the precursor ion and nitrobenzene ETD reagent in the trap cell, inducing fragmentation, prior to mobility separation. Figure 41 shows clear mobility separation of the c/z-ions generated from the triply charged peptide RHPEYAVSVLLR; with the c-ion series having a higher mobility and therefore more compact gas phase structure. When considering the 'bulky' side chains present in the residues located near the *N*-terminus this is at first surprising. However, the *N*-terminal arginine residue resulting from a miscleavage site can form an internal salt-bridging interaction with the nearby glutamic acid, thus causing a reduction in collision cross section. This conclusion is supported by the c_3-ion having a mobility closer to that of the z-ion trend than the remainder of the ion series. The c_3-ion does not incorporate the glutamic acid residue meaning that the formation of an internal salt-bridging interaction is far less feasible.

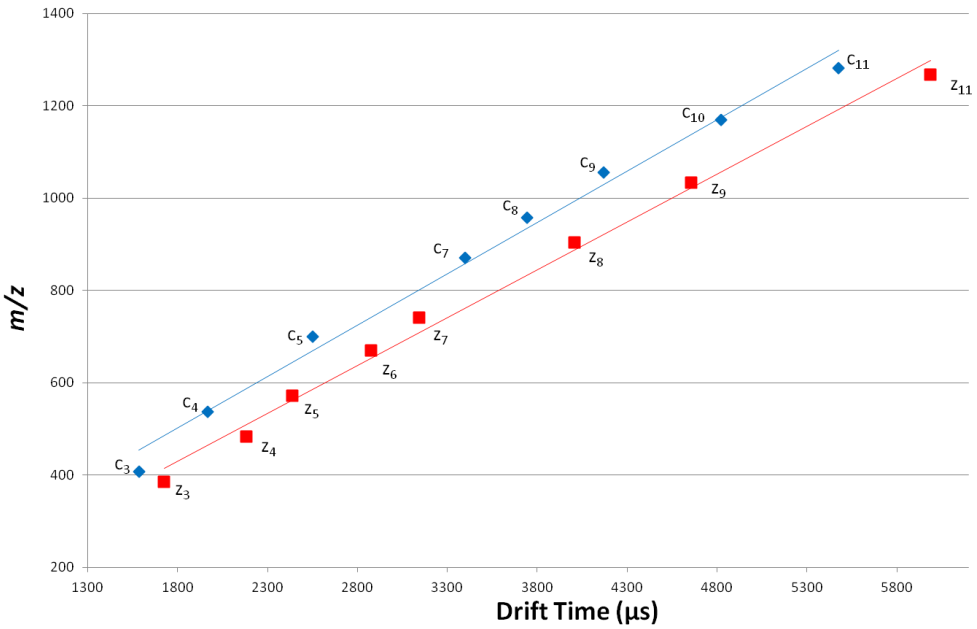

Figure 41. T-Wave ion mobility separation of c/z-ions generated from ETD of the $[M+3H]^{3+}$ peptide RHPEYAVSVLLR.

Mobility separation of the c/z-ions generated from ETD of triply charged Substance P (RPKPQQFFGLM) also showed a similar trend (Figure 42); the peptide incorporates two proline residues which are known to induce a kink in peptide structure [174]. It is possible that the separation of c/z-ions generated from both Substance P and RHPEYAVSVLLR are assisted by the presence of this proline motif in the peptide structure.

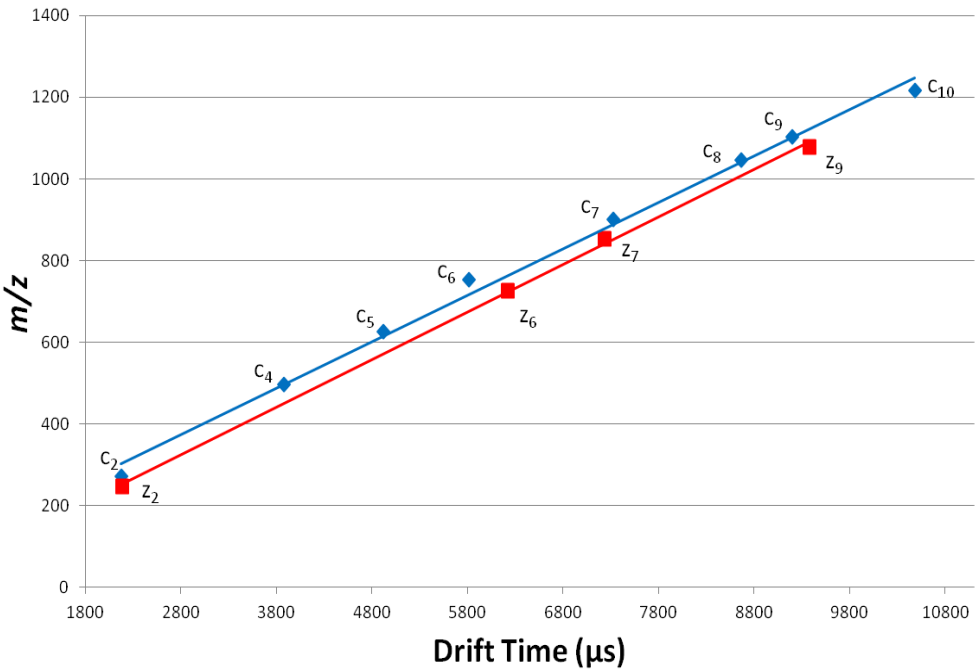

Figure 42. T-Wave ion mobility separation of c/z-ions generated from ETD of $[M+3H]^{3+}$ Substance P (RPKPQQFFGLM).

T-Wave ion mobility separation of c/z-ions however appears to be more difficult in the absence of internal interactions as the achievable mobility resolution is insufficient to distinguish between the ion series according to primary structure. The triply charged peptide SLHTLFGDELC*K does not contain a proline residue and the capability to form an internal salt-bridge is much reduced within the c-ion series. There is no *N*-terminal arginine residue present and a bridging interaction between the internal H and D/E residues is far less energetically favourable due to their mid-chain location. Figure 43 shows the mobility separation achieved for this peptide and in contrast to the peptides discussed thus far the smaller members of the z-ion series appear to exhibit a higher mobility relative to the c-ion series.

