The University of Edinburgh

Structural Elucidation of Peptides and Proteins by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Nicolas Camille Polfer

Submitted in partial fulfilment for the degree of Doctor of Philosophy

November 2003

Fir t'Omo
de Raym
t'Misch
t'Catherine
an de Mathias

Der Deutsche hat für den Komplex des Daseins eines wirklichen Wesens das Wort Gestalt. Er abstrahiert bei diesem Ausdruck, von dem Beweglichen, er nimmt an, dass ein Zusammengehöriges festgestellt, abgeschlossen und in seinem Charakter fixiert sei. Betrachten wir aber alle Gestalten, besonders die organischen, so finden wir, dass nirgend ein Bestehendes, nirgend ein Ruhendes, ein Abgeschlossenes vorkommt, sondern dass vielmehr alles in einer steten Bewegung schwanke. Daher unsere Sprache das Wort Bildung sowohl von dem Hervorgebrachten, als von dem Hervorgebrachtwerdenden gehörig genug zu brachen pflegt.

Wollen wir also Morphologie einleiten, so dürfen wir nicht von Gestalt sprechen; sondern, wenn wir das Wort gebrauchen, uns allenfalls dabei nur die Idee, den Begriff oder in der Erfahrung nur für den Augenblick Festgehaltenes denken.

J. W. v. Goethe, Morphologie (1807)

For the whole existential complexity of the actual phenomenon German has the word Gestalt ('shape', 'form', character'). The expression enables one to abstract from fluidity, to assume that something that belongs together is thereby identified, detached, and fixed in its character. But if we actually contemplate all the forms (Gestalten) around us, especially the organic ones, nowhere do we find anything stable, at rest, or detached; rather, we find that everything is in constant flux. For that reason, the German language uses the word Bildung ('form', 'formation', 'shaping', 'constitution') to designate, appropriately enough, both the product and the process.

Translation by R. H. Stephenson; Goethes Conception of Knowledge and Science, Edinburgh University Press, ©1995.

Acknowledgements

The University of Edinburgh

Structural Elucidation of Peptides and Proteins by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Submitted in November 2003 in partial fulfilment of the degree of PhD

By Nicolas Camille Polfer

Abstract

This thesis describes the application of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) to the study of small proteins and peptides and their structures. Experiments were carried out using two instruments, a 3.0 Tesla FT-ICR mass spectrometer and a 9.4 Tesla FT-ICR mass spectrometer. Ions were produced via electrospray ionisation and fragmented in the ICR cell by either sustained off-resonance irradiation collisionally activated dissociation (SORI CAD) or electron capture dissociation (ECD). The latter technique has attracted considerable interest recently. It is an electronically mediated dissociation technique that provides complementary information to vibrational excitation techniques (e.g. SORI CAD). More importantly, ECD has been shown to be a powerful tool for the identification and localisation of post-translational modifications, such as glycosilations and phosphorylations.

Two unique features of FT-ICR mass spectrometry are high mass resolution and mass accuracy. These features have been exploited in a novel approach to the study of metal exchange reactions in metallo-proteins using stable isotope labelling. The system chosen, Zn_4-SmtA, is a cyanobacterial metallothionein (MT). Deconvoluted ESI-FT-ICR spectra for Zn_4-SmtA containing Zn isotopes in natural abundance as well as enriched with 93% ^{67}Zn show that the Zn_4 cluster of SmtA, in contrast to the structurally-analogous cluster of mammalian MT, contains a kinetically-inert Zn site, a feature which can be related to its secondary and tertiary structure, and which is of potential importance to its biological function.

The main body of this work is a systematic study of the dissociation, via ECD, of doubly protonated peptides, such as luteinising hormone releasing hormone (LHRH), bradykinin and bombesin. The most common fragments reported in the

literature occur from cleavage of the N-C_α backbone bond, yielding c' (N-terminal) and z$^\bullet$ (C-terminal) fragments. One striking observation in the spectra of the peptides studied is the presence of prominent radical c fragments (labelled herein c$^\bullet$) and even-electron z fragments (labelled z'). A study of the correlation between the relative abundance of each of the fragment ions as a function of the amino acid sequence for each of the peptides is shown.

These experimental studies have been complemented by molecular mechanics modelling, using the Amber force field, to correlate with putative gas-phase structures for these doubly protonated peptides. Here, a hypothesis is put forward that the relative ECD fragmentation pattern is related to the gas-phase structure of the peptide ion <u>prior to</u> electron capture, and in particular that hydrogen bonding of the protonated site to the backbone enhances backbone cleavage at that site (referred to as 'ECD/structure correlation'). The relative ECD fragmentation pattern for bombesin and four LHRH variants are compared to their modelled gas-structures. A detailed study of the LHRH variants with sequences pEHWSYGLRPG-OH {1} and pEHWSYGLRPG-NH$_2$ {2}, and bombesin {3}, showed good agreement between the relative ECD fragment abundances and predicted Amber structures for {1} and {3}, but not for {2}. This may imply that the postulated model is not a full description of the ECD process.

Articles and Related Papers in this Thesis

Articles included as part of this thesis (see Appendix 5, page 243)

Polfer, N. C.; Haselmann, K. F.; Zubarev, R. A.; Langridge-Smith, P. R. R. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 936-943.

Blindauer, C. A.; Polfer, N. C.; Keiper, S. E.; Harrison, M. D.; Robinson, N. J.; Langridge-Smith, P. R. R.; Sadler, P. J. *J. Am. Chem. Soc.* **2003**, *125*, 3226-3227.

Related articles not included as part of this thesis

Haselmann, K. F.; Budnik, B. A.; Kjeldsen, F.; Polfer, N. C.; Zubarev, R. A. *Eur. J. Mass Spectrom.* **2002**, *8*, 461-469

Conferences and meetings attended

Electron Capture Dissociation of Biomolecules and Related Phenomena, Stockholm, Sweden, October 19th-22nd, 2003.
'The role of gas-phase basicity in the electron capture dissociation mechanism' (oral).
16th International mass spectrometry conference, International conference centre, Edinburgh, UK, Sep 1st-5th, 2003.
'Mechanistic studies of electron capture dissociation' (oral).
Final year postgraduate talks seminar, School of Chemistry, University of Edinburgh, Edinburgh, UK, May 7th, 2003.
'Electron capture dissociation reveals secondary structure of peptides' (oral). Leanne Green memorial prize for best postgraduate talk.
Manipulation of electrons, atoms and molecules, School of Chemistry, University of Nottingham, Nottingham, UK, April 8th, 2003.
Conference on Ion Chemistry and Mass Spectrometry, UCLA conference center, Lake Arrowhead, California, January 17th-19th, 2003.
'Electron capture dissociation: a probe of polypeptide gas-phase structure?' (oral).
BMSS Trapped Ion Meeting, Laboratory of the Government Chemist, Teddington, UK, December 5th, 2002.
'Electron capture dissociation as a probe of the polypeptide gas-phase structure' (oral).
Bruker FTICR MS User Meeting, Laboratory of the Government Chemist, Teddington, December 4th, 2002.
'The basics of FTICR MS' (oral).
10th Anniversary of Department of Immunology, Laboratoire National de Santé, Luxembourg, October 30th-31st, 2002.
'Protein characterisation by Fourier transform ion cyclotron resonance mass spectrometry' (oral)
BMSS Annual Meeting, University of Loughborough, Loughborough, UK, September 9th-11th, 2002.

'Electron-mediated dissociation of polypeptides in FTICR MS' (oral). Barber Bardoli Award for best student talk.

ASMS Conference, Mariott Orlando World Center, Orlando, Florida, June 2nd-6th, 2002.

'Zn self-exchange in the metalloprotein Zn_4SmtA by FTICR MS' (poster).

'Electron kinetic energy measurements in electron capture dissociation: A retarding field analysis' (poster).

'Profiling of glucocorticoids by Fourier Transform Ion Cyclotron Resonance Mass Sepctrometry' (poster).

Table of contents

Chapter 1 Introduction

Chapter 2 Experimental set-up and preliminary results

Chapter 3 Mechanistic studies of electron capture dissociation

Chapter 4 Molecular mechanics force field modelling of peptide gas-phase ions

Chapter 5 Zn isotope exchange in the metallothionein Zn₄SmtA

Chapter 6 Further avenues in the study of the ECD mechanism

Abbreviations

BIRD	black-body infrared radiative dissociation
CAD	collisionally activated dissociation
CID	collisionally-induced dissociation
DFT	density functional theory
ε	dielectric constant or relative permittivity
ECD	electron capture dissociation
E_S	strain energy
ESI	electrospray ionisation
ESP	electrostatic potential
eV	electron volt
FT-ICR	Fourier transform ion cyclotron resonance
GB	gas-phase basicity
GnRH	gonadropin releasing hormone
Hbond	hydrogen bond
HF	Hartree Fock
IRMPD	infrared multi-photon dissociation
LHRH	luteinizing hormone releasing hormone
kcal	kilo calorie
kJ	kilo joule
MALDI	matrix-assisted laser desorption ionisation
mol	mole
PA	proton affinity
RESP	restrained electrostatic potential
RMS	root-mean-square
SID	surface-induced dissociation
SORI	sustained off-resonance irradiation

Units

	1 eV	1 kcal mol^{-1}	1 kJ mol^{-1}
eV	1	0.04838	0.01156
kcal mol^{-1}	20.670	1	0.2390
kJ mol^{-1}	86.485	4.184	1

Chapter 1

Introduction

1.1 Peptide and protein mass spectrometry - the big picture

In the past few years a milestone in genetics has been achieved with the deciphering of the human genome [1]. However the rewards from this achievement in terms of understanding cellular biology and obtaining medical breakthroughs may take many more years to materialise. While the genetic map of an organism may be viewed as an instruction book on how proteins should be expressed by a cell, the system is highly complex and dynamic. For the 30 thousand genes that are thought to exist for the human species, there may be as many as 1 million proteins if one includes their modifications (e.g. post-translational modifications) that can regulate the biological activity of a protein [2]. Since proteins, contrary to deoxyribonucleic acid (DNA), cannot be amplified (i.e. multiplied) analytical techniques are required that are sensitive enough to study biological systems. Hence, the study of proteins, which is also termed *proteomics*, is a more difficult project than *genomics* (i.e. study of genes) [3].

Mass spectrometry is one of the most sensitive analytical techniques available that gives qualitative and (with careful analysis) quantitative information for complex protein mixtures [4,5]. The common feature for mass spectrometric techniques is that molecules are transferred into the gas-phase as ions where they can be manipulated by electrostatic and/or magnetic fields to determine their mass-to-charge ratio (*m/z*). In the incipient days, mass spectrometry was applied to volatile compounds which can be easily transferred into the gas-phase by heating and which can then be ionised with electron ionisation (EI) [6]. Desorption techniques with fast energy deposition onto the sample enabled mass spectrometry to be applied to non-volatile compounds. These 'energy-sudden' methods include field desorption (FD) [7], plasma desorption (PD) [8], secondary ion mass spectrometry (SIMS) [9], fast atom bombardment (FAB) [10] and laser desorption (LD) [11]. However, for

macromolecules, such as proteins, more energy is required to transfer the molecule into the gas-phase than energy needed to break up the molecule.

Two ionisation techniques have enabled macromolecules to be studied by mass spectrometry. In matrix-assisted laser desorption ionisation (MALDI) [12] the laser photons heat a matrix which aids transfer of the analyte into the gas-phase as ions. Koichi Tanaka was first to extend this technique to biomolecules [13], for which he was awarded the Nobel Prize in Chemistry in 2002. In electrospray ionisation (ESI) in the presence of an electrostatic field a Taylor cusp is formed which causes formation of droplets, from which ions pass into the gas-phase (a more detailed account of ESI will be given in section 1.3.2). John B. Fenn also shared the Nobel Prize in Chemistry in 2002 for his contribution in developing electrospray ionisation for biological mass spectrometry [14].

A biological sample (e.g. cell extract) may contain tens of thousands of proteins, the abundances of which may vary by a factor of 10^6 [15]. Hence sophisticated separation techniques are required to reduce the complexity of the sample prior to mass spectrometric analysis. The cutting-edge technology for protein separation is two-dimensional polyacrylamide gel electrophoresis (2D PAGE), where proteins are supported on a polymer and separated in one dimension based on their masses by the application of an electric field gradient [16,17]. Proteins are separated in the second dimension based on their isoelectric point (i.e. pH where protein is neutral). The isolated proteins are cut into peptide fragments by an enzyme (e.g. trypsin), which is technically known as *enzymatic digestion* of proteins.

Mass spectrometry assists protein identification by two main approaches. In the 'peptide- mass mapping' (or 'peptide-mass fingerprint') approach [18,19] the enzymatic peptide fragments are mass analysed and these masses are compared to theoretical mass spectra predicted from databases that contain the genetic sequence information. Ionisation of peptides by MALDI coupled with time-of-flight (TOF) mass analysis has been shown to be readily amenable to automated analysis [20,21]. The second approach to protein identification consists of peptide ion mass isolation (MSMS) and fragmentation in a mass spectrometer for the purpose of obtaining the primary structure (i.e. sequence) of the peptide [22,23]. These peptide sequences are

scanned against known protein sequences contained in databases in order to establish the identity of the protein. While this approach provides more reliable identification of proteins, the implementation of automated peptide fragmentation in a mass spectrometer is more technically challenging than the 'peptide-mass mapping' approach.

A limitation of 2D PAGE (two-dimensional polyacrylamide gel electrophoresis) is its insensitivity for many of the lower abundance proteins (i.e. < 1,000 copies per cell) [15]. Recent advances in improving the sensitivity of protein identification by 2D PAGE include the use of affinity tags (e.g. antibodies) to purify protein samples [24,25]. Since 2D PAGE is a labour-intensive and expensive technique several research groups are developing approaches to separate protein mixtures by chromatographic techniques on-line to mass spectrometric analysis [26] and fragmentation in the mass spectrometer [27,28]. The role of mass spectrometry in *proteomics* and current challenges have been reviewed in the literature [5,15,29].

1.2   Contribution of this work

This work presents a new approach to study the kinetic metal exchange in metalloproteins using metal isotope exchange for the whole cyanobacterial metallothionein Zn_4-SmtA (see Chapter 5) [Blindauer Claudia, 2003 #72]. One of the Zn metal sites appears to be inert to exchange, which is linked to the tertiary conformation of the protein and may be of biological significance. The high resolving power and mass accuracy of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) are essential to establish the amount of Zn that has exchanged.

The main body of this work involves the study of a relatively novel peptide fragmentation technique, electron capture dissociation (ECD) [30], which is a promising tool to identify and localise post-translational modifications [31-34] and has so far only been demonstrated on FT-ICR MS instruments. The equipment used in this work and the approach used to produce reliable ECD mass spectra is shown in Chapter 2. Systematic studies of ECD on model polypeptide ions are discussed in mechanistic terms in Chapter 3. Some of these peptides are modelled using the Amber mechanical force field [35,36] to determine if there is a correlation between

the modelled gas-phase structures and the ECD fragmentation pattern (Chapter 4). The role of a peptide fragmentation technique is to provide primary structure information (i.e. sequence) of a peptide in order to identify it; ECD may also reveal secondary structure information of polypeptides. Chapter 6 hints at further mechanistic studies of ECD with synthetic peptides and gives a summary of the present work on ECD.