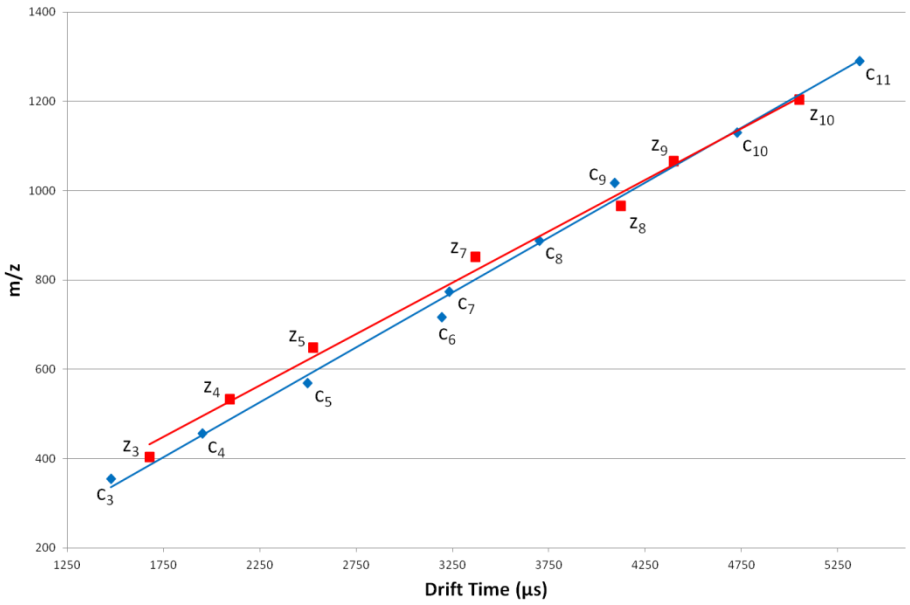

Figure 43. T-Wave ion mobility separation of c/z-ions generated from ETD of $[M+3H]^{3+}$ peptide SLHTLFGDELC*K, where C* represents carbamidomethylated cysteine.

Again, this observation can be rationalised by considering the potential for salt-bridge formation, on this occasion between the *C*-terminal lysine residue and internal glutamic/aspartic acid residues located in close proximity. The z_3-ion does not incorporate an acidic residue meaning that salt-bridge formation is far less feasible and consequently the mobility of this ion is similar to that of the c-ion series. There is a clear shift to a higher mobility for the z_4-ion once a single acidic residue is incorporated within the structure; suggesting that once a salt-bridging interaction can be readily formed a more compact structure is adopted. This mobility trend is then continued for the z_5-ion incorporating two acidic residues.

Fragment ions generated by ETD do not appear to be readily amenable to universal separation by ion mobility. This is highlighted by the data shown in Figure 44 whereby the combined mobility data generated from a selection of peptides (LFTFHADIC*TLPDTEK, SLHTLFGDELC*K, RHPEYAVSVLLR, KVPQVSTPTLVEVSR and RPKPQQFFGLM) does not exhibit substantial distinction between the product ion series. The data shows a slight trend toward z-ion species adopting a more compact structure, particularly at higher *m/z*. However, this observation may simply be a direct result of those sequences that incorporate a basic

residue at the *C*-terminus, meaning that as previously discussed, salt-bridging interactions can form with acidic residues in close proximity.

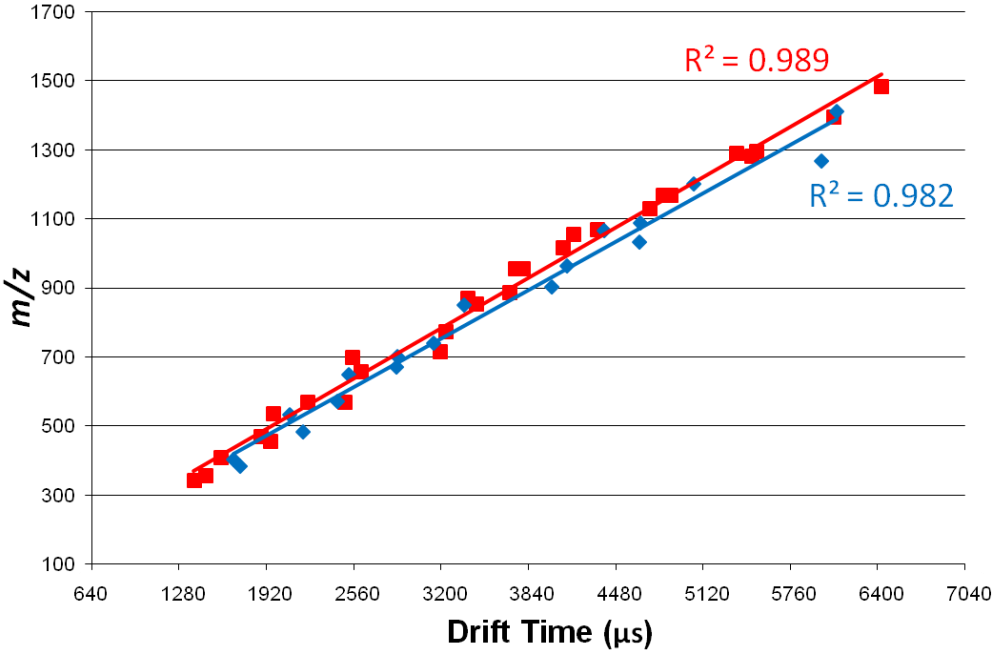

Figure 44. T-Wave ion mobility separation of c-ions (blue diamonds) and z-ions (red squares) generated by ETD of a selection of peptides produced by trypsin/Lys-C proteolysis.

7.3. Conclusions

At present bioinformatic based peptide identification from complex sample mixtures can impact significantly upon an analytical workflow due to the required time for data processing. Added analytical dimensionality, for example through use of the orthogonal technique ion mobility, has potential to reduce the complexity of data produced and hence the time taken for such data processing. The studies undertaken in this chapter aimed to investigate whether intramolecular interactions present within peptide and peptide fragment ion structures impact significantly upon the mobility of these ions during T-Wave analysis. The site of protonation in singly charged tryptic peptides does not appear to significantly affect the conformation adopted by the structure.