Thus, FT-ICR MS is ideally suited for structural elucidation of biologically relevant molecules, and of peptides and proteins in particular. Further, the high resolution and mass accuracy as well as ion trapping abilities of FT-ICR MS enable more interesting biological problems to be studied, which cannot be studied with any other mass spectrometric technique.

1.3 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

1.3.1 History and developments

A Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) works on the premise that ions in the presence of a uniform magnetic field describe a circular motion in the perpendicular plane to the magnetic field direction, named the cyclotron motion (see section 1.3.2 for cyclotron theory). The cyclotron frequency of each ion, which is dependent on its mass-to-charge ratio (m/z), is measured as an induced current on the detection plates in a detection cell.

Cyclotron motion was first observed by Lawrence and Livingstone in 1932 [37]. The first mass spectrometer to make use of the ICR principle was the omegatron [38,39], which was constructed in 1949. In the omegatron and the ICR mass spectrometers developed thereafter [40-42] the magnetic field is scanned to generate mass spectra. In the 1970's two fundamental modifications to ICR gave rise to the field of FT-ICR MS. The first was the idea by McIver to trap ions using an electric field combined with a magnetic field [43]; the second improvement was the introduction of the Fourier Transform algorithm by Marshall and Comisarow [44]. The former improvement enabled ions to be trapped in three dimensions and therefore allow the measurement of ions for longer periods of time, whereas the latter

improvement made it possible to detect ions of different mass-to-charge ratios at the same time and de-convolute their cyclotron frequencies. In FT-ICR MS, as opposed to ICR MS, the magnetic field strength is held constant. The advent of superconductive magnets in the 1980's increased the achievable mass resolution and extended the upper mass limit.

Whereas the first ICR mass spectrometers used internal electron ionisation (EI) to form ions [37], nowadays the ionisation is generally performed external to the magnetic field. External ionisation has a number of advantages over internal ionisation, including easier sample introduction and easier changeover from one ionisation technique to another. Some ionisation techniques, and in particular electrospray ionisation (ESI), which is operated at atmospheric pressure, are in fact incompatible with the low pressure present in the detection cell during ion detection ($< 10^{-9}$ mbar). In order to bridge the pressure difference between the ionisation source and the detection cell, multiple pumping stages are required (differential pumping). ESI was first interfaced with FT-ICR in 1986 [45]. Since then other ionisation techniques which are important to the study of biological molecules have been interfaced with the technique, including matrix assisted laser desorption ionisation (MALDI) [46] and smaller volume ESI techniques: nano ESI [47] and micro ESI [48]. ESI works well in conjunction with FT-ICR, since the most abundant charge states in the mass spectra of proteins are generally observed between *m/z* 500 to 2500, which is a *m/z* range where high mass resolving power is readily achievable. Furthermore, the ion population generated by ESI is very reproducible and controllable in contrast to MALDI, thereby making the external calibration of spectra much more accurate.

The advantages of FT-ICR compared to other mass analysers are its ultra-high mass-resolving power (10^6), high mass accuracy ($<$ ppm), good sensitivity, non-destructive detection, long ion trapping times (minutes to hours) and tandem mass spectral (MS/MS) capabilities in time. Mass resolutions in excess of 1 million have been reported [49], and protein concentrations at the attomole (10^{-18} mol) level have been detected [50]. Belov *et al.* have even achieved sub-attomole detection limits by using an ion funnel [51], which increases the overall ionisation efficiency of electrospray. Some methodological advances, such as stored waveform inverse

Fourier transform (SWIFT) [52] have made FT-ICR much more performant with respect to tandem mass spectrometry (MSMS).

Many instrument development improvements have been proposed, such as using a dual-trap design, where one detection cell accumulates ions prior to introducing the ions into the actual detection cell [53]. This approach makes quadrupolar excitation axialisation (QEA) much easier to operate. QEA has been shown to be effective at reaxialising the ions to the centre of the trap, which increases the mass accuracy of the measurement and increases the number of peaks observed in the spectrum [54-56]. Since FT-ICR has fairly long duty cycles (seconds), such a dual-trap would also minimise sample loss when interfacing FT-ICR with a chromatographic run. The long duty cycle of FT-ICR is in fact one of the major drawbacks of the technique with respect to chromatographic separation techniques. Nevertheless, FT-ICR has already been successfully interfaced with chromatographic separation techniques, such as capillary electrophoresis (CE) ESI to detect endogenous levels of hemoglobin from cell lysates [57]. A historical perspective of FT-ICR MS developments is given by Alan Marshall in his 'milestones in FT-ICR MS development' review paper [58].

1.3.2 Fundamentals of FT-ICR MS

The principles of FT-ICR MS have been reviewed abundantly in the literature [59-62]. In the presence of a uniform magnetic field ions undergo a circular motion in the plane perpendicular to the direction of the magnetic field, which is called the cyclotron motion. The fundamental reason for the cyclotron motion is the Lorentz force (depicted in orange in Figure 1.1), which acts on a positive ion, as shown in Figure 1.1. It should be noted that for negative ions the Lorentz force acts in the opposite direction, causing the cyclotron motion of negative ions to occur in the opposite direction.
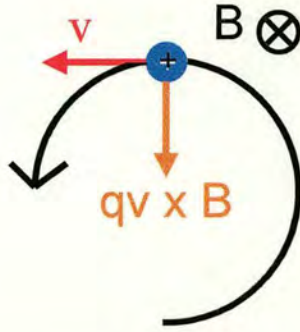
Figure 1.1 Diagrammatic representation the cyclotron motion, showing the Lorentz force (qv x B) acting on a positive ion, where v is the velocity vector of the ion, B is the magnetic field vector and q is the charge of ion.

The unperturbed angular cyclotron frequency ω_c (Equation 1.1) (radians s^{-1}) is dependent on the magnetic field strength B_0 (Tesla) and the inverse of the mass-to-charge ratio m/q (kg C^{-1}). This implies that higher mass-to-charge species have lower cyclotron frequencies.

$$\omega_c = \frac{qB_0}{m}$$

Equation 1.1

Figure 1.2 shows the diagram of a cubic ICR cell taken from Marshall and co-workers [62]. The directions for this cell are defined such that the magnetic field direction is co-axial with the z direction. This means that the Lorentz force contains the ions radially (xy plane) and in order to trap ions in 3 dimensions, i.e. also in the direction of the magnetic field (z direction), electrostatic trapping plates are required.
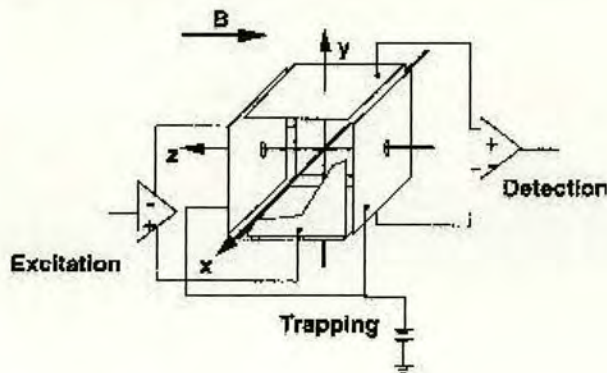
Figure 1.2 Cubic ICR cell, B showing direction of magnetic field [62].

The effect of the electric field generated by the trapping plates is to give rise to other ion motions besides the cyclotron motion: the *trapping* motion (angular frequency ω_T – Equation 1.2) in the z direction and the *magnetron* motion (angular frequency ω_m – Equation 1.3), which takes place in the xy plane. The combined trajectory of these motions is similar to a spiral, to and from each trapping plate. A visual representation of the fundamental motions in the ICR cell is given in section 1.3.3.1 (Figures 1.3, 1.4 and 1.5). It has been shown that the frequencies of these motions are given by the following equations:

$$\omega_T = \sqrt{\frac{q\alpha V_T}{\pi^2 m a^2}} \qquad\qquad \text{[62]} \quad \textbf{Equation 1.2}$$

$$\omega_m = \frac{\alpha V_T}{a^2 B_0} \qquad\qquad \text{[63]} \quad \textbf{Equation 1.3}$$

where V_T is the trapping plate potential (V), a is the length of the cell (cm) and α is a term representing the geometry of the cell. Magnetron motion, which takes place in the xy plane in opposite direction of the cyclotron motion, causes a reduction in the observed cyclotron frequency, resulting in the *reduced* angular cyclotron frequency ω_+ (Equation 1.4). The *reduced* cyclotron frequency is also known as the *observed* cyclotron frequency, since it is the measured frequency in FT-ICR MS.

$$\omega_+ = \omega_c - \omega_m \qquad\qquad\qquad\qquad \textbf{Equation 1.4}$$

In contrast to the cyclotron frequency the magnetron frequency is virtually mass-to-charge <u>independent</u>. At normal mass-to-charge ratios (m/z) the magnitude of the magnetron frequency (100's Hz) is much lower than the cyclotron frequency (10's kHz), however at high values of m/z ($> 10^5$) the cyclotron frequency approaches the magnetron frequency. The magnetron frequency is proportional to the trapping plate voltages V_T, thus it is desirable to employ low trapping voltages.

When the ions are trapped in the detection cell they are positioned close to the centre of the trap (in the xy plane) and the cyclotron motions of ions of a particular m/z are not in phase with one another. In order to detect ions in the form of an induced current on the detection plates, the cyclotron radii of the ions need to be

increased. This is effected by employing a radiofrequency (rf) dipolar excitation[1] on opposing excitation plates of the same frequency as the cyclotron frequency of the ions. Equation 1.5 shows the expression for the post-excitation radius r_{exc} (cm) for a dipolar excitation. r_{exc} is dependent on the length of excitation t_{exc} (s), the distance between the excitation plates d (cm) and the peak-to-peak voltage V_{p-p} (V), and inversely proportional to the magnetic field strength B_0 Importantly, r_{exc} is independent of m/z.

$$r_{exc} = \frac{V_{p-p} t_{exc}}{2 d B_0} \qquad [62] \qquad \textbf{Equation 1.5}$$

Another reason for employing a dipolar excitation is that ions of a particular m/z need to be brought into phase with one another, so that they can be detected as a coherent ion packet on the detection plates. As shown in Figure 1.2, the detection plates are electrically connected and hence the movement of ions between these two plates causes the appearance of induced current (or image current) between the plates. The amplitude of this current is dependent on the charge moving between the two detection plates, according to Earnshaw's theorem [62]. In practical terms this means that the amplitude of the detected current is dependent on the charge and number of ions of a particular m/z. The amplitude of the induced current is also dependent on the distance of the charge to the detection plates, however, since the post-excitation radius is independent of m/z, it is the same for all ions. This feature makes FT-ICR a quantitative technique.

In order to excite ions of different m/z, a sweep rf dipolar excitation is applied on the excitation plates. The resulting complex induced current signal on the detection plates encodes the frequencies of all ions that are excited. A Fourier Transform of this signal gives the individual frequency components of the induced current signal. An easily accessible tutorial for the operation of FT-ICR MS has been published by Amster [64].

---

[1] In a dipolar excitation at any moment in time the voltage on one excitation plate is 180° out of phase with the voltage on the other excitation plate. If the dipolar excitation frequency is identical to the cyclotron frequency, then the ion is accelerated (i.e. cyclotron radius increases) due to resonant absorption (see section 1.3.3.2, Figure 1.7).

The induced current signal on the detection plates is referred to as the *time-domain ICR signal* [62]. The amplitude of the *time-domain ICR signal* (which is also termed *free induced decay* or *transient*) decays exponentially as a result of loss of coherence of the ion packet due to collisions with background gas. The *time-domain ICR signal* is digitised using an analog-to-digital converter and then Fourier transformed into a frequency domain spectrum. The digitisation rate needs to be at least twice as large as the largest cyclotron frequency to be measured (lowest *m/z*) according to the Nyquist criterion; the Nyquist criterion states that at least 2 points in a sine (or cosine) wave period have to be known in order to unambiguously determine its frequency. The frequency domain spectrum is converted into a mass-to-charge spectrum using the fundamental cyclotron frequency equation (Equation 1.1).

The mass resolution in FT-ICR MS is defined as the full width of the spectral peak at half-maximum peak height (FWHM), which is referred to as $\Delta\omega_{50\%}$ in the frequency domain and $\Delta m_{50\%}$ in the mass domain. Generally, the mass resolving power $\omega/\Delta\omega_{50\%}$ ($= -m/\Delta m_{50\%}$) is more widely used as a measure of peak resolution in FT-ICR MS. In the low-pressure limit (i.e. ICR time-domain signal persists undamped during acquisition period) the mass resolving power is proportional to the cyclotron frequency of the ion and the length of the acquisition time. This implies that lower *m/z* ions, having higher cyclotron frequencies, also have higher peak resolutions than higher *m/z* species (for the same transient length). Although the peak width $\Delta\omega_{50\%}$ remains constant with increasing *m/z* the cyclotron frequency varies inversely with *m/z*, which means that the mass resolving power varies inversely with *m/z* [62]. Longer acquisition times also result in higher peak resolution. This explains the requirement for an ultra-high vacuum in the detection cell in order to minimise collisions with background gas.

While mass resolution is important, mass accuracy is arguably an even more important requirement in mass spectrometry. To obtain very accurate mass data (i.e. < ppm), the data needs to be calibrated against standard known *m/z*. The calibration takes into account cyclotron frequency shifts due to the magnetron frequency and space-charge effects between ions. To maximise mass accuracy low trapping voltages should be employed (< 1V) and not too many ions should be introduced into

the detection cell to minimise space-charge effects. The mass accuracy expression is given by:

$$\text{mass accuracy} = (m_{exp} - m_{theor})/m_{theor} \times 10^6 \text{ (in ppm)}$$ **Equation 1.6**

Ledford *et al.* proposed the following equation to calibrate the mass spectra, which is generally used in FT-ICR [65].

$$\frac{m}{q} = \frac{A}{\omega_+} + \frac{B}{\omega_+^2}$$ **Equation 1.7**

where A and B are constants which are derived from FT-ICR measured frequencies of standards of known *m/z*.

Coalescence of peaks that are very close in mass is a feature often observed in FT-ICR MS, even if the apparent mass resolving power should be high enough to resolve every peak. It was found that employing very low trapping potentials (e.g. 0.2 V) is the best approach to minimise peak coalescence [66], since it results in a reduction of space-charge effects.

In practice, the largest limitation to high resolution and mass accuracy are space charge-effects between ions (in the low pressure limit). The highest mass that has effectively been unit mass resolved is 112 kDa [67], however for molecules above 50 kDa it is difficult to generate high mass resolving power (i.e. $> 10^5$) spectra due to the very large number of isotopic peaks present [68].