Proteases such as Lys-C and Lys-N are becoming more frequently utilised as alternatives to trypsin, which has traditionally been used in proteomic studies. The advent of ETD which is a complimentary technique to CID is especially useful for the study of peptides when preservation of labile PTMs is required. ETD exhibits excellent fragmentation efficiency for multiply charged peptides/proteins and consequently sequences with internal arginine residues are ideal for study using the technique. If it were possible to effectively distinguish between peptide/peptide fragment ion structures produced by these enzymes on a basis of their gas phase conformation this could be used to enhance bioinformatic searches. The data discussed in this chapter shows there is only a small change in the calculated collision cross section of singly charged Lys-C/Lys-N peptide sequences differing solely by the termini at which the lysine residue is located. However, much greater differences in collision cross section were observed for doubly charged Lys-C/Lys-N sequences; suggesting that columbic repulsion between positive charges can have a cogent effect upon the adopted conformation.

Universal mobility separation of b/y-ions or c/z-ions generated from fragmentation of a complex mixture does not appear to be possible with the current achievable mobility resolution; combined data from a selection of peptides does not show significant mobility difference between ion series. Factors including internal interactions, proline residues and the position of ‘bulky’ amino acids all influence the mobility separation and, when combined, result in a broad distribution of drift times for a particular ion series and consequent loss of separation. Distinction between fragment ion series is however possible for individual peptides where such contributions toward the adopted conformation may result in separation by T-Wave ion mobility.

Chapter 8

References

1. Tyers, M. and M. Mann, *From genomics to proteomics.* Nature, 2003. **422**(6928): p. 193-197.
2. Doherty, M.K., et al., *Turnover of the Human Proteome: Determination of Protein Intracellular Stability by Dynamic SILAC.* Journal of Proteome Research, 2009. **8**(1): p. 104-112.
3. Angel, T.E., et al., *Mass spectrometry-based proteomics: existing capabilities and future directions.* Chemical Society Reviews, 2012. **41**(10): p. 3912-3928.
4. Steen, H. and M. Mann, *The ABC's (and XYZ's) of peptide sequencing.* Nature Reviews Molecular Cell Biology, 2004. **5**(9): p. 699-711.
5. Wisniewski, E.S., D.K. Rees, and E.W. Chege, *Proteolytic-Based Method for the Identification of Human Growth Hormone.* Journal of Forensic Sciences, 2009. **54**(1): p. 122-127.
6. Boersema, P.J., S. Mohammed, and A.J.R. Heck, *Phosphopeptide fragmentation and analysis by mass spectrometry.* Journal of Mass Spectrometry, 2009. **44**(6): p. 861-878.
7. Zubarev, R.A., N.L. Kelleher, and F.W. McLafferty, *Electron capture dissociation of multiply charged protein cations. A nonergodic process.* Journal of the American Chemical Society, 1998. **120**(13): p. 3265-3266.
8. Syka, J.E.P., et al., *Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(26): p. 9528-9533.
9. Breuker, K., et al., *Top-down identification and characterization of biomolecules by mass spectrometry.* Journal of the American Society for Mass Spectrometry, 2008. **19**(8): p. 1045-1053.
10. Han, X.M., et al., *Extending top-down mass spectrometry to proteins with masses greater than 200 kilodaltons.* Science, 2006. **314**(5796): p. 109-112.
11. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-3567.

12. Eng, J.K., A.L. Mccormack, and J.R. Yates, *An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database.* Journal of the American Society for Mass Spectrometry, 1994. **5**(11): p. 976-989.
13. Craig, R. and R.C. Beavis, *TANDEM: matching proteins with tandem mass spectra.* Bioinformatics, 2004. **20**(9): p. 1466-1467.
14. Harrison, A.G., *Peptide Sequence Scrambling Through Cyclization of b(5) Ions.* Journal of the American Society for Mass Spectrometry, 2008. **19**(12): p. 1776-1780.
15. Bleiholder, C., et al., *Sequence-Scrambling Fragmentation Pathways of Protonated Peptides.* Journal of the American Chemical Society, 2008. **130**(52): p. 17774-17789.
16. Cox, G.B., *Preparative Reversed-Phase Chromatography of Proteins.* Chromatography in Biotechnology, 1993. **529**: p. 165-182.
17. Xu, P., D.M. Duong, and J.M. Peng, *Systematical Optimization of Reverse-Phase Chromatography for Shotgun Proteomics.* Journal of Proteome Research, 2009. **8**(8): p. 3944-3950.
18. Washburn, M.P., D. Wolters, and J.R. Yates, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nature Biotechnology, 2001. **19**(3): p. 242-247.
19. LoBrutto, R., et al., *Effect of the eluent pH and acidic modifiers in high-performance liquid chromatography retention of basic analytes.* Journal of Chromatography A, 2001. **913**(1-2): p. 173-187.
20. King, R., et al., *Mechanistic investigation of ionization suppression in electrospray ionization.* Journal of the American Society for Mass Spectrometry, 2000. **11**(11): p. 942-950.
21. Gabelica, V. and E. De Pauw, *Internal energy and fragmentation of ions produced in electrospray sources.* Mass Spectrometry Reviews, 2005. **24**(4): p. 566-587.
22. Dole, M., L.L. Mack, and R.L. Hines, *Molecular Beams of Macroions.* Journal of Chemical Physics, 1968. **49**(5): p. 2240-&.
23. Fenn, J.B., *Electrospray wings for molecular elephants (Nobel lecture).* Angewandte Chemie-International Edition, 2003. **42**(33): p. 3871-3894.