1.3.3 Simulation of ion motion in the ICR cell

Given the complexity of an ICR cell, which comprises electrostatic and magnetic trapping of the ions, as well as radiofrequency (rf) fields for dipolar and quadrupolar excitation of the ions, it would be of great benefit to be able to visualise the trajectory of the ions in the ICR cell in order to gain an intuitive understanding. SIMION[2] is an ion optics simulation program (developed by David Dahl working at the Idaho National Engineering Laboratory (Idaho, USA)) that allows the construction of ion optics in the program, ion trajectory simulations in these ion

---

[2] http://www.simion.com/

optics and visualisation of the ion motions [69]. A review of the development of the SIMION program from a historical perspective has been given by David Dahl [70]. This review includes the present version of SIMION (version 7.0) that was used in these simulations and explains many of the concepts that can only be briefly explained here.

A brief explanation of the theory underlying the SIMION program is given in Appendix 1, as well as a description of the open–ended ICR cell which was designed by Dr. Michael Easterling. This section is primarily concerned with giving a brief description of ion motions in the ICR cell.

The shown simulations are all for a positive ion of m/z 1000 (mass = 1000 amu, charge = $+1.6 \times 10^{-19}$ C) and a magnetic field strength of 4.7 T. The trapping plate voltage is +1 V unless otherwise specified.

1.3.3.1 Fundamental motions in the ICR cell

Figure 1.3 shows a cut-open cell with the trajectory of a positive ion of m/z of 1000. The spiral-like trajectory is a superposition of the cyclotron and trapping motions. The Lorentz force due to the magnetic field constrains the ion in two dimensions (xy). The voltages on the trapping plates cause a potential energy well in the centre of the cell (see Figure 1.4), which enables trapping of the ions in the z direction (i.e. axis of magnetic field).
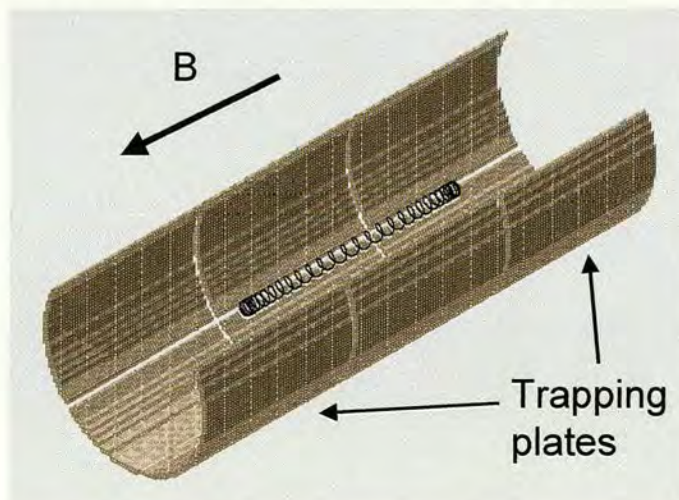
Figure 1.3 Ion trajectory simulation in an ICR cell for 4.7 T magnetic field strength showing cyclotron motion (in xy direction) and trapping motion (in z direction). Ion m/z = 1000, trapping plate voltages = +1 V, cyclotron kinetic energy 5 eV.
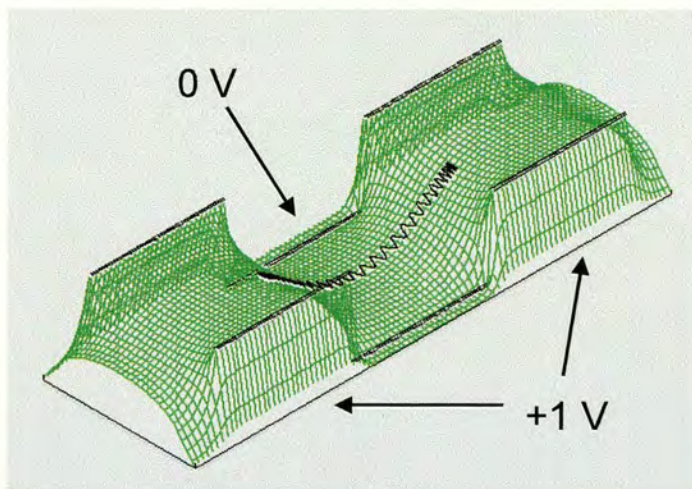
Figure 1.4 Potential energy view of Figure 1.3 for xz plane. The trapping frequency depends on the trapping plate voltages (see Equation 1.2).

The magnetron motion also occurs in the xy plane, but since its frequency is much lower (i.e. 100's Hz) it cannot be seen in Figure 1.3. By employing much higher trapping plate voltages (e.g. 500 V) than would normally be employed in FT-ICR and placing the ion off-centre in the xy direction, the magnetron motion can be visualised (see Figure 1.5).
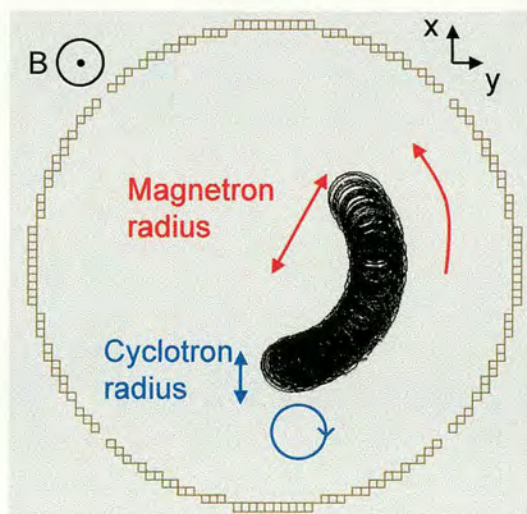
Figure 1.5 Ion trajectory simulation showing magnetron (red colour code) and cyclotron (blue colour code) motions (cross-section of cell in xy direction). Magnetic field direction B is out of plane of paper.

1.3.3.2 Dipolar excitation

When applying a dipolar excitation on the excitation plates of the same frequency as the cyclotron frequency of the ion (on-resonance excitation), the cyclotron radius of the ion increases linearly with time (see Figure 1.6).
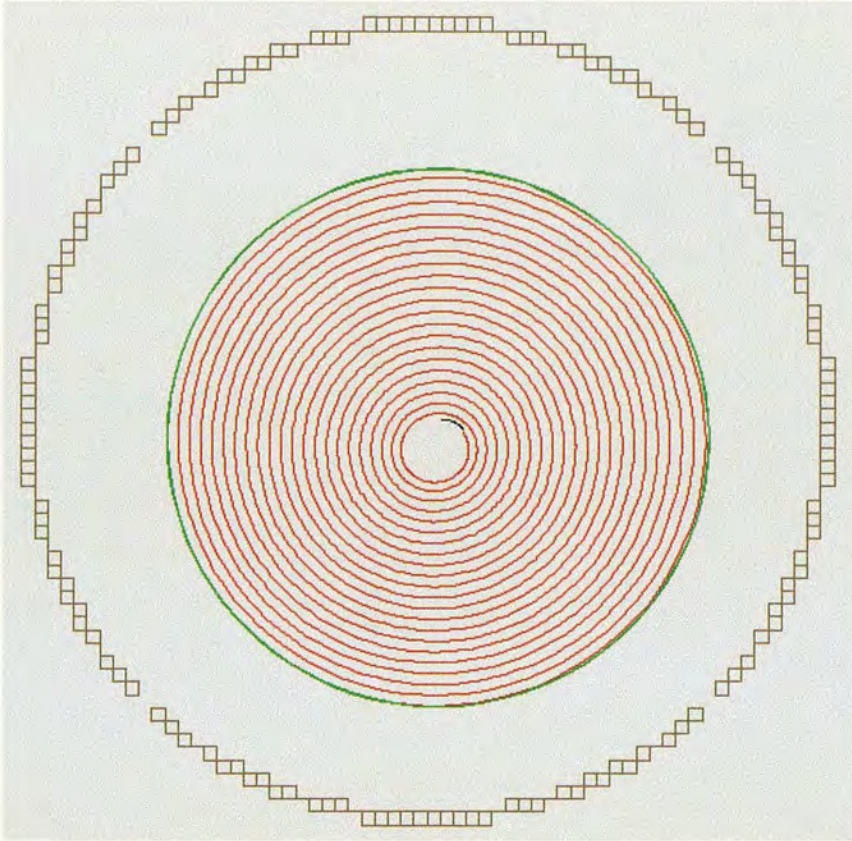
Figure 1.6 Ion trajectory for on-resonance dipolar excitation. xy cross-section of cell. V_{p-p} 20 V, excitation time = 300 μs. Excitation (red colour code) and post-excitation radius (green colour code). Initial cyclotron kinetic energy = 5 eV, post-excitation cyclotron kinetic energy ~ 580 eV.

If, conversely, an off-resonance dipolar excitation pulse is applied, the ion goes through cycles of increases and decreases of the cyclotron radius (see Figure 1.7). The cyclotron radius increases when the excitation and cyclotron phases are in resonance, but decreases when both are out of phase.
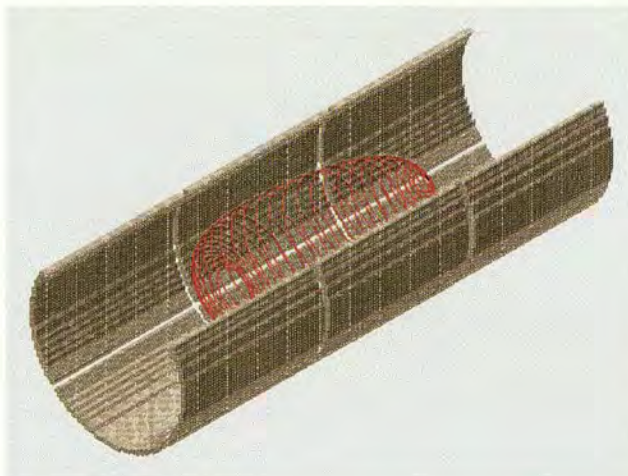
Figure 1.7 Ion trajectory for off-resonance dipolar excitation. $V_{p-p} = 40V$, off-resonance = +2000 Hz.

For a 2000 Hz off-set of the excitation frequency, the cyclotron and excitation phases are in resonance every 500 μs. This implies that the cyclotron radius will increase to a maximum value after 250 μs of excitation. Off-resonance dipolar excitation is used in sustained off-resonance irradiation collisionally activated dissociation (SORI CAD) [71]. In SORI CAD the pressure in the ICR cell is raised by pulsing a gas (e.g. Ar). The increase in the number of collisions manifests itself in an increase in the internal energy of ions. The maximum collision energy to the ion can be adjusted by changing the excitation frequency off-set, length of excitation and the peak-to-peak voltage of excitation. SORI CAD will be discussed again in the section concerned with polypeptide fragmentation (section 1.4).

1.4 Protein and polypeptide structure

1.4.1 Primary structure

The basic building blocks of peptides or proteins are the amino acid. There are 20 different α-amino acids encoded by DNA. These amino acids have a chiral carbon centre (Figure 1.8), apart from glycine, and are therefore optically active. All naturally occurring amino acids are L stereoisomers, apart from a very few exceptions. Table A2.1 (Appendix 2) summarises the 20 amino acids found in nature and their residue masses (i.e. mass of -NH-CHR-CO-, where R is side-chain).

In a protein amino acids are chemically bound by an amide/peptide bond. Molecules of less than 20 amino acids are generally referred to as polypeptides, whereas larger entities are called proteins. By convention, one always starts reading the primary structure (i.e. sequence) of a protein from the free amine site (named the N-terminus) and proceeds to the carboxylic acid group (named the C-terminus). The primary structure of a protein is the fundamental property by which it is identified.

$$H_2N-\underset{\underset{H}{|}}{\overset{\overset{R}{|}}{C^*}}\overset{O}{-\overset{\diagup}{\underset{\diagdown}{}}}OH$$
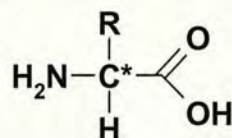
Figure 1.8 General structure for α-amino acid, where C* denotes the chiral carbon.

1.4.2 Differences in solution and gas-phase structures

In solution protein folding is driven by the difference between the intramolecular interactions and the protein-solvent interactions. Both of these interactions have large values, whereas their difference is fairly small. This makes the task of modelling such systems extremely difficult. This challenge is commonly referred to as the *protein folding problem*. By removing the solvent, it is possible to study the intramolecular interactions separately and then add solvent molecules one at a time to study the effect of solvation [72].

The secondary structure of proteins [73] (e.g. α-helix, β-sheet)[3] and higher structures (tertiary folding, multiple protein assemblies) determine biological activity. How well the gas-phase structure mirrors the solution phase structure is a matter of debate. Differences in gas-phase and solution phase structures arise due to the fact that in solution the basic sites are generally located on the outside of the protein, in order to maximise protein-solvent interaction, whereas in the gas-phase the basic sites are generally involved in back solvation and hence located inside the structure. However, due to overall conformational constraints (i.e. large energy

---

[3] For simple introduction course, see http://www.expasy.org/swissmod/course/text/chapter1.htm and http://www.expasy.org/swissmod/course/text/chapter2.htm.

16

barriers between conformations), such a complete re-arrangement may not be possible and hence the solution phase structure could be 'frozen' into the gas-phase. This may especially apply to ionisation techniques that give little energy to the molecule in the ionisation process, such as electrospray ionisation.

1.4.3 Electrospray ionisation

Electrospray ionisation (ESI) is a soft ionisation technique that allows even non-covalent interactions to be conserved into the gas-phase [74]. The exact mechanism in ESI as to how ions pass from droplets into the gas-phase is still controversial. Dole *et al.* [75,76] suggested that solvent evaporation from the droplet causes a series of Coulomb explosions as the Rayleigh limit is reached resulting in ever smaller droplets and ending in the formation of free molecular ions (this is referred to as the *charged residue model*). Conversely, Iribane and Thompson [77] proposed that ions could directly desorb from the droplet surface into the gas-phase; multiple Coulombic explosions of droplets give rise to very small droplets with very high charge densities and hence the resulting electrostatic field at the droplet surface could transfer ions directly into the gas-phase (this is referred to as the *ion evaporation model*). Fenn used the ion evaporation model to explain the typically 'bell-shaped' charge state distribution that is often observed for larger molecules, such as proteins [78]; since the number of charges on an ion is probably related to the number of charges available at the droplet surface at the time of evaporation, ions with lower charge states are formed earlier in the evaporation process than ions with higher charge states. Nevertheless, for much larger molecules Nohmi and Fenn have shown that the charge residue model can explain ion formation [79].

Recently Beauchamp and co-workers studied the droplet evaporation and discharge dynamics for electrosprayed droplets in a beautiful experiment involving Doppler interferometry to determine the droplet size and mobility in a constant field to determine the charge [80]. They inferred from these experiments that droplet instability occurs near the Rayleigh limit of charge resulting in discharge losses of 15-40% of the total charge (depending on the solvent used) and only small mass losses.

1.4.4 Solution and gas-phase basicity

Whereas acid-base functionality in solution is defined by pKa values[4], in the gas-phase this concept is expressed by the gas-phase basicity or proton affinity (see Table A2.2, Appendix 2). The gas-phase basicity and proton affinity are defined as - ΔG and $-\Delta H$ of the reaction $[M+H]^+ \leftrightarrow MH^+$ [81]. Polar solvents (e.g. water) stabilise protonated bases in solution relative to protonated bases in the gas-phase due to favourable charge-dipole moment interactions (i.e. electrostatic interactions) [82]. This effect is more pronounced for ammonia than for imidazole, which explains why lysine is a much stronger base than histidine in solution (i.e. $pKa_{lysine} > pKa_{histdine}$), whereas the order of basicity is reversed in the gas-phase (i.e. $GB_{histidine} > GB_{lysine}$) (see Table A2.2, Appendix 2). Hence, the order of basicity in solution and the gas-phase are not equivalent. Water also has a high dielectric constant[5] ($\varepsilon_{H2O} = 80$) which favours stabilisation of charges on amino acids that are not stabilised in the gas-phase [82].

The gas-phase basicity is in general determined by two methods - the *deprotonation* and the *kinetic* methods. In the *deprotonation* method [83,84] the protonated amino acid is reacted with a neutral standard compound of known gas-phase basicity to abstract the proton. If the gas-phase basicity of this standard is higher this will result in deprotonation of the amino acid, if it is lower no proton transfer will take place. Thus, the gas-phase basicity of the amino acid can be bracketed in between the known gas-phase basicities of two standards.

In the *kinetic* method [84-86] a proton-bound dimer of the unknown and the standard compound is dissociated (e.g. by collision-induced dissociation). The proton is expected to be retained by the species with higher gas-phase basicity. The advantage of the *kinetic* method over the *deprotonation* method is that the activation barrier to proton transfer is lower and hence a more accurate determination of the gas-phase basicity can be achieved.

---

[4] The pKa value is defined as the pH value where [HA] = [A⁻] for the acid-base reaction HA $\leftrightarrow$ A⁻ + H⁺.

[5] The dielectric constant of a medium is unitless and is also known as the relative permittivity $\varepsilon_r = \varepsilon/\varepsilon_0$, where ε is the permittivity of the medium and ε_0 the permittivity of free space.

Much controversy has surrounded the gas-phase basicity of histidine, for which different values were determined by different workers [83,84] (which may be due to the use of different ionisation techniques). It is nowadays accepted that arginine has the highest gas-phase basicity, followed by histidine and lysine [87] (see Table A2.2, Appendix 2). While these accepted gas-phase basicities concern amino acids on their own, the gas-phase basicities of individual amino acids increase in larger peptides [84]. This is due to favourable intramolecular non-bonding interactions (see section 1.4.5) resulting in a stabilisation of charges [84], in analogy to water molecules stabilising charges in solution.

1.4.5 Non-bonding interactions

Non-bonding interactions (or intermolecular interactions) between molecules refer to through-space interactions between partial charges of polar molecules that do not involve covalent bonding. The repulsive part of these interactions prevents nuclei from collapsing into one another. The attractive contribution of non-bonding interactions determines three-dimensional folding of peptides and proteins. Intermolecular forces vary with respect to their typical potential stabilisation energies and the distance dependence of their interaction energies (see Table 1.1).

An intermolecular interaction that deserves a special mention is hydrogen bonding [88-90] since it plays a particularly important role in the determination of protein structure, which in turn affects the biological function of proteins. A hydrogen bond is defined as the intermolecular interaction between A-H...B, where A, B are electronegative elements such as N, O or F. The typical interaction energy for a hydrogen bond is 20 kJ mol^{-1} [91], which is relatively large compared to other intermolecular interactions. However, the strength of a hydrogen bond varies significantly with the system studied. Tareste *at al.* have deduced the strength of a hydrogen bond between a polar group and water molecule to be 0.5 kcal mol^{-1} [92], which they determined from the adhesion energy between monolayers of lipids and water. Huang *et al.* have calculated typical hydrogen bond strengths of complexes of biological interest in the range of 5-30 kcal mol^{-1} using quantum chemical calculations [93].

Interaction type	Distance dependence of potential energy	Typical energy $(kJ\ mol^{-1})$	Comment
Ion-ion	$1/r$	250	Between ions
Ion-dipole	$1/r^2$	15	
Dipole-dipole	$1/r^3$	2	Between stationary polar molecules
	$1/r^6$	0.6	Between rotating polar molecules
London (dispersion)	$1/r^6$	2	Between all types of molecules

Table 1.1 Typical stabilisation energies and distance dependencies of intermolecular interactions taken from [91].

In solution protein folding intermolecular hydrogen bonds between polar groups and water molecules are broken and replaced by intramolecular hydrogen bonds depending on the energy balance of these processes (difference between intramolecular interactions and the protein-solvent interactions, as mentioned before). Conversely, in the gas-phase (for fully de-hydrated peptide or protein ions) only intramolecular hydrogen bonding is present and hence the gas-phase *protein folding problem* is arguably easier to solve.

1.4.6 Techniques to study gas-phase structure

Relatively few techniques are available to study the structure of large ions in the gas-phase. The most common techniques include H/D exchange [94-96] (that gives information about surface availability of proteins), dissociation studies [97-99] (that give mainly primary structure information) and ion mobility [100,101] [102] (that gives information about the size of ions). Hydrogen bonding in the gas-phase has been probed by infrared spectroscopy, however, this has only been implemented on small molecular systems with very strong hydrogen bonding (e.g. BrHBr⁻.Ar

clusters) [103]. An excellent review of recent advances in the gas-phase structure of peptides and proteins has been published by Jarrold [81].

1.4.7 Computational modelling

Since the experimental techniques described above only give limited information about the gas-phase structure of ions, computational modelling of gas-phase structures is essential. In ion mobility mass spectrometry the experimental cross-section of a peptide is compared to calculated cross-sections for modelled peptide gas-phase structures in order to identify a structure [81]. Since peptides are large entities by the standards of quantum chemical models and the calculations are therefore computationally expensive, molecular mechanical models are generally employed for this modelling work.

1.5 Molecular mechanics

Molecular mechanics models are based on Newtonian physics and are used to simulate structures by calculating the strain energy of the molecule. The strain energy E_S is defined as the sum of the bond stretching, angle bending, torsion strain and non-bonding interaction energies (see Equation 1.9). These energies are dependent on the positions of the atoms. Each atom position represents a variable in a multi-dimensional potential energy surface (PES) and the task of the force field model is to determine potential energy minima.

$$E_S = E_{bond\ stretching} + E_{angle\ bending} + E_{torsion\ strain} + E_{non-bonding\ interactions} \qquad \textbf{Equation 1.9}$$

The temperature of a model system defines the kinetic energy of the atoms.[6] The potential and kinetic energy of the system can interchange in analogy to a ball on an inclined plane, and hence these models are referred to as (Newtonian) mechanical force field models. When during a simulation kinetic energy is taken out of the system (i.e. decreasing temperature) the molecule is forced to move to a minimum on

---

[6] The temperature expresses the kinetic energy (K.E.) of the atoms. K.E. = $\frac{1}{2}$ m v^2 = N_f k T/2, where m = mass of atom, v = velocity of atom, N_f = degrees of freedom (N_f = 3, for perfect gas), k = Boltzmann constant and T = temperature.

the PES. If this procedure is repeated a number of times different minima are found with different relative energies, which serve as potential candidate structures.

Various force field models have been developed including CHARMm [104], GROMOS [105], MM2 [106], MM3 [107] and Amber [108,109]. The Amber force field has been developed particularly for biological molecules such as proteins and deoxyribonucleic acid (DNA). Further developments of Amber have tried to improve the procedure for calculating atomic point charges [36]. These atom-centred point-charges are meant to accurately describe the electrostatic potential of the molecule at given multiples of the Van der Waals radii (1.4, 1.6, 1.8 and 2.0) in order to model non-bonding interactions (i.e. charge-charge, charge-dipole, dipole-dipole and dispersion interactions, as well as hydrogen bonding; see Table 1.1). Atomic charge fitting in Amber was derived using Gaussian 90 [110] and Gaussian 92 [111] using Hartree-Fock level theory and a 6-31G* basis set. The restrained electrostatic potential (RESP) charge fitting procedure involves a least-squares fitting of the atomic point charges (as determined by Gaussian) in order to minimise higher than expected charges on 'buried' atoms (e.g. sp^3 carbons) and to give the same charge to chemically equivalent atoms [112]. Atomic charges of the 20 naturally occurring amino acids with typical protonation sites that are observed in solution are included in the AMBER[7] suite of programs [36].

The strain energy (E_S) expression used in Amber is shown in Equation 1.10; the parameters and variables used are listed in Table 1.2. In order for force field modelling to accurately simulate molecular structures, the correct parameterisation of the force field is of crucial importance (i.e. garbage in, garbage out). The bond stretching and angle bending interactions are described by Hooke's law (in analogy to a spring). The equilibrium bond distance r_{eq} and equilibrium bond angle θ_{eq} are often obtained from X-ray crystallography data. The dihedral angle ϕ refers to the torsional rotation. Many of the force constants are obtained from quantum chemical calculations. The final term in this expression describes non-covalent interactions, such as the Van der Waals forces, hydrogen bonding and the electrostatic

---

[7] Note that AMBER refers to the software modelling package, whereas Amber refers to the mechanical force field used in this package.

interactions. The Van der Waals interaction is modelled using a Lennard-Jones potential with an attractive force proportional to R_{ij}^{-6} (R_{ij} = distance between atoms i and j) and a repulsive force proportional to R_{ij}^{-12}. The electrostatic interaction (i.e. charge-charge interaction) depends on the magnitude of the charges q_i and q_j and their distance R_{ij}.

$$E_S = E_{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 +$$

$$\sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}} \right]$$

Equation 1.10

Symbol	Description
K_r	Bond distance force constant
r	Bond distance
r_{eq}	Equilibrium bond distance
K_θ	Bond angle force constant
θ	Bond angle
θ_{eq}	Equilibrium bond angle
V_n	Torsion force constant
ϕ	Torsion angle
γ	Phase angle
A_{ij}	Repulsion term
B_{ij}	Dispersion term
R_{ij}	Distance between atom i and j
q_i	Charge on atom i
q_j	Charge on atom j
ε	Permittivity of medium

Table 1.2 Summary of terms used in Equation 1.10.

Since electrostatic interactions only decrease as a function of R_{ij}^{-1} it is the longest range intermolecular interaction (see also Table 1.1) and when modelling gas-phase structures this is particularly true, given the higher permittivity of free space[8]. Dipole moments have to be taken into account when calculating the electrostatic interactions. An alignment of the peptide dipoles, such as occurs in α-helices, results in a higher energy which is stabilised by hydrogen bonding between the amide oxygen (i.e. C=O) of one amino acid and the amide hydrogen (i.e. N-H) of another amino acid. Polyalanine (Ala_n) has one of the highest helical propensities (i.e. preference to form α-helices). Simulations by Vila *et al.* have shown that the propensity for Ala_n to form α-helices increases as the dielectric constant of the solvent is lowered [113], suggesting that the gas-phase is a suitable environment to study the secondary structure of peptides (i.e. folding motives such as α-helices or β-sheets). Modelling by Hudgins *et al.* on $Ac\text{-}Ala_n\text{-}LysH^+$ and $Ac\text{-}LysH^+\text{-}Ala_n$ revealed that the former adopts an α-helical gas-phase structure, whereas the latter gave a globular structure, which was confirmed by ion mobility measurements [114,115]. This shows that the accuracy of mechanical force fields has reached a stage where the secondary structure of gas-phase peptides can be predicted accurately.

Nevertheless, it is important to realise the limitations of force field models; it is not possible to change the configuration of the molecule during modelling or take into account electronic effects. This means that transition states cannot be determined. Another consequence of ignoring electronic effects is that the common force fields do not take into account polarisability of molecules[9]. Marten *et al.* have argued that there is a poor correlation between hydrogen bonding interactions modelled by classical electrostatic interactions (as used in force fields) as compared to hydrogen bonding interactions derived from quantum chemical calculations [116]. The use of polarisable force fields should allow for a more rigorous treatment of

---

[8] Dielectric constant of water is 80, which means that the permittivity of water is almost two orders of magnitude lower than for a vacuum.

[9] Polarisability of a molecule refers to the induced dipole moment generated in a molecule due to an applied electrostatic field.

hydrogen bonding interactions [117]. Accurate modelling of hydrogen bonds [93,117] can only be achieved by high-level *ab initio* calculations (MP2) [118] which take into account electron correlation [119], electrostatic, induction and dispersion effects [120]. Such high-level calculations are not yet applicable to larger systems such as polypeptides, for which force field modelling is the most appropriate approach. A comparison of various force fields with quantum chemical methods to describe hydrogen bonding in DNA (deoxyribonuclic acid) by Hobza *et al.* has shown that the best performance was exhibited by the Amber force field [121]. Moreover, calculations by Meyer *et al.* [122] using the B3LYP type exchange and correlation functionals [123] hybrid density functional method (DFT) [124] and Amber on intermolecularly hydrogen-bonded nucleic acid tetrads showed good agreement for strongly bonded complexes, but poor agreement for weakly hydrogen-bonded complexes.

1.6 Fragmentation of polypeptides

1.6.1 Nomenclature of backbone fragmentation

Fragmentation of polypeptides with the aim of determining primary structure information (i.e. sequence) plays an important part in protein identification [22]. This sequencing is achieved by electronically and/or vibrationally exciting the molecule in order to induce bond cleavage. Figure 1.9 shows the nomenclature for the peptide fragmentation pattern. Three distinct bonds in the backbone exist where cleavage could occur: the C_α-C bond, the peptide bond C-N and the N-C_α bond. Of these the peptide bond (or amide bond) is generally the weakest and in vibrational excitation techniques of which collisionally activated dissociation (CAD) [125] is one example, b and y fragments generally are most abundant. Conversely, in electron capture dissociation (ECD) [30] c and z fragments are generally the most abundant fragments. In order to contrast the different fragmentation mechanisms in vibrational excitation techniques and electron capture dissociation (ECD), the widely known mechanisms for both are introduced; vibrational excitation techniques – *mobile proton* model [126]; ECD – *hot hydrogen re-arrangement* mechanism [127].
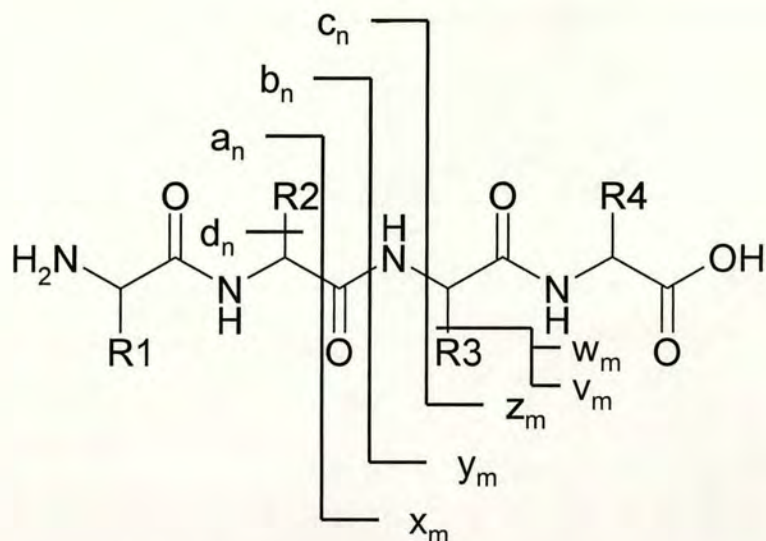
Figure 1.9 Nomenclature for polypeptide fragmentation based on Roepstorff and Fohlmann [128], where n is the number of residues from the N-terminus and m is the number of residues from the C-terminus.