24. Fenn, J.B., et al., *Electrospray Ionization for Mass-Spectrometry of Large Biomolecules.* Science, 1989. **246**(4926): p. 64-71.
25. Fenn, J.B., *Ion Formation from Charged Droplets - Roles of Geometry, Energy, and Time.* Journal of the American Society for Mass Spectrometry, 1993. **4**(7): p. 524-535.
26. Gaskell, S.J., *Electrospray: Principles and practice.* Journal of Mass Spectrometry, 1997. **32**(7): p. 677-688.
27. Vanberkel, G.J., S.A. Mcluckey, and G.L. Glish, *Electrochemical Origin of Radical Cations Observed in Electrospray Ionization Mass-Spectra.* Analytical Chemistry, 1992. **64**(14): p. 1586-1593.
28. Xu, X.M., S.P. Nolan, and R.B. Cole, *Electrochemical Oxidation and Nucleophilic-Addition Reactions of Metallocenes in Electrospray Mass-Spectrometry.* Analytical Chemistry, 1994. **66**(1): p. 119-125.
29. Wilm, M.S. and M. Mann, *Electrospray and Taylor-Cone Theory, Doles Beam of Macromolecules at Last.* International Journal of Mass Spectrometry, 1994. **136**(2-3): p. 167-180.
30. Mack, L.L., et al., *Molecular Beams of Macroions .2.* Journal of Chemical Physics, 1970. **52**(10): p. 4977-&.
31. Iribarne, J.V. and B.A. Thomson, *Evaporation of Small Ions from Charged Droplets.* Journal of Chemical Physics, 1976. **64**(6): p. 2287-2294.
32. Thomson, B.A. and J.V. Iribarne, *Field-Induced Ion Evaporation from Liquid Surfaces at Atmospheric-Pressure.* Journal of Chemical Physics, 1979. **71**(11): p. 4451-4463.
33. Kebarle, P., *A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry.* Journal of Mass Spectrometry, 2000. **35**(7): p. 804-817.
34. Cole, R.B., *Some tenets pertaining to electrospray ionization mass spectrometry.* Journal of Mass Spectrometry, 2000. **35**(7): p. 763-772.
35. Karas, M. and F. Hillenkamp, *Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10000 Daltons.* Analytical Chemistry, 1988. **60**(20): p. 2299-2301.
36. Ehring, H., M. Karas, and F. Hillenkamp, *Role of Photoionization and Photochemistry in Ionization Processes of Organic-Molecules and Relevance*

for Matrix-Assisted Laser Desorption Ionization Mass-Spectrometry. Organic Mass Spectrometry, 1992. **27**(4): p. 472-480.

37. Zhang, J.Y., Y.F. Kong, and Y.F. Zhu, *Ionization mechanism of oligonucleotides in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.* Rapid Communications in Mass Spectrometry, 2001. **15**(1): p. 57-64.
38. Paul, W. and H. Steinwedel, **Ein Neues Massenspektrometer Ohne Magnetfeld.* Zeitschrift Fur Naturforschung Section a-a Journal of Physical Sciences, 1953. **8**(7): p. 448-450.
39. Hoffmann, E.d. and V. Stroobant, *Mass Spectrometry: Principles and Applications, 3rd edn., Wiley, New York.* 2007.
40. Todd, J.F.J., *Ion Trap Mass-Spectrometer - Past, Present, and Future.* Mass Spectrometry Reviews, 1991. **10**(1): p. 3-52.
41. March, R.E., *Quadrupole Ion Traps.* Mass Spectrometry Reviews, 2009. **28**(6): p. 961-989.
42. Cooks, R.G. and R.E. Kaiser, *Quadrupole Ion Trap Mass-Spectrometry.* Accounts of Chemical Research, 1990. **23**(7): p. 213-219.
43. Zhang, L.K., et al., *Accurate mass measurements by Fourier transform mass spectrometry.* Mass Spectrometry Reviews, 2005. **24**(2): p. 286-309.
44. Weickhardt, C., F. Moritz, and J. Grotemeyer, *Time-of-flight mass spectrometry: State-of-the-art in chemical analysis and molecular science.* Mass Spectrometry Reviews, 1996. **15**(3): p. 139-162.
45. Chernushevich, I.V., A.V. Loboda, and B.A. Thomson, *An introduction to quadrupole-time-of-flight mass spectrometry.* Journal of Mass Spectrometry, 2001. **36**(8): p. 849-865.
46. Guilhaus, M., *Principles and Instrumentation in Time-of-Flight Mass-Spectrometry - Physical and Instrumental Concepts.* Journal of Mass Spectrometry, 1995. **30**(11): p. 1519-1532.
47. Mamyrin, B.A., et al., *Mass-Reflectron a New Nonmagnetic Time-of-Flight High-Resolution Mass-Spectrometer.* Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki, 1973. **64**(1): p. 82-89.
48. Boesl, U., R. Weinkauf, and E.W. Schlag, *Reflectron Time-of-Flight Mass-Spectrometry and Laser Excitation for the Analysis of Neutrals, Ionized*

Molecules and Secondary Fragments. International Journal of Mass Spectrometry and Ion Processes, 1992. **112**(2-3): p. 121-166.

49. Mamyrin, B.A., *Laser-Assisted Reflectron Time-of-Flight Mass-Spectrometry.* International Journal of Mass Spectrometry and Ion Processes, 1994. **131**: p. 1-19.
50. Johnson, J.V., et al., *Tandem-in-Space and Tandem-in-Time Mass-Spectrometry - Triple Quadrupoles and Quadrupole Ion Traps.* Analytical Chemistry, 1990. **62**(20): p. 2162-2172.
51. Roepstorff, P. and J. Fohlman, *Proposal for a Common Nomenclature for Sequence Ions in Mass-Spectra of Peptides.* Biomedical Mass Spectrometry, 1984. **11**(11): p. 601-601.
52. Biemann, K., *Contributions of Mass-Spectrometry to Peptide and Protein-Structure.* Biomedical and Environmental Mass Spectrometry, 1988. **16**(1-12): p. 99-111.
53. Mann, M., R.C. Hendrickson, and A. Pandey, *Analysis of proteins and proteomes by mass spectrometry.* Annual Review of Biochemistry, 2001. **70**: p. 437-473.
54. Ballard, K.D. and S.J. Gaskell, *Sequential Mass-Spectrometry Applied to the Study of the Formation of Internal Fragment Ions of Protonated Peptides.* International Journal of Mass Spectrometry and Ion Processes, 1991. **111**: p. 173-189.
55. Scoble, H.A., S.A. Martin, and K. Biemann, *Peptide Sequencing by Magnetic Deflection Tandem Mass-Spectrometry.* Biochemical Journal, 1987. **245**(2): p. 621-622.
56. Summerfield, S.G., M.S. Bolgar, and S.J. Gaskell, *Promotion and stabilization of b(1) ions in peptide phenylthiocarbamoyl derivatives: Analogies with condensed-phase chemistry.* Journal of Mass Spectrometry, 1997. **32**(2): p. 225-231.
57. Gross, J.H., *Mass Spectrometry, 2nd edn., Springer, Heidelberg.* 2011.
58. Jennings, K.R., *Collision-induced decompositions of aromatic molecular ions.* International Journal of Mass Spectrometry and Ion Physics 1968. **1**: p. 227-235.