1.6.2 Vibrational excitation

1.6.2.1 Mobile proton model

Fragmentation of peptides by vibrational excitation techniques, such as CAD [125], sustained off-resonance irradiation collisionally activated dissociation (SORI CAD) [71], surface-induced dissociation (SID) [129], infrared multi-photon dissociation (IRMPD) [130] or black-body infrared radiative dissociation (BIRD) [131] generally cause cleavage of the peptide bond, resulting in b and y fragment ions. The common feature in these techniques is that energy is transferred to the molecule in the form of collisions (CAD, SORI CAD, SID) or infrared absorption (IRMPD, BIRD), which causes the molecule to be in a vibrationally excited state prior to dissociation. McLafferty and co-workers refer to these techniques as the 'heating' techniques and they suggested that as the energy in the molecule increases one of the weaker bonds in the molecule dissociates. This would explain the high cleavage on the N-terminal side of proline, since the strain on the peptide chain due to proline makes the N-terminal peptide bond of proline particularly weak.

Wysocki and co-workers observed that the number of ionising protons relative to the number of basic residues in peptides where acidic residues are present

influences the fragmentation behaviour [132]. Conversely, when the number of ionising protons is higher than the number of arginine residues the cleavage is nonselective. This implies that the proton is involved in cleavage of the peptide bond and hence that 'charge directed' cleavage occurs. In order to test their hypothesis, they carried out fragmentation studies on model peptides using SID. SID works on the premise that an ion collides with an organic surface. In contrast to other fragmentation techniques, SID allows accurate deposition of energy onto the molecule with a narrow energy spread, since the kinetic energy of the ion can be specifically set.

Wysocki and co-workers showed fragmentation efficiency curves for the model peptides RPPGFSPF and PPGFSPFR in the 1+ and 2+ charge states [126]. The onset of fragmentation for the 1+ charge state for both peptides is equivalent. The preferred site of protonation is unequivovally on the arginine side-chain, since the gas-phase basicity of that site is highest [87]. This observation led to the development of the *mobile proton* model, which states that the energy required to effect fragmentation depends on the location of the proton; i.e. it is dependent on how much energy is required to make the proton mobile. The doubly protonated form of RPPGFSPF shows an earlier onset of fragmentation than does the corresponding 2+ charge state of PPGFSPFR. This is due to the fact that the second most likely protonation site for PPGFSPFR is on the proline at the N-terminus, which is a more basic site and therefore binds the proton more strongly than any of the protonation sites in RPPGFSPF, apart from arginine. A correlation exists between the gas-phase basicity of the protonated site and the energy required for fragmentation.

The fragmentation efficiency curves for the 2+ charge states of RRPPGFSPF and PPGFSPFRR show a similar onset energy of fragmentation. With respect to the *mobile proton* model, this implies that the proton is bound with the same 'tightness' in both cases, which basically means that both arginine side-chains are protonated. Given the closeness of the charge sites and ensuing Coulombic repulsion, one would expect the onset of fragmentation to be at a lower energy. The higher fragmentation onset for doubly protonated RRPPGFSPF (and PPGFSPFRR) relative to singly or doubly charged RPPGFSPF (or PPGFSPFR) can only be explained by the *mobile*

proton model and implies that Coulombic repulsion plays an insignificant role. Studies by Jockusch *et al.* [97] on the fragmentation of different charge states of bovine ubiquitin by BIRD also suggested higher activation energies for higher charge states, which runs contrary to the belief that Coulombic repulsion more readily gives rise to fragmentation.

1.6.2.2 Structure of fragment ions

The question as to where and how the proton initiates backbone cleavage is still controversial, especially with respect to the structure of the resulting b fragment ion. The proton could either move to the amide nitrogen or the carbonyl oxygen. Since the carbonyl oxygen is more basic, it would be logical to assume that this is where the proton attaches. Lehmann and co-workers have proposed that protonation of the backbone carbonyl would result in nucleophilic attack from the carbonyl oxygen N-terminal to the protonated carbonyl to give rise to an oxalozone b ion and a linear y fragment [133] (see Figure 1.10). This is consistent with the observation that b_1 ions are generally not observed, unless the N-terminus is acylated or carries a side-chain to generate a cyclic b_1 ion [134]. Further, the oxazolone structure is more basic than the corresponding y fragment linear peptide and would therefore retain the proton [135].
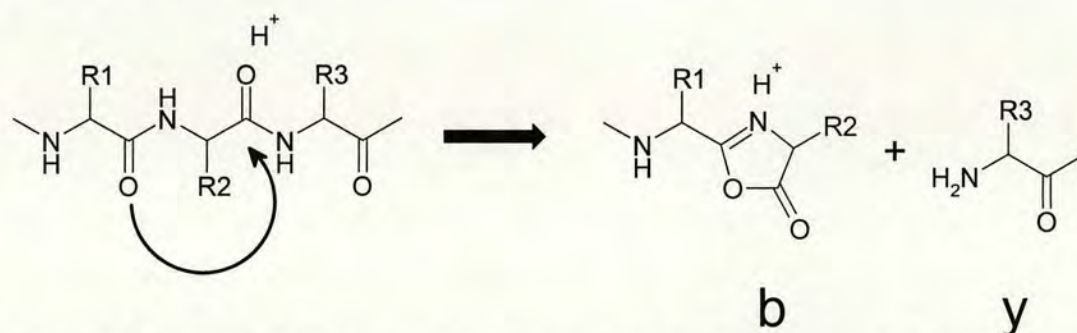
Figure 1.10 Reaction schematic for formation of oxazolone b ion and linear y ion following nucleophilic attack by N-terminal carbonyl oxygen [133].

Conversely, protonation at the amide nitrogen would lead to a lengthening of the C(O)-N bond [98,136] giving rise to an acylium b ion (see Figure 1.11). Haselmann *et al.* used electron capture dissociation on b^{2+} ions, which gave an abundant loss of CO, which they associated with the acylium structure [137]. These

results are in disagreement with *ab initio* calculations by Yalcin et al. [138] and Paizs et al. [139] which propose the oxazolone structure as the most energetically favourable structure for b ions.
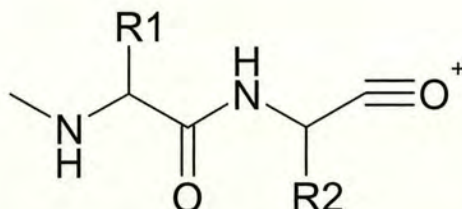
Figure 1.11 Structure of acylium b ion.

The fact that cleavage C-terminal to the proline residue does generally not occur, Schlosser and Lehmann have taken as evidence that the oxazolone structure is more likely [133]. In this case, due to the cyclic structure of proline, the oxazolone structure cannot be formed (see Figure 1.12).
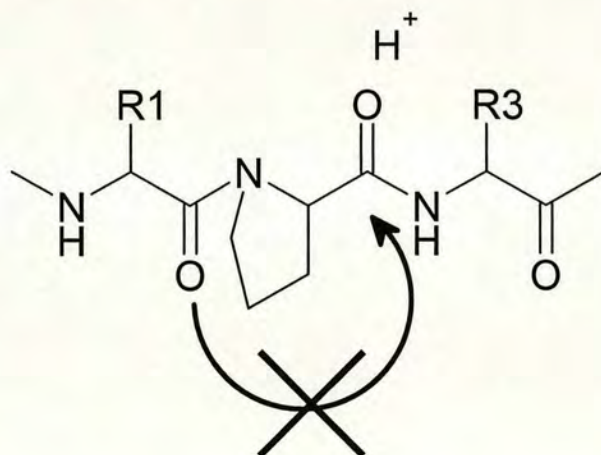
Figure 1.12 Reaction scheme showing that oxazolone cannot be formed due to the cyclic structure of proline. Adapted from [133].

The controversy surrounding the structure of b ions will probably not be resolved until a spectroscopic technique establishes its structure. The acylium structure was in fact the first one to be proposed [140] and it may be that both structures are possible, especially for larger fragments, where high level theory *ab initio* calculations are not yet possible. Paizs *et al.* evaluated the dissociation energy

required to cleave the peptide bond, based on the oxazolone structure, which they determined to be 74 kJ mol^{-1} (equivalent to 0.767 eV) [139]. This compares to onset fragmentation energies of 30-80 eV for the model peptides that Wysocki and co-workers studied [98]. The total available energy in the molecule far exceeds the energy required to break a peptide bond and therefore more dissociation pathways may be possible.

1.4.2.3 Relative fragmentation

Theoretical studies of SID and multiple collision activation collisionally induced dissociation (MCA-CID) [141] on small polyalanine peptides by Laskin *et al.* showed that both fragmentation techniques gave different abundances of fragment ions [142]. Due to the slower heating of the molecule in MCA-CID higher activation energy pathways, which are dominant in SID, have to compete against stepwise ion activations. As a result, the higher activation fragmentation channels are less likely to occur and low-energy dissociation channels predominate. Energy-resolved fragmentation efficiency curves also show that the relative abundance of fragments changes as a function of the mean collision energy.

Predicting the relative fragmentation behaviour in vibrational excitation techniques is therefore not a simple task, since it depends on a whole variety of parameters: secondary structure of the peptide [132], the availability of the *mobile proton* [98], the charge state of the precursor ion [98,143], the size of the ion [144], the excitation energy, the time-scale of excitation and the pre-excitation internal energy of the ion [142] (and hence also the ionisation technique). While certain enhanced cleavages are observed, such as on the N-terminal side proline and on the C-terminal side of aspartate and glutamate [145], the fragmentation in vibrational excitation techniques is in general too dynamic to be totally predictable.

1.4.3 Electron capture dissociation

1.4.3.1 Hot hydrogen re-arrangement mechanism

The dissociation mechanism in electron capture dissociation (ECD) in contrast to the vibrational excitation techniques is believed to be *non-ergodic* [30]; i.e. dissociation occurs prior to energy randomisation in the molecule. This implies

that the dissociation occurs on a very fast time-scale (< 1ps). Evidence supporting the *non-ergodic* mechanism of ECD are the fact that the peptide backbone N-C$_\alpha$ is broken preferentially in favour of more labile side-chain groups [31-34] and that even non-covalent intermolecular bonding is conserved in the presence of ECD [146]. *Ab initio* calculations by Carpenter have shown that the whole recombination energy must be contained in a small molecular region in order to account for N-C$_\alpha$ bond cleavage, suggesting *non-ergodic* behaviour [127].

The capture of an electron by a peptide cation leads to a reduction of the charge by one to give rise to the *reduced* odd-electron species $[M+nH]^{(n-1)+\bullet}$ (see Scheme 1). This implies that the charge of the protonated species needs to be at least 2+ in order to detect a charged reduced species.

$$[M+nH]^{n+} + e^- \rightarrow [M+nH]^{(n-1)+\bullet} \qquad \textbf{Scheme 1.1}$$
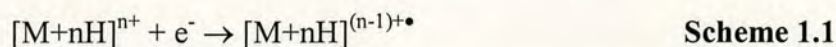
The capture of an electron is an exothermic process which is dependent on the recombination energy (see Figure 1.13) [147]. The total energy released upon electron capture is approximated to be on the order of 4-7 eV, depending mainly on the proton affinity (PA) of $[M+(n-1)H]^{(n-1)+}$, since the hydrogen atom affinity is generally small (< 1 eV). The energy gain is sufficient to cleave any bond in the molecule (typical bond strength 3-4 eV).

Figure 1.13 Thermodynamic cycle of ECD, showing recombination energy (RE), proton affinity (PA), ionisation energy of H atom (IE) and hydrogen atom affinity (HA). Adapted from [147].

McLafferty and co-workers postulated the *hot hydrogen re-arrangement* mechanism to explain how electron capture would give rise to backbo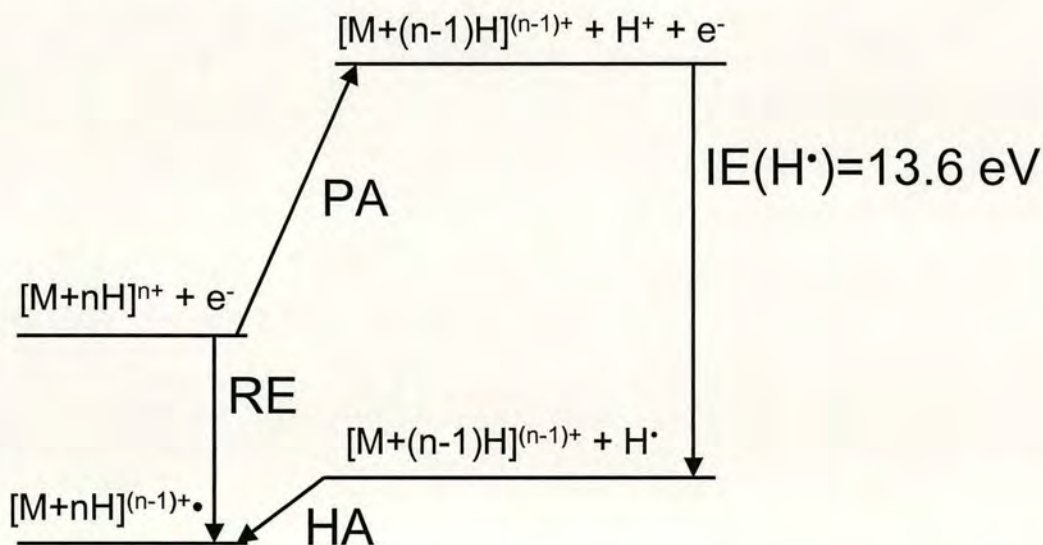ne dissociation [127]. The mechanism works on the premise that a peptide cation captures a slow electron (< 0.2 eV) at the site of protonation, giving rise to a 'hot hydrogen' $H^{\bullet}$ (~ 6 eV) [148]. This hydrogen radical can then be re-captured at a high hydrogen radical affinity site, such as a backbone carbonyl, to induce backbone alpha-cleavage (Figure 1.14) [30]. As can be seen in Figure 1.14 the c fragment ion contains an additional H, which explains the fact that it is 1 Da higher in mass than predicted from the residue masses. Kjeldsen *et al.* suggested a new notation for these fragments, c' and $z^{\bullet}$ [149], which will be used throughout this thesis to avoid confusion. Another piece of evidence in favour of the hydrogen radical transfer is the loss of 1 Da from the reduced species, which can only be explained by the loss of a hydrogen radical (hydrogen atom desorption). The preferential cleavage in ECD of the $N-C_{\alpha}$ bond resulting in c' and $z^{\bullet}$ ions also explains why no ECD cleavage is observed on the N-terminal side of proline.
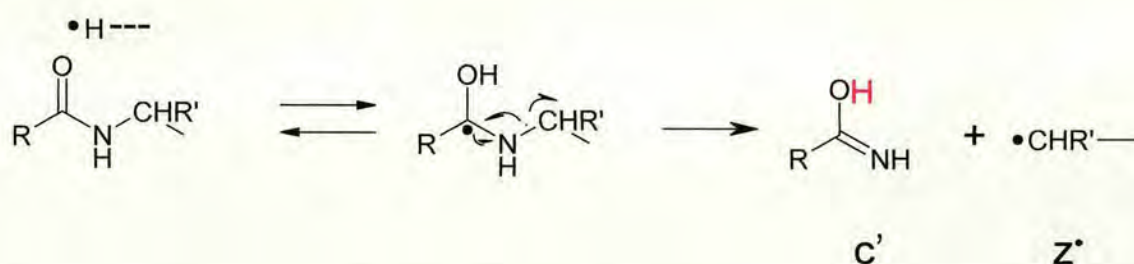
Figure 1.14 Schematic showing *hot hydrogen re-arrangement* mechanism with subsequent radical backbone alpha-cleavage resulting in c' and $z^{\bullet}$ fragment ions. Additional H on c fragment marked in red (H) explaining the fact that c fragment 1 Da higher in mass than predicted from residue masses of amino acids. Adapted from [127] and [149].

McLafferty and co-workers also observed high cleavage efficiencies of cysteine bridges (i.e. disulphide bridges) (see Scheme 1.2) [127]. *Ab initio* calculations gave higher hydrogen radical affinities for CH_3SSCH_3 than $CH_3CONHCH_3$, supporting the 'hot hydrogen' re-capture hypothesis.
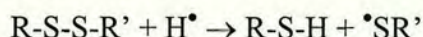
$$R\text{-}S\text{-}S\text{-}R' + H^\bullet \rightarrow R\text{-}S\text{-}H + {}^\bullet SR'$$ **Scheme 1.2**

More minor fragment ions that are observed in ECD include a$^\bullet$ and y ions (see Figure 1.15), which can be rationalised by hydrogen atom transfer to an amide nitrogen atom. Incidentally, hydrogen atom transfer to a nitrogen atom on the backbone, instead of a carbonyl oxygen, has been put forward by Zubarev and co-workers as a alternative mechanism to explain N-C$_\alpha$ bond cleavage [150].
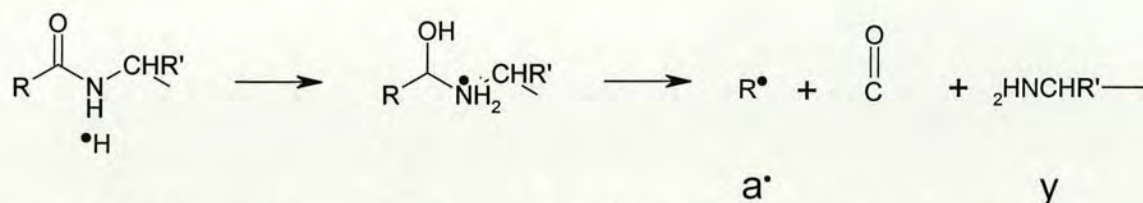
Figure 1.15 Proposed reaction schematic for a$^\bullet$ and y ion formation in ECD [127].

Other unusual lower abundance fragments, such as c$^\bullet$ and z' have been been observed by a number of groups [32,127,148,151,152], and have been rationalised by hydrogen atom transfer to a C$_\alpha$ [148].

Another phenomenon in ECD are neutral losses from the reduced species that have been linked to cleavages of particular amino acid side-chains [153,154]. Such neutral losses can give information about the presence of certain amino acids, which is particularly useful in cases where the sequencing is incomplete. This information can also be useful when there is ambiguity about N or C-terminal fragments.

A review of the understanding of the ECD mechanism has been given by Zubarev and co-workers [150] in a historical context of related fragmentation techniques such as dissociative recombination [155], neutralisation-reionisation [156] and electron impact excitation of ions from organics (EIEIO) [157]. Many of the assumptions made in the *hot hydrogen re-arrangement* mechanism are based on empirical evidence from observed fragments and *ab initio* calculations. However, no direct empirical evidence exists for the *hot hydrogen re-arrangement* mechanism, which may be difficult to obtain if the mechanism is as fast as assumed.

1.4.3.2 Other ECD mechanisms

It is likely that other reaction pathways compete against the *hot hydrogen re-arrangement* mechanism in ECD. The cleavage at disulphide bridges for example has also been rationalised by direct dissociative electron capture [127]. Hudgins *et al.* [158] have shown that fully deuterated peptides and sodiated peptides give rise to backbone cleavage upon electron capture. While this observation does not disprove the *hot hydrogen re-arrangement* mechanism, it opens up the possibility of other mechanisms. A comparison of doubly protonated versus doubly sodiated polyglycol ions showed less fragmentation for the sodiated species [159], which is consistent with a lower recombination energy for sodiated ions than for protonated ions.

O'Connor and co-workers have shown that ECD for cyclic peptides gave rise to amino acid losses, which can only be explained by two backbone cleavages to have taken place upon one electron capture [160]. Two backbone cleavages are also thought to give rise to a$^\bullet$ and y fragment ions (see Figure 1.15). It remains to be seen whether two (or more) backbone cleavages are a dominant process in ECD or whether this is a low abundance dissociation channel.

The *non-ergodic* mechanism in ECD is subject of some controversy. Since an FT-ICR MS experiment is much too long (100's ms – s) to establish a kinetic profile of the reaction, ECD has been investigated by quantum mechanical molecular modelling. Uggerud and co-workers have shown using *ab initio* calculations for model pepides that a fast ECD mechanism ($< 10^{-12}$ s) may apply for a minority of ions (~10%), while the majority of ions are subject to a much slower mechanism (~10^{-5} s) [161]. Their calculations also indicate that disulphide bond cleavage proceeds without an activation barrier in the presence of a hydrogen atom [161]. *Ab initio* calculations by Tureček and co-workers have also shown that N-C$_\alpha$ bond cleavage may not have to be non-ergodic and their predicted unimolecular rate constant is ~10^5 s^{-1} [162], which is in agreement with Uggerud and co-workers (see above). Further, electron location on the backbone carbonyl carbon (see Figure 1.14) causes a weakening of the N-C$_\alpha$ bond [162], which has also been confirmed by *ab initio* calculations by Konishi *et al* [163]. Calculations on protonated GlySer gave an estimated recombination energy of electron capture of 66.66 kcal mol^{-1} (278.91 KJ

mol^{-1}, 2.89 eV) through Franck-Condon factor (i.e. no atomic movement), whereas the N-C$_\alpha$ bond cleavage transition state was estimated at merely 6.6 kcal mol^{-1} (27.61 KJ mol^{-1}, 0.286 eV) [163]. In comparison, a typical N-C bond energy in the ground state is approximately 300 kJ mol^{-1} ($\sim$ 3 eV).

In these *ab initio* calculations, hydrogen bonding between the protonated site (e.g. basic site on amino acid side-chain of arginine, histidine or lysine) and a backbone carbonyl oxygen is thought to play an important part in explaining the hydrogen atom transfer [163]. In fact, *ab initio* calculations by Tureček and co-workers have predicted unimolecular rate constants for this transfer as high as 10^{12} s^{-1} [164]. Although, their current belief is that the electron moves along the chain to the carbonyl carbon, whereas the proton is transferred separately from the protonated site to the backbone carbonyl oxygen [165]. This separate electron and proton transfer, resulting in effective hydrogen atom transfer, was first suggested by Siegbahn et al. [166].

While *ab initio* calculations may be able to explain the ECD mechanism in terms of an ergodic mechanism, all quantum mechanical calculations suffer from the limitation of calculating energy levels at ground electronic states. When the electron is initially captured, this may occur in a high-lying electronic Rydberg state, in analogy to dissociative recombination [167], which is the current view of McLafferty and co-workers. Nevertheless, how the electron relaxes from this Rydberg state to lower electronic states and how the electronic recombination energy is transferred into nuclear motion to induce backbone cleavage cannot be accurately modelled even by state-of-the-art quantum mechanical calculations. Hence, the controversy about mechanisms in ECD is likely to go on for some time.

1.4.3.3 Other electron-mediated dissociation techniques

By varying the kinetic energy of the electrons it is possible to explore other ion-electron reactions, which have been pioneered largely by Zubarev and co-workers. The electron capture cross-section is maximised for low-energy electrons (< 0.2 eV), but has a local maximum around 7 eV [149]. Such a kinetic energy regime for the electrons has been termed as 'hot electron' capture dissociation (HECD) [149]. Electron excitation dissociation (EED) [168] has been proposed for singly-

charged species and electron detachment dissociation (EDD) [169] has been applied to multiply deprotonated peptides (peptide anions). The applicability of these electron-mediated dissociation techniques for polypeptides, the required electron kinetic energy and their general fragmentation behaviour are summarised in Table 1.3. Reactions of polypeptide ions with electrons in the gas-phase have also recently been reviewed by Zubarev [170].

Name	Abbreviation	Ions	e⁻ kinetic energy (eV)	Fragments
Electron Capture Dissociation	**ECD**	$\geq 2+$	< 0.2	c' and $z^\bullet$
Hot Electron Capture Dissociation	**HECD**	$\geq 2+$	7-10	All fragments
Electron Detachment Dissociation	**EDD**	$\geq 2-$	> 20	a, $a^\bullet$, x, $x^\bullet$, c' and $z^\bullet$
Electron Excitation Dissociation	**EED**	$1+$	Fast and slow e⁻	a, c' and $z^\bullet$

Table 1.3 Table summarising electron-mediated dissociation techniques for polypeptides proposed by Zubarev and co-workers, adapted from [171].

1.4.3.4 ECD of proteins

The 'top-down' approach of sequencing whole proteins has been proposed by McLafferty and co-workers as a more efficient way to identify proteins [172,173]. The common 'bottom-up' approach involves enzymatic digestion of the protein followed by mass spectrometric analysis [22]. The disadvantage with the 'bottom-up' approach is that one cannot be sure how large the whole protein was before digestion. If a peptide fragment carrying a post-translational modification is not detected in the enzymatic digest mass spectrum, this information is lost. Further, tandem mass spectrometry of whole proteins requires very little sample ($< 10^{-17}$ mol) [47] and is therefore very sensitive.

ECD of peptides generally gives different sequence information to vibrational excitation techniques, and hence both sequencing techniques can be highly complementary [151]. Given the fact that b and c fragmentation at the same residue gives fragments that are 17.03 Da different in mass, whereas y and $z^\bullet$ fragmentation at the same residue gives fragments that are 16.02 Da different in mass, it would thus be possible to get information on which fragments are N or C-terminal. This is of particular interest in *de novo* sequencing and hence this approach has been termed the 'golden complementary set' for sequencing [174].

The challenge for ECD to sequence proteins is that non-covalent interactions may hold the molecule together, thus yielding no sequence information. In order to break non-covalent interactions McLafferty and co-workers used vibrational excitation prior to ECD, which they termed 'in beam' or 'activated-ion' ECD (AI ECD) [175].

A very interesting application of ECD has recently been pursued by McLafferty and co-workers to investigate the folding of protein ions. Since more extensive non-covalent bonding in the molecule results in fewer ECD fragment ions, the temperature-dependent folding of protein ions can be studied [99,176]. The extent of hydrogen atom desorption has also been shown to be linked to how 'open' the gas-phase structure of the protein is [177]. Nevertheless, the higher hydrogen atom desorption for a higher charge state (13+ charge state as opposed to 7+ charge state for bovine ubiquitin, see [177]) could also be due to a higher recombination energy, due to a decreased proton affinity for the higher charge state, as argued by Zubarev [170]. Hence, these results may be open to multiple interpretation and more fundamental work needs to be carried out before using ECD as a tool to investigate gas-phase structure of ions.

Breuker and McLafferty have recently observed ECD-type c and y fragments in electrospray ionisation spectra of cytochrome c. This protein carries a heme group with iron (III) and since no electrons were used to effect electron capture, the authors have postulated that an electron transfer takes place in the protein, resulting in ECD. Breuker and McLafferty have named this phenomenon 'native electron capture dissociation' (NECD) [178]. The relative yield of c and y fragments was found to

increase with increasing cytochrome c concentration and decreasing pH, which Breuker and McLafferty have rationalised by dimer formation and asymmetric charge partitioning between the dimer parts, as observed by Jurchen and Williams [179]. Breuker and McLafferty believe that the charge asymmetry on the dimer may be the driving force for the proposed electron transfer, resulting in NECD.

1.4.3.5 Relative fragmentation

ECD is believed to give more extensive sequence coverage than vibrational excitation techniques [148,180]. While higher cleavage efficiencies have been reported at the C-terminal side of tryptophan and lower cleavage at the N-terminal side of cysteine, the cleavage efficiencies in ECD have been found to be more homogeneous than for CAD [180]. Budnik *et al.* have proposed that ECD is a more 'reducible' fragmentation technique, which means that the cleavage efficiency is related to the immediate chemical vicinity and not to the rest of the molecule [181]. In practical terms, this implies that the relative cleavage efficiency in ECD does not depend on the molecular size of the peptide, the charge state, the basicities of the protonated sites or the gas-phase conformation of the peptide (i.e. gas-phase structure).

This assumption only holds true in the absence of hydrogen bonding. In fact, as mentioned before, *ab initio* calculations by Tureček and co-workers have indicated that a one-step transfer of the hydrogen atom to a backbone carbonyl would be preferred over the two-step process (hydrogen atom desorption followed by re-capture) as proposed by McLafferty [164]. The predicted rate constants for this one-step process were as high as 10^{12} s^{-1}. This implies that hydrogen bonding of the protonated site to particular backbone carbonyls would result in enhanced transfer. One of the primary aims of this thesis is to investigate the relative cleavage efficiency in ECD and to compare this to structures modelled using the Amber force field model.

References

(1)   Venter, C. In *219th American Chemical Society National Meeting*; American Chemical Society: San Francisco, CA, 2000.

(2)   Xu, Y. *Cell Death and Diff.* **2003**, *10*, 400-403.

(3)   Dove, A. *Nature Biotech.* **1999**, *17*, 233-236.