59. Haddon, W.F. and F.W. McLafferty, *Metastable Ion Characteristics .7. Collision-Induced Metastables.* Journal of the American Chemical Society, 1968. **90**(17): p. 4745-&.
60. McLuckey, S.A., *Principles of Collisional Activation in Analytical Mass-Spectrometry.* Journal of the American Society for Mass Spectrometry, 1992. **3**(6): p. 599-614.
61. Mayer, P.M. and C. Poon, *The Mechanisms of Collisional Activation of Ions in Mass Spectrometry.* Mass Spectrometry Reviews, 2009. **28**(4): p. 608-639.
62. McLuckey, S.A. and J.M. Wells, *Mass analysis at the advent of the 21st century.* Chemical Reviews, 2001. **101**(2): p. 571-606.
63. Claeys, M., et al., *Comparison of high- and low-energy collision-induced. Dissociation tandem mass spectrometry in the analysis of glycoalkaloids and their aglycons.* Journal of the American Society for Mass Spectrometry, 1996. **7**(2): p. 173-181.
64. Thorne, G.C., K.D. Ballard, and S.J. Gaskell, *Metastable Decomposition of Peptide [M + H]+ Ions Via Rearrangement Involving Loss of the C-Terminal Amino-Acid Residue.* Journal of the American Society for Mass Spectrometry, 1990. **1**(3): p. 249-257.
65. Zou, S., J. Oomens, and N.C. Polfer, *Competition between diketopiperazine and oxazolone formation in water loss products from protonated ArgGly and GlyArg.* International Journal of Mass Spectrometry, 2012. **316**: p. 12-17.
66. Zhang, Y., et al., *Optimized Orbitrap HCD for Quantitative Analysis of Phosphopeptides.* Journal of the American Society for Mass Spectrometry, 2009. **20**(8): p. 1425-1434.
67. Bean, M.F., et al., *Tandem Mass-Spectrometry of Peptides Using Hybrid and 4-Sector Instruments - a Comparative-Study.* Analytical Chemistry, 1991. **63**(14): p. 1473-1481.
68. Lau, K.W., et al., *Observations on the detection of b- and y-type ions in the collisionally activated decomposition spectra of protonated peptides.* Rapid Communications in Mass Spectrometry, 2009. **23**(10): p. 1508-1514.
69. Yates, J.R., et al., *Method to compare collision-induced dissociation spectra of peptides: Potential for library searching and subtractive analysis.* Analytical Chemistry, 1998. **70**(17): p. 3557-3565.

70. Harrison, A.G., et al., *Scrambling of sequence information in collision-induced dissociation of peptides.* Journal of the American Chemical Society, 2006. **128**(32): p. 10364-10365.
71. Polfer, N.C., et al., *On the dynamics of fragment isomerization in collision-induced dissociation of peptides.* Journal of Physical Chemistry A, 2008. **112**(6): p. 1286-1293.
72. Dongre, A.R., et al., *Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: Evidence for the mobile proton model.* Journal of the American Chemical Society, 1996. **118**(35): p. 8365-8374.
73. Wysocki, V.H., et al., *Special feature: Commentary - Mobile and localized protons: a framework for understanding peptide dissociation.* Journal of Mass Spectrometry, 2000. **35**(12): p. 1399-1406.
74. Somogyi, A., V.H. Wysocki, and I. Mayer, *The Effect of Protonation Site on Bond Strengths in Simple Peptides - Application of Ab-Initio and Modified Neglect of Differential-Overlap Bond Orders and Modified Neglect of Differential-Overlap Energy Partitioning.* Journal of the American Society for Mass Spectrometry, 1994. **5**(8): p. 704-717.
75. Paizs, B. and S. Suhai, *Fragmentation pathways of protonated peptides.* Mass Spectrometry Reviews, 2005. **24**(4): p. 508-548.
76. Mueller, D.R., M. Eckersley, and W.J. Richter, *Hydrogen Transfer-Reactions in the Formation of Y+2 Sequence Ions from Protonated Peptides.* Organic Mass Spectrometry, 1988. **23**(3): p. 217-222.
77. Cordero, M.M., J.J. Houser, and C. Wesdemiotis, *The Neutral Products Formed during Backbone Fragmentations of Protonated Peptides in Tandem Mass-Spectrometry.* Analytical Chemistry, 1993. **65**(11): p. 1594-1601.
78. Yalcin, T., et al., *Why are B ions stable species in peptide spectra?* Journal of the American Society for Mass Spectrometry, 1995. **6**(12): p. 1165-1174.
79. Yague, J., et al., *Peptide rearrangement during quadrupole ion trap fragmentation: Added complexity to MS/MS spectra.* Analytical Chemistry, 2003. **75**(6): p. 1524-1535.
80. Saminathan, I.S., et al., *The Extent and Effects of Peptide Sequence Scrambling Via Formation of Macrocyclic b Ions in Model Proteins.* Journal of the American Society for Mass Spectrometry, 2010. **21**(12): p. 2085-2094.

81. Yu, L., et al., *On the Relevance of Peptide Sequence Permutations in Shotgun Proteomics Studies.* Journal of Proteome Research, 2011. **10**(5): p. 2409-2416.
82. Goloborodko, A.A., et al., *Sequence Scrambling in Shotgun Proteomics is Negligible.* Journal of the American Society for Mass Spectrometry, 2011. **22**(7): p. 1121-1124.
83. Baba, T., et al., *Electron capture dissociation in a radio frequency ion trap.* Analytical Chemistry, 2004. **76**(15): p. 4263-4266.
84. Silivra, O.A., et al., *Electron capture dissociation of polypeptides in a three-dimensional quadrupole ion trap: Implementation and first results.* Journal of the American Society for Mass Spectrometry, 2005. **16**(1): p. 22-27.
85. Ding, L. and F.L. Brancia, *Electron capture dissociation in a digital ion trap mass spectrometer.* Analytical Chemistry, 2006. **78**(6): p. 1995-2000.
86. Chi, A., et al., *Analysis of intact proteins on a chromatographic time scale by electron transfer dissociation tandem mass spectrometry.* International Journal of Mass Spectrometry, 2007. **259**(1-3): p. 197-203.
87. Zubarev, R.A., et al., *Towards an understanding of the mechanism of electron-capture dissociation: a historical perspective and modern ideas.* European Journal of Mass Spectrometry, 2002. **8**(5): p. 337-349.
88. Zubarev, R.A., et al., *Electron capture dissociation of gaseous multiply-charged proteins is favored at disulfide bonds and other sites of high hydrogen atom affinity.* Journal of the American Chemical Society, 1999. **121**(12): p. 2857-2862.
89. Turecek, F., *N-C(alpha) bond dissociation energies and kinetics in amide and peptide radicals. Is the dissociation a non-ergodic process?* Journal of the American Chemical Society, 2003. **125**(19): p. 5954-5963.
90. Nagaraj, N., et al., *Feasibility of Large-Scale Phosphoproteomics with Higher Energy Collisional Dissociation Fragmentation.* Journal of Proteome Research, 2010. **9**(12): p. 6786-6794.
91. Haydon, C.E., et al., *Identification of novel phosphorylation sites on Xenopus laevis aurora A and analysis of phosphopeptide enrichment by immobilized metal-affinity chromatography.* Molecular & Cellular Proteomics, 2003. **2**(10): p. 1055-1067.