(4)   Aebersold, R.; Gygi, S. P.; Grithin, T. J.; Han, D. K. M.; Yelle, M. J. *Am. Gen. Proteom. Technol.* **2001**, *1*, 22, 24, 24-27.

(5)   Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198-207.

(6)   Bleakney, W. *Phys. Rev.* **1929**, *34*, 157.

(7)   Beckey, H. D. In *International series in analytical chemistry*; Belcher, R., Frieser, H., Eds.; Pergamon Press: Oxford, 1977; Vol. 62.

(8)   Torgerson, D. F.; Skrowronski, R. P.; Macfarlane, R. D. *Biochem. Biophys. Res. Commun. Mass Spectrom.* **1974**, *60*, 616.

(9)   Benninghoven, A.; Sichtermann, W. K. *Anal. Chem.* **1978**, *50*, 1180.

(10)   Barber, M.; Bordoli, R. S.; Elliott, G. J.; Sedgwick, R. D.; Tyler, A. N. *Anal. Chem.* **1982**, *54*, 645.

(11)   Posthumus, M. A.; Kistemaker, P. G.; Meuzelaar, H. C. L.; Ten Nuever de Brauw, M. C. *Anal. Chem.* **1978**, *50*, 985.

(12)   Karas, M.; Bachmann, D.; Hillenkamp, F. *Anal. Chem.* **1985**, *57*, 2935.

(13)   Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151.

(14)   Yamashita, M.; Fenn, J. B. *J. Am. Chem. Soc.* **1984**, *88*, 4451-4459.

(15)   Pandey, A.; Mann, M. *Nature* **2000**, *405*, 837-846.

(16)   Wilkins, M. R.; Williams, K. L.; Apple, R. D.; Hochstrasser, D. F. In *Proteome Research: New Frontiers in Functional Genomics*; Springer: Berlin, 1997; pp 1-243.

(17)   Anderson, N. G.; Anderson, N. L. *Electrophoresis* **1996**, *17*, 443-453.

(18)   Henzel, W. J.; Billeci, T. M.; Stuls, J. T.; Wong, S. C. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5011-5015.

(19)   Henzel, W. J.; Watanabe, C.; Stuls, J. T. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 931-942.

(20)   Jensen, O. N.; Mortensen, P.; Vorm, O.; Mann, M. *Anal. Chem.* **1997**, *69*, 1706-1714.

(21)   Berndt, P.; Hobohm, U.; Langen, H. *Electrophoresis* **1999**, *20*, 3521-3526.

(22)   Chait, B. T.; Wang, R.; Beavis, R. C.; Kent, S. B. *Science (Washington, D. C.)* **1993**, *262*, 89-92.

(23)   Kinter, M.; Sherman, N. E. *Protein Sequencing and Identification Using Tandem Mass Spectrometry*; John Wiley and Sons: New York, 2000.

(24)   Jacobowitz, O.; Iyengar, R. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10630-10634.

(25)   Fassina, G.; Verdoliva, A.; Palombo, G.; Ruvo, M.; Cassani, G. *J. Mol. Recogn.* **1998**, *11*, 128-133.

(26)   Jensen, P. K.; Pasa-Tolic, L.; Anderson, G. A.; Horner, J. A.; Lipton, M. S.; Bruce, J. E.; Smith, R. D. *Anal. Chem.* **1999**, *71*, 2076-2084.

(27)   Mørtz, E.; O'Connor, P. B.; Roepstorff, P.; Kelleher, N. L.; Wood, T. D.; McLafferty Fred, W.; Mann, M. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8264-8267.

(28)   Li, W.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Anal. Chem.* **1999**, *71*, 4397-4402.

(29)   Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001**, *101*, 269-295.

(30)   Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

(31)   Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250-4253.

(32)   Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.

(33)   Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.

(34)   Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.

(35)   Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K., M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R.; Duan, Y.; Pitera, J.; Massova, I. G.; Seibel, G. L.; Singh, U. C.; Weiner, P. J.; Kollmann, P. A. In;; University of California: San Francisco.

(36) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K., M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollmann, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

(37) Lawrence, E. O.; Livingston, M. S. *Phys. Rev. Lett.* **1932**, *40*, 555.

(38) Hipple, J. A.; Sommers, H.; Thomas, H. A. *Phys. Rev.* **1949**, *76*, 1877.

(39) Sommers, H.; Thomas, H. A.; Hipple, J. A. *Phys. Rev.* **1951**, *82*, 697.

(40) Wobschall, D.; Graham, J. R. J.; Malone, D. P. *Phys. Rev.* **1963**, *131*.

(41) Llewellyn, P. M. In;: USA, 1969.

(42) Baldeschwieler, J. D. *Science (Washington, D. C.)* **1968**, *159*, 263.

(43) McIver, R. T. *Rev. Sci. Instrum.* **1970**, *41*, 555.

(44) Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *25*, 282.

(45) Henry, K. D.; Williams, E. R.; Wang, B.-H.; McLafferty, F. W.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. USA* **1986**, *86*, 9075-9078.

(46) Hettich, R. L.; Buchanan, M. V. *Int. J. Mass Spectrom. Ion Proc.* **1991**, *111*, 365-380.

(47) Valaskovic, G. A.; Kelleher, N. L.; McLafferty Fred, W. *Science (Washington, D. C.)* **1996**, *273*, 1199-1202.

(48) Emmett, M. R.; White, F. M.; Hendrickson, C. L.; Shi, S. D. H.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 333-340.

(49) Bruce, J. E.; Anderson, G. A.; Hofstadler, S. A.; Winger, B. E.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 700.

(50) Quencer, T. L.; Emmett, M. R.; Hendrickson, C. L.; Kelly, P. H.; Marshall, A. G. *Anal. Chem.* **2001**, *73*, 1721-1725.

(51) Belov, M. E.; Gorshkov, M. V.; Udseth, H. R.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 2271-2279.

(52) Marshall, A. G.; Wang, T. C.; Ricca, T. L. *J. Am. Chem. Soc.* **1985**, *107*, 7893-7897.

(53) Gorshkov, M.; Pasa-Tolic, L.; Bruce, J. E.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **1997**, *69*, 1307.

(54) Bruce, J. E.; Anderson, G. A.; Hofstadler, S. A.; Van Orden, S. L.; Sherman, M. S.; Rockwood, A. L.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 914.

(55) Hendrickson, C. L.; Laude, D. A. J. *Anal. Chem.* **1995**, *67*, 1717.

(56)   Pitsenberger, C. C.; Easterling, M. L.; Amster, I. J. *Anal. Chem.* **1996**, *68*, 3732.

(57)   Hofstadler, S. A.; Swanch, F. D.; Gale, D. C.; Ewing, A. G.; Smith, R. D. *Anal. Chem.* **1995**, *67*, 1477.

(58)   Marshall, A. G. *Int. J. Mass Spectrom.* **2000**, *200*, 331-356.

(59)   Pardue, H. L. In *Analyt. Chim. Acta*, 1985; Vol. 178, p 158 pp.

(60)   Comisarow, M. B.; Nibbering, N. M. M. In *Int. J. Mass Spectrom. Ion Proc.*, 1986; Vol. 72, p 222 pp.

(61)   Marshall, A. G.; Verdun, F. R. *Fourier Transforms in NMR, Optical, and Mass Spectrometry: A User's Handbook*; First Edition ed.; Elsevier: Amsterdam, 1990.

(62)   Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.

(63)   Marshall, A. G.; Guan, S. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1819-1823.

(64)   Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325-1337.

(65)   Ledford, E. B.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **1984**, *56*, 2744-2748.

(66)   Stults, J. T. *Anal. Chem.* **1997**, *69*, 1815-1819.

(67)   Kelleher, N. L.; Senko, M. W.; Siegel, M. M.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 380.

(68)   Bruce, J. E.; Anderson, G. A.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1998**, *70*, 519-525.

(69)   Dahl, D. A.; Delmore, J. E.; Appelhans, A. D. *Rev. Sci. Instrum.* **1990**, *61*, 607-609.

(70)   Dahl, D. A. *Int. J. Mass Spectrom.* **2000**, *200*, 3-25.

(71)   Senko, M. W.; Speir, J. P.; McLafferty Fred, W. *Anal. Chem.* **1994**, *66*, 2801-2808.

(72)   Klassen, J. S.; Blades, A. T.; Kebarle, P. *J. Phys. Chem.* **1995**, *99*, 15509.

(73)   Branden, C.; Tooze, J. *Introduction to Protein Structure*; 2nd Ed. ed.; Garland Publishing Inc.: New York, 1999.

(74)   Smith, R. D.; Light-Wahl, K. J.; Winger, B. E.; Loo, J. A. *Org. Mass Spectrom.* **1992**, *27*, 811.

(75)   Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys.* **1968**, *49*, 2240.

(76)   Mach, L. L.; Kralik, P.; Rheude, A.; Dole, M. *J. Chem. Phys.* **1970**, *52*, 4977.

(77)   Iribarne, J. V.; Thompson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287.

(78)   Fenn, J. B. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 524.

(79)   Nohmi, T.; Fenn, J. B. *J. Am. Chem. Soc.* **1992**, *114*, 3241.

(80)   Smith, J. N.; Flagan, R. C.; Beauchamp, J. L. *J. Phys. Chem. A* **2002**, *106*, 9957-9967.

(81)   Jarrold, M. F. *Ann. Rev. Phys. Chem.* **2000**, *51*, 179-207.

(82)   Fulscher, M. P.; Mehler, E. L. *Chem. Phys.* **1996**, *204*, 403-410.

(83)   Gorman, G. S.; Speir, J. P.; Turner, C. A.; Amster, I. J. *J. Am. Chem. Soc.* **1992**, *114*, 3986-3988.

(84)   Carr, S. R.; Cassady, C. J. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 1203-1210.

(85)   Wu, Z.; Fenselau, C. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 777-780.

(86)   Boejesen, G.; Breindahl, T. J. *Journal of the Chemical Society Perkin Translations* **1994**, *2*, 1029-1037.

(87)   Harrison, A. G. *Mass Spectrom. Rev.* **1997**, *16*, 201-217.

(88)   Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: New York, 1997.

(89)   Scheiner, S. *Hydrogen Bonding: A Theoretical Perspective*; Oxford University Press: New York, 1997.

(90)   Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond*; Oxford University Press: Oxford, 1999.

(91)   Atkins, P. W. *Physical Chemistry*; 6th Edition ed.; Oxford University Press: Oxford, 1998.

(92)   Tareste, D.; Pincet, F.; Perez, E.; Rickling, S.; Mioskowski, C.; Lebeau, L. *Biophys. Chem.* **2002**, *83*, 3675-3681.

(93)   Huang, N.; MacKerell, A. D. J. *J. Phys. Chem.* **2002**, *106*, 7820-7827.

(94)   Wood, T. D.; Chorush, R. A.; Wampler, F. M.; Little, D. P.; O'Connor, P. B.; McLafferty, F. W. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2451.

(95)   McLafferty, F. W.; Guan, Z.; Haupts, U.; Wood, T. D.; Kelleher, N. L. *J. Am. Chem. Soc.* **1998**, *120*, 4732.

(96)   Valentine, S. J.; Clemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 3558.

(97)   Jockusch, R. A.; Schnier, P. D.; Price, W. D.; Strittmatter, E. F.; Demirev, P.; Williams, E. R. *Anal. Chem.* **1997**, *69*, 1119-1126.

(98)   Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.

(99)   Breuker, K.; Oh, H.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124*, 6407-6420.

(100)   Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 954.

(101)   Mao, Y.; Woenckhaus, J.; Kolafa, J.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1999**, *121*, 2712-2721.

(102)   Wyttenbach, T.; Bowers, M. T. *Top. Curr. Chem.* **2003**, *225*, 207-232.

(103)   Pivonka, N. L.; Kaposta, C.; Brummer, M.; von Helden, G.; Meijer, G.; Woste, L.; Neumark, D. M.; Asmis, K. R. *J. Chem. Phys.* **2003**, *118*, 5275-5278.

(104)   Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; Slater, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187-217.

(105)   van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulations (GROMOS) Library Manual*; Biomos: Groningen, 1987.

(106)   Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127-8134.

(107)   Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551-8582.

(108)   Weiner, S. J.; Kollmann, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta Jr., S.; Weiner, P. J. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.

(109)   Weiner, S. J.; Kollmann, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230-252.

(110)   Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Schlegel, H. B.; Raghavachari, K.; Robb, M. A.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seger, R.; Melius, C. F.; Baker, J.; Martin, L. R.; Kahn, L. R.; Stewart, J. J. P.; Topiol, S.; Pople, J. *Gaussian 90*; Gaussian Inc.: Pittsburgh, PA, 1990.

(111)   Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.;

Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. *Gaussian 92*; Gaussian Inc.: Pittsburgh, PA, 1992.

(112) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollmann, P. A. *J. Phys. Chem.* **1993**, *97*, 10269-10280.

(113) Vila, J. A.; Ripoll, D. R.; Scheraga, H. A. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13075-13079.

(114) Hudgins, R. R.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1998**, *120*, 12974-12975.

(115) Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1999**, *121*, 3494-3501.

(116) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 11775-11788.

(117) Kim, K.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 12952-12961.

(118) Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(119) Chalasinski, G.; Szczesniak, M. M. *Chem. Rev.* **2000**, *100*, 4227.

(120) Krishnan, R.; Pople, J. A. *Int. J. Quantum Chem.* **1978**, *14*, 91.

(121) Hobza, P.; Hubálek, F.; Kabelác, M.; Mejzlik, P.; Šponer, J.; Vondrášek, J. *Chem. Phys. Lett.* **1996**, *257*, 31-35.

(122) Meyer, M.; Schneider, C.; Brandl, M.; Suhnel, J. *J. Phys. Chem. A* **2001**, *105*, 11560-11573.

(123) Becke, A. J. *J. Chem. Phys.* **1993**, *98*, 5648-5652.

(124) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864-872.

(125) Loo, J. A.; Udseth, H. R.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 207.

(126) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399-1406.

(127) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn David, M.; Carpenter Barry, K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.

(128) Roepstorff, P.; Fohlmann, J. *J. Biomed. Mass Spectrom.* **1984**, *11*, 601.

(129) Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *J. Mass Spectrom.* **1996**, *31*, 339-350.

(130)   Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

(131)   Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859-866.

(132)   Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 5142-5154.

(133)   Schlosser, A.; Lehmann, W. D. *J. Mass Spectrom.* **2000**, *35*, 1382.

(134)   Tu, Y.-P.; Harrison, A. G. *Rapid Commun. Mass Spectrom.* **1998**, *12*.

(135)   Polce, M. J.; Ren, D.; Wesdemiotis, C. *J. Mass Spectrom.* **2000**, *35*, 1391-1398.

(136)   Somogyi, A.; Wysocki, V. H.; Mayer, I. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 704-717.

(137)   Haselmann, K. F.; Budnik, B. A.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 2242-2246.

(138)   Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **1995**, *7*, 233.