92. Bodenmiller, B., et al., *An integrated chemical, mass spectrometric and computational strategy for (quantitative) phosphoproteomics: application to Drosophila melanogaster Kc167 cells.* Molecular Biosystems, 2007. **3**(4): p. 275-286.
93. Lemeer, S. and A.J.R. Heck, *The phosphoproteomics data explosion.* Current Opinion in Chemical Biology, 2009. **13**(4): p. 414-420.
94. Mikesh, L.M., et al., *The utility of ETD mass spectrometry in proteomic analysis.* Biochimica Et Biophysica Acta-Proteins and Proteomics, 2006. **1764**(12): p. 1811-1822.
95. Sobczyk, M., et al., *Coulomb-assisted dissociative electron attachment: Application to a model peptide.* Journal of Physical Chemistry A, 2005. **109**(1): p. 250-258.
96. Syrstad, E.A. and F. Turecek, *Toward a general mechanism of electron capture dissociation.* Journal of the American Society for Mass Spectrometry, 2005. **16**(2): p. 208-224.
97. Savitski, M.M., et al., *Hydrogen rearrangement to and from radical z fragments in electron capture dissociation of peptides.* Journal of the American Society for Mass Spectrometry, 2007. **18**(1): p. 113-120.
98. Pitteri, S.J., P.A. Chrisman, and S.A. McLuckey, *Electron-transfer ion/ion reactions of doubly protonated peptides: Effect of elevated bath gas temperature.* Analytical Chemistry, 2005. **77**(17): p. 5662-5669.
99. Iavarone, A.T., K. Paech, and E.R. Williams, *Effects of charge state and cationizing agent on the electron capture dissociation of a peptide.* Analytical Chemistry, 2004. **76**(8): p. 2231-2238.
100. Swaney, D.L., et al., *Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors.* Analytical Chemistry, 2007. **79**(2): p. 477-485.
101. Scarff, C.A., et al., *Travelling wave ion mobility mass spectrometry studies of protein structure: biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements.* Rapid Communications in Mass Spectrometry, 2008. **22**(20): p. 3297-3304.

102. Smith, D.P., et al., *Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies.* European Journal of Mass Spectrometry, 2009. **15**(2): p. 113-130.
103. Erba, E.B., et al., *Ion Mobility-Mass Spectrometry Reveals the Influence of Subunit Packing and Charge on the Dissociation of Multiprotein Complexes.* Analytical Chemistry, 2010. **82**(23): p. 9702-9710.
104. Bernstein, S.L., et al., *Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease.* Nature Chemistry, 2009. **1**(4): p. 326-331.
105. Smith, D.P., S.E. Radford, and A.E. Ashcroft, *Elongated oligomers in beta(2)-microglobulin amyloid assembly revealed by ion mobility spectrometry-mass spectrometry.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(15): p. 6794-6798.
106. Cole, H.L., et al., *Characterizing Early Aggregates Formed by an Amyloidogenic Peptide by Mass Spectrometry.* Angewandte Chemie-International Edition, 2010. **49**(49): p. 9448-9451.
107. Thalassinos, K., et al., *Characterization of Phosphorylated Peptides Using Traveling Wave-Based and Drift Cell Ion Mobility Mass Spectrometry.* Analytical Chemistry, 2009. **81**(1): p. 248-254.
108. Shvartsburg, A.A., et al., *Separation of Peptide Isomers with Variant Modified Sites by High-Resolution Differential Ion Mobility Spectrometry.* Analytical Chemistry, 2010. **82**(19): p. 8327-8334.
109. Mie, A., et al., *Enantiomer separation of amino acids by complexation with chiral reference compounds and high-field asymmetric waveform ion mobility spectrometry: Preliminary results and possible limitations.* Analytical Chemistry, 2007. **79**(7): p. 2850-2858.
110. McCooeye, M., et al., *Separation and quantitation of the stereoisomers of ephedra alkaloids in natural health products using flow injection-electrospray ionization-high field asymmetric waveform ion mobility spectrometry-mass spectrometry.* Analytical Chemistry, 2003. **75**(11): p. 2538-2542.
111. Mason, E.A. and E.W. McDaniel, *Transport Properties of Ions in Gases, 1st edn., Wiley, New York.* 1988.

112. Bohrer, B.C., et al., *Biomolecule Analysis by Ion Mobility Spectrometry.* Annual Review of Analytical Chemistry, 2008. **1**: p. 293-327.
113. Wildgoose, J., et al., *Using a novel travelling wave ion mobility device coupled with a time-of-flight mass spectrometer for the analysis of intact proteins.* Molecular & Cellular Proteomics, 2006. **5**(10): p. S14-S14.
114. Leary, J.A., et al., *Methodology for Measuring Conformation of Solvent-Disrupted Protein Subunits using T-WAVE Ion Mobility MS: An Investigation into Eukaryotic Initiation Factors.* Journal of the American Society for Mass Spectrometry, 2009. **20**(9): p. 1699-1706.
115. Scarff, C.A., et al., *Probing Hemoglobin Structure by Means of Traveling-Wave Ion Mobility Mass Spectrometry.* Journal of the American Society for Mass Spectrometry, 2009. **20**(4): p. 625-631.
116. Giles, K., et al., *Applications of a travelling wave-based radio-frequencyonly stacked ring ion guide.* Rapid Communications in Mass Spectrometry, 2004. **18**(20): p. 2401-2414.
117. Pringle, S.D., et al., *An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument.* International Journal of Mass Spectrometry, 2007. **261**(1): p. 1-12.
118. Riba-Garcia, I., et al., *Evidence for structural variants of a- and b-type peptide fragment ions using combined ion Mobility/Mass spectrometry.* Journal of the American Society for Mass Spectrometry, 2008. **19**(4): p. 609-613.
119. Garcia, I.R., et al., *Studies of Peptide a- and b-Type Fragment Ions Using Stable Isotope Labeling and Integrated Ion Mobility/Tandem Mass Spectrometry.* Journal of the American Society for Mass Spectrometry, 2008. **19**(12): p. 1781-1787.
120. Ruotolo, B.T., et al., *Ion mobility-mass spectrometry analysis of large protein complexes.* Nature Protocols, 2008. **3**(7): p. 1139-1152.
121. Williams, J.P. and J.H. Scrivens, *Coupling desorption electrospray ionisation and neutral desorption/extractive electrospray ionisation with a travelling-wave based ion mobility mass spectrometer for the analysis of drugs.* Rapid Communications in Mass Spectrometry, 2008. **22**(2): p. 187-196.

122. Valentine, S.J., A.E. Counterman, and D.E. Clemmer, *A database of 660 peptide ion cross sections: Use of intrinsic size parameters for bona fide predictions of cross sections.* Journal of the American Society for Mass Spectrometry, 1999. **10**(11): p. 1188-1211.
123. Clemmer, D.E., Cross Section Database http://www.indiana.edu/~clemmer/Research/cross%20section%20database/cs%20database.htm.
124. Harvey, D.J., et al., *Ion Mobility Mass Spectrometry for Extracting Spectra of N-Glycans Directly from Incubation Mixtures Following Glycan Release: Application to Glycans from Engineered Glycoforms of Intact, Folded HIV gp120.* Journal of the American Society for Mass Spectrometry, 2011. **22**(3): p. 568-581.
125. Williams, J.P., et al., *Ion mobility augments the utility of mass spectrometry in the identification of human hemoglobin variants.* Rapid Communications in Mass Spectrometry, 2008. **22**(20): p. 3179-3186.
126. Purves, R.W. and R. Guevremont, *Electrospray ionization high-field asymmetric waveform ion mobility spectrometry-mass spectrometry.* Analytical Chemistry, 1999. **71**(13): p. 2346-2357.
127. Guevremont, R., et al., *Analysis of a tryptic digest of pig hemoglobin using ESI-FAIMS-MS.* Analytical Chemistry, 2000. **72**(19): p. 4577-4584.
128. Bradford, M.M., *Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding.* Analytical Biochemistry, 1976. **72**(1-2): p. 248-254.
129. Bush, M.F., et al., *Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology.* Analytical Chemistry, 2010. **82**(22): p. 9557-9565.
130. Xuan, Y., et al., *High-field asymmetric waveform ion mobility spectrometry (FAIMS) coupled with high-resolution electron transfer dissociation mass spectrometry for the analysis of isobaric phosphopeptides.* Rapid Communications in Mass Spectrometry, 2009. **23**(13): p. 1963-1969.
131. Shvartsburg, A.A. and R.D. Smith, *Fundamentals of Traveling Wave Ion Mobility Spectrometry.* Analytical Chemistry, 2008. **80**(24): p. 9689-9699.

132. Eyers, C.E., et al., *QCAL - a novel standard for assessing instrument conditions for proteome analysis.* Journal of the American Society for Mass Spectrometry, 2008. **19**(9): p. 1275-1280.
133. Eyers, C.E., et al., *CONSeQuence: Prediction of Reference Peptides for Absolute Quantitative Proteomics Using Consensus Machine Learning Approaches.* Molecular & Cellular Proteomics, 2011. **10**(11).
134. McCullough, B.J., et al., *Development of an ion mobility quadrupole time of flight mass spectrometer.* Analytical Chemistry, 2008. **80**(16): p. 6336-6344.
135. Wildgoose, J.L., et al., *Seattle.* 2006.
136. Jafari, M.T., *Improved design for high resolution electrospray ionization ion mobility spectrometry.* Talanta, 2009. **77**(5): p. 1632-1639.
137. Giles, K., J.P. Williams, and I. Campuzano, *Enhancements in travelling wave ion mobility resolution.* Rapid Communications in Mass Spectrometry, 2011. **25**(11): p. 1559-1566.
138. Tabb, D.L., et al., *Statistical characterization of ion trap tandem mass spectra from doubly charged tryptic peptides.* Analytical Chemistry, 2003. **75**(5): p. 1155-1163.
139. Breci, L.A., et al., *Cleavage N-terminal to proline: Analysis of a database of peptide tandem mass spectra.* Analytical Chemistry, 2003. **75**(9): p. 1963-1971.
140. Grewal, R.N., et al., *Fragmentation of protonated tripeptides: The proline effect revisited.* Journal of Physical Chemistry B, 2004. **108**(15): p. 4899-4908.
141. Smith, L.L., K.A. Herrmann, and V.H. Wysocki, *Investigation of gas phase ion structure for proline-containing b(2) ion.* Journal of the American Society for Mass Spectrometry, 2006. **17**(1): p. 20-28.
142. Clauser, K.R., P. Baker, and A.L. Burlingame, *Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS MS and database searching.* Analytical Chemistry, 1999. **71**(14): p. 2871-2882.
143. Chalkley, R.J., et al., *Statistical Analysis of Peptide Electron Transfer Dissociation Fragmentation Mass Spectrometry.* Analytical Chemistry, 2010. **82**(2): p. 579-584.

144. Nishikaze, T. and M. Takayama, *Influence of Charge State and Amino Acid Composition on Hydrogen Transfer in Electron Capture Dissociation of Peptides.* Journal of the American Society for Mass Spectrometry, 2010. **21**(12): p. 1979-1988.
145. Chawner, R., S.J. Gaskell, and C.E. Eyers, *Proposal for a common nomenclature for peptide fragment ions generated following sequence scrambling during collision-induced dissociation.* Rapid Communications in Mass Spectrometry, 2012. **26**(2): p. 205-206.
146. Vazquez, J., et al., *Peptide rearrangement during quadrupole ion trap fragmentation: Added complexity to MS/MS spectra.* Analytical Chemistry, 2003. **75**(6): p. 1524-1535.
147. Jia, C., W. Qi, and Z. He, *Cyclization Reaction of Peptide Fragment Ions during Multistage Collisionally Activated Decomposition: An Inducement to Lose Internal Amino-Acid Residues.* Journal of the American Society for Mass Spectrometry, 2007. **18**(4): p. 663-678.
148. Enjalbal, C., et al., *Low energy peptide fragmentations in an ESI-Q-Tof type mass spectrometer.* Journal of Proteome Research, 2007. **6**(4): p. 1378-1391.
149. Maitre, P., B.J. Bythell, and B. Paizs, *Cyclization and Rearrangement Reactions of a(n) Fragment Ions of Protonated Peptides.* Journal of the American Chemical Society, 2010. **132**(42): p. 14766-14779.
150. Molesworth, S.P. and M.J. Van Stipdonk, *Apparent Inhibition by Arginine of Macrocyclic b Ion Formation from Singly Charged Protonated Peptides.* Journal of the American Society for Mass Spectrometry, 2010. **21**(8): p. 1322-1328.
151. Molesworth, S., S. Osburn, and M. Van Stipdonk, *Influence of Amino Acid Side Chains on Apparent Selective Opening of Cyclic b(5) Ions.* Journal of the American Society for Mass Spectrometry, 2010. **21**(6): p. 1028-1036.
152. Bythell, B.J., et al., *Effect of the His Residue on the Cyclization of b Ions.* Journal of the American Society for Mass Spectrometry, 2010. **21**(8): p. 1352-1363.
153. Molesworth, S., S. Osburn, and M. Van Stipdonk, *Influence of Size on Apparent Scrambling of Sequence During CID of b-Type Ions.* Journal of the American Society for Mass Spectrometry, 2009. **20**(11): p. 2174-2181.

154. Goodlett, D.R., et al., *On the Relevance of Peptide Sequence Permutations in Shotgun Proteomics Studies.* Journal of Proteome Research, 2011. **10**(5): p. 2409-2416.
155. Siu, K.W.M., et al., *The Extent and Effects of Peptide Sequence Scrambling Via Formation of Macrocyclic b Ions in Model Proteins.* Journal of the American Society for Mass Spectrometry, 2010. **21**(12): p. 2085-2094.
156. Zubarev, R.A., et al., *Sequence Scrambling in Shotgun Proteomics is Negligible.* Journal of the American Society for Mass Spectrometry, 2011. **22**(7): p. 1121-1124.
157. Gaskell, S.J., et al., *Evidence for structural variants of a- and b-type peptide fragment ions using combined ion Mobility/Mass spectrometry.* Journal of the American Society for Mass Spectrometry, 2008. **19**(4): p. 609-613.
158. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.
159. Carroll, K.M., et al., *Absolute Quantification of the Glycolytic Pathway in Yeast: DEPLOYMENT OF A COMPLETE QconCAT APPROACH.* Molecular & Cellular Proteomics, 2011. **10**(12).
160. Burlet, O., et al., *Charge Promotion of Low-Energy Fragmentations of Peptide Ions.* Rapid Communications in Mass Spectrometry, 1992. **6**(11): p. 658-662.
161. Erlekam, U., et al., *Infrared Spectroscopy of Fragments of Protonated Peptides: Direct Evidence for Macrocyclic Structures of b(5) Ions.* Journal of the American Chemical Society, 2009. **131**(32): p. 11503-11508.
162. Atik, A.E. and T. Yalcin, *A Systematic Study of Acidic Peptides for b-Type Sequence Scrambling.* Journal of the American Society for Mass Spectrometry, 2011. **22**(1): p. 38-48.
163. Farrugia, J.M. and R.A.J. O'Hair, *Involvement of salt bridges in a novel gas phase rearrangement of protonated arginine-containing dipeptides which precedes fragmentation.* International Journal of Mass Spectrometry, 2003. **222**(1-3): p. 229-242.
164. Paizs, B., et al., *Ab initio and MS/MS studies on protonated peptides containing basic and acidic amino acid residues I. Solvated proton vs. salt-bridged structures and the cleavage of the terminal amide bond of protonated*

RD-NH2. International Journal of Mass Spectrometry, 2002. **219**(1): p. 203-232.

165. Yalcin, T. and A.G. Harrison, *Ion chemistry of protonated lysine derivatives.* Journal of Mass Spectrometry, 1996. **31**(11): p. 1237-1243.
166. Kish, M.M. and C. Wesdemiotis, *Selective cleavage at internal lysine residues in protonated vs. metalated peptides.* International Journal of Mass Spectrometry, 2003. **227**(1): p. 191-203.
167. Brancia, F.L., S.G. Oliver, and S.J. Gaskell, *Improved matrix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-containing peptides.* Rapid Communications in Mass Spectrometry, 2000. **14**(21): p. 2070-2073.
168. Gonzalez, J., et al., *Effect of the position of a basic amino acid on C-terminal rearrangement of protonated peptides upon collision-induced dissociation.* Journal of Mass Spectrometry, 1996. **31**(2): p. 150-158.
169. She, Y.M., et al., *Formation of (b(n-1)+H2O) ions by collisional activation of MALDI-formed peptide [M+H](+) ions in a QqTOF mass spectrometer.* Journal of the American Society for Mass Spectrometry, 2007. **18**(6): p. 1024-1037.
170. Bythell, B.J., et al., *Proton-Driven Amide Bond-Cleavage Pathways of Gas-Phase Peptide Ions Lacking Mobile Protons.* Journal of the American Chemical Society, 2009. **131**(39): p. 14057-14065.
171. Tsaprailis, G., et al., *Influence of secondary structure on the fragmentation of protonated peptides.* Journal of the American Chemical Society, 1999. **121**(22): p. 5142-5154.
172. Counterman, A.E. and D.E. Clemmer, *Cis-trans signatures of proline-containing tryptic peptides in the gas phase.* Analytical Chemistry, 2002. **74**(9): p. 1946-1951.
173. Bythell, B.J., et al., *Gas-Phase Structure and Fragmentation Pathways of Singly Protonated Peptides with N-Terminal Arginine.* Journal of Physical Chemistry B, 2010. **114**(46): p. 15092-15105.
174. Rankenberg, J.M., et al., *Proline Kink Angle Distributions for GWALP23 in Lipid Bilayers of Different Thicknesses.* Biochemistry, 2012. **51**(17): p. 3554-3564.