(139)   Paizs, B.; Lendvay, G.; Vekey, K.; Suhai, S. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 525.

(140)   McLafferty, F. W.; Turecek, F. *Interpretation of Mass Spectra*; 4th ed.; University Science Books: Sausalito, CA, 1993.

(141)   McLuckey, S. A.; Goeringer, D. E. *J. Mass Spectrom.* **1997**, *32*, 461.

(142)   Laskin, J.; Densiov, E.; Futrell, J. H. *Int. J. Mass Spectrom.* **2002**, *219*, 189-201.

(143)   Tang, X.-J.; Thibault, P.; Boyd, R. K. *Anal. Chem.* **1993**, *65*, 2824-2834.

(144)   Downard, K. M.; Biemann, K. *Int. J. Mass Spectrom. Ion Proc.* **1995**, *148*, 191-202.

(145)   Yu, W.; Vath, J. E.; Huberty, M. C.; Martin, S. A. *Anal. Chem.* **1993**, *65*, 3015.

(146)   Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Anal. Chem.* **2001**, *73*, 2998-3005.

(147)   Zubarev, R. A. In *Dissociative Recombination, Theory, Experiments and Applications*; Larsson, M., Mitchell, J. B. A., Schneider, I. F., Eds.; World Scientific: Singapore, 2000; Vol. IV, pp 214-217.

(148)   Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563-573.

(149)   Kjeldsen, F.; Budnik, B. A.; Haselmann Kim, F.; Jensen, F.; Zubarev, R. A. *Chem. Phys. Lett.* **2002**, *356*, 201-206.

(150)   Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. *Eur. J. Mass Spectrom.* **2002**, *8*, 337-349.

(151)   Kruger, N. A.; Zubarev, R. A.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom.* **1999**, *185,186,187*, 787-793.

(152)   Hakansson, K.; Emmett, M. R.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2001**, *73*, 3605-3610.

(153)   Cooper, H. J.; Hudgins, R. R.; Hakansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 241-249.

(154)   Haselmann, K. F.; Budnik, B. A.; Kjeldsen, F.; Polfer, N. C.; Zubarev, R. A. *Eur. J. Mass Spectrom.* **2002**, *8*, 461-469.

(155)   Bates, D. R. *Phys. Rev.* **1950**, *77*, 718.

(156)   Wesdemiotis, C.; McLafferty, F. W. *Chem. Rev.* **1987**, *87*, 495.

(157)   Cody, R. B.; Freiser, B. S. *Anal. Chem.* **1979**, *51*, 547.

(158)   Hudgins, R. R.; Hakansson, K.; Quinn, J. P.; Hendrickson, C. L.; Marshall, A. G. In *50th ASMS Conference on Mass Spectrometry and Allied Topics*: Orlando, FL, USA, 2002.

(159)   Cerda, B. A.; Horn, D. M.; Breuker, K.; Carpenter, B. K.; McLafferty, F. W. *Eur. Mass Spectrom.* **1999**, *5*, 335-338.

(160)   Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8949-8958.

(161)   Bakken, V.; Helgaker, T.; Uggerud, E. In *16th IMSC*: Edinburgh, UK, 2003.

(162)   Turecek, F. *J. Am. Chem. Soc.* **2003**, *125*, 5954-5963.

(163)   Konishi, H.; Yokotake, Y.; Ishibashi, T. *J. Mass Spectrom. Soc. Jpn.* **2002**, *50*, 229-232.

(164)   Turecek, F.; Syrstad, E. A. *J. Am. Chem. Soc.* **2003**, *125*, 3353-3369.

(165)   Turecek, F.; Syrstad, E. A.; Seymour, J. L.; Chen, X. In *16th IMSC*: Edinburgh, UK, 2003.

(166)   Siegbahn, P. E. M.; Blomberg, M. R. A.; Crabtree, R. H. *Theor. Chem. Acc.* **1997**, *97*, 289-300.

(167)   Datz, S. *J. Phys. Chem. A* **2001**, *105*, 2369-2373.

(168)   Nielsen, M. L.; Budnik, B. A.; Haselmann, K. F.; Olsen, J. V.; Zubarev, R. A. *Chem. Phys.Lett.* **2000**, *330*, 558-562.

(169)   Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.

(170)   Zubarev, R. A. *Mass Spectrom. Rev.* **2003**, *22*, 57-77.

(171)   Polfer, N., C.; Haselmann, K. F.; Langridge-Smith, P. R. R. In *BMSS*: Loughborough, UK, 2002.

(172)   Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty Fred, W. *J. Am. Chem. Soc.* **1999**, *121*, 806-812.

(173)   Sze, S. K.; Ge, Y.; Oh, H.; McLafferty, F. W. *Anal. Chem.* **2003**, 1599-1603.

(174)   Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 10313-10317.

(175)   Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.

(176)   Horn, D. M.; Breuker, K.; Frank, A. J.; McLafferty, F. W. *J. Am. Chem. Soc.* **2001**, *123*, 9792-9799.

(177)   Breuker, K.; Oh, H.; Cerda, B. A.; Horn, D. M.; McLafferty, F. W. *Eur. J. Mass Spectrom.* **2002**, *8*, 177-180.

(178)   Breuker, K.; McLafferty, F. W. *Angew. Chem. Int. Ed. Engl.* **2003**, accepted.

(179)   Jurchen, J. C.; Williams, E. R. *J. Am. Chem. Soc.* **2003**, *125*, 2817-2826.

(180)   Kruger, N. A.; Zubarev, R. A.; Carpenter Barry, K.; Kelleher, N. L.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom. Ion Proc.* **1999**, *182,183*, 1-5.

(181)   Budnik, B. A.; Nielsen, M. L.; Olsen, J. V.; Haselmann, K. F.; Hörth, P.; Haehnel, W.; Zubarev, R. A. *Int. J. Mass Spectrom.* **2002**, *219*, 283-294.

Chapter 2
Experimental set-up and preliminary results

2.1 Introduction

The elemental constituents of a Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) are a strong magnet, an ultra-high vacuum, an ICR cell and an advanced data processing station [1]. Magnetic fields have historically been widely used in mass spectrometry and every mass spectrometer needs an advanced data processing station. The features that make FT-ICR MS truly unique compared to other mass spectrometric techniques are the non-destructive detection and the very long ion trapping times that can be achieved (hours).

Whereas historically ions were made inside the ICR cell mostly by electron ionisation (EI) [2], nowadays the ionisation is generally performed external to the magnetic field. There are multiple advantages to external ion generation, including easier sample introduction, easier changeover between different ionisation techniques and the fact that the most popular ionisation methods these days, electrospray ionisation (ESI) [3] and matrix-assisted laser desorption ionisation (MALDI) [4,5], operate on pressures which are incompatible with those required in the ICR cell during excitation/detection. The difference in pressure (ESI – atmospheric pressure; intermediate pressure MALDI – 10^{-3} atm) constitutes an engineering challenge which has been overcome by the use of multiple pumping stages.

In this Chapter the two FT-ICR MS instruments employed in this work are described. Preliminary electron capture dissociation (ECD) data obtained are presented, as well as the method used to optimise the parameters to carry out ECD on these instruments. An investigation of the electron kinetic energy distribution from a rhenium filament and an indirectly heated dispenser cathode is shown. The aim of this Chapter is to show how 'purer' ECD mass spectra (i.e. ECD spectra showing abundant peaks due to electron capture dissociation and minimising fragments from other fragmentation processes such as vibrational excitation) can be achieved. 'Pure'

ECD spectra are a prerequisite for mechanistic studies of ECD, which are discussed in Chapter 3.

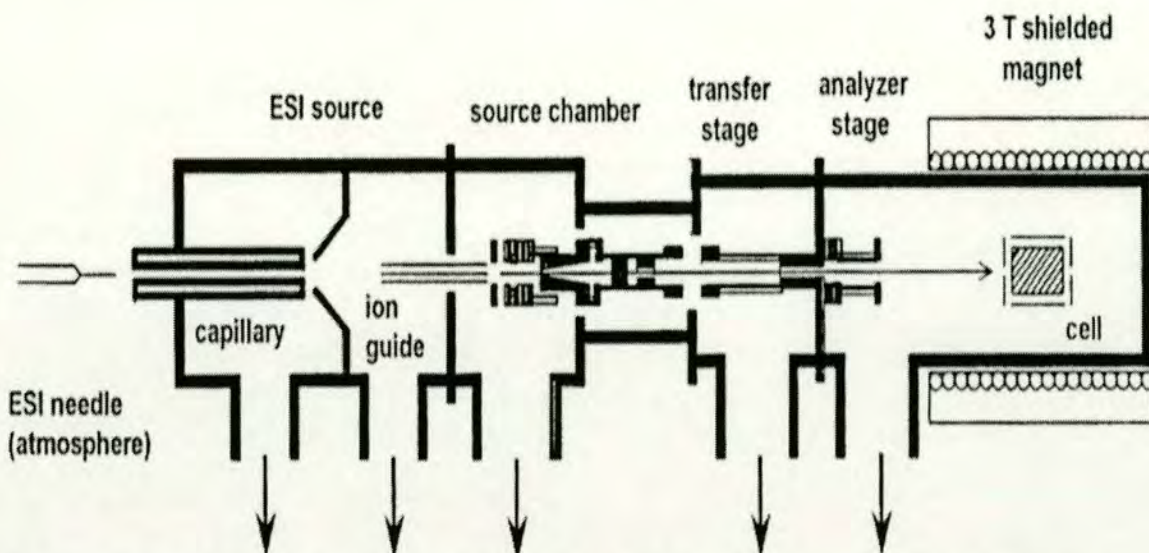## 2.1.1 3 Tesla Apex II Bruker FT-ICR MS

Figure 2.1 Diagram of Bruker Daltonics FT-ICR MS set-up.[1] Arrows indicate pumping stages.

Figure 2.1 shows a diagram of the experimental set-up of the 3 Tesla passively shielded APEX II Bruker Daltonics (Billerica, MA, USA) FT-ICR MS instrument equipped with an electrospray ionisation source (Analytica, Branford, CT, USA). The ions are formed at atmospheric pressure, then pass through a heated glass capillary with both ends metallised. This capillary (length 180 mm, ID 0.3 mm) also limits the leak to atmosphere of the first roughing stage, where the pressure is measured to be $\sim 10^{-2}$ mbar. Ions are accumulated in a storage hexapole located in the source chamber (at a pressure of 5×10^{-6} mbar), prior to pulsed transfer to the ICR detection cell (analyser stage). This is located in the homogeneous region of the 3 T super-conducting magnet and is at the ultra-high vacuum, measured to be $< 10^{-10}$ mbar in optimal conditions.

In between the source chamber and the analyser stage is the transfer stage which is made up of electrostatic lenses. In this region the ions are accelerated to ~3

---

[1] Diagram taken from Bruker Daltonics BioApex handbook.

keV and then focused using an Einzel lens [6]. The reason for focusing the ions is to overcome the magnetic mirror effect [7] which occurs in the fringing magnetic field and applies a force opposite to the direction of travel (z direction, see Figure 2.3). This effect is minimised when the ions enter the fringing field tangent to the magnetic field lines. Given the requirement for low trapping plate voltages (< 1V) in the ICR cell in order to achieve high resolution, the kinetic energy of the ions needs to be finely adjusted in order to be able to trap them. Other instrument manufacturers (e.g. IonSpec, Irvine, CA) use multipole ion guides to focus the ions in [8], which arguably allow simpler tuning of the ion transfer.
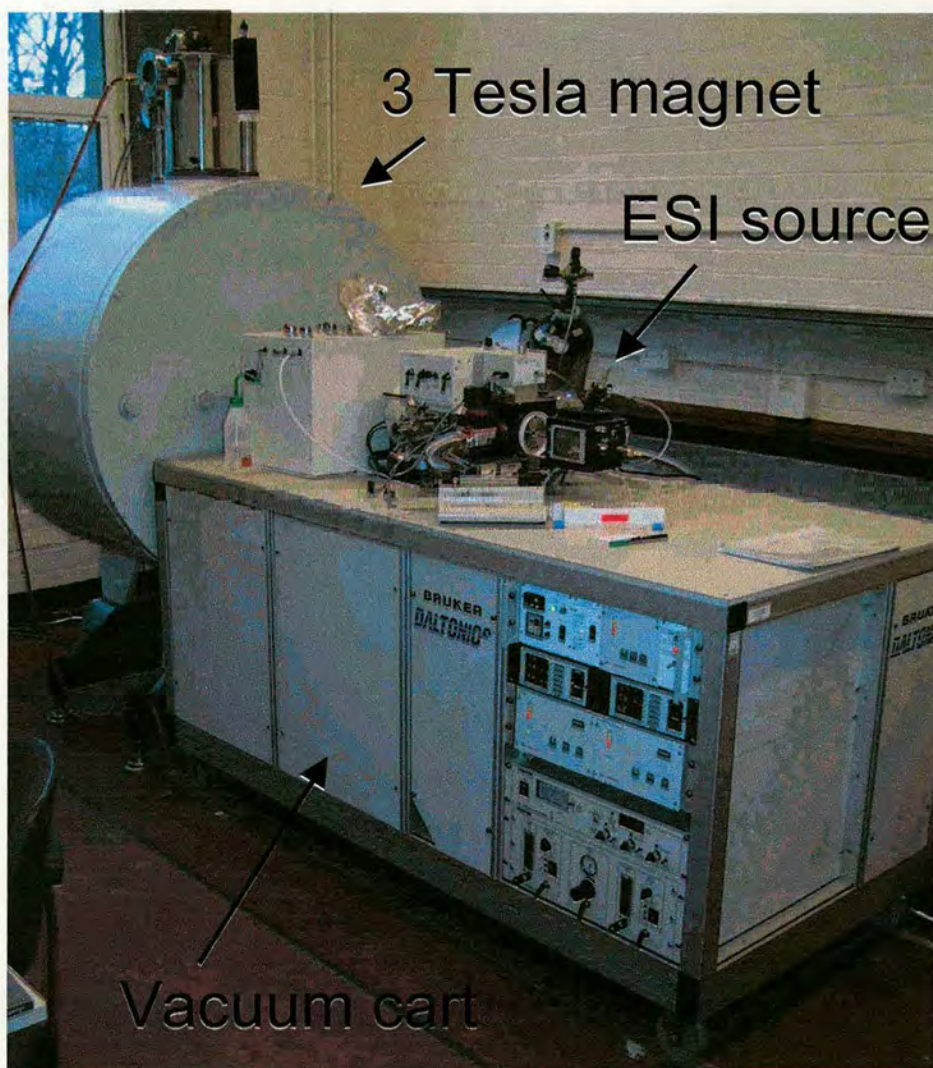
Figure 2.2 Picture of 3 Tesla Bruker Daltonics FT-ICR instrument.

Figure 2.2 shows a picture of the 3 Tesla instrument showing the superconducting magnet (Magnex, Oxford, UK) and the vacuum cart equipped with the ESI source. A syringe pump (Cole Parmer, IL, USA) used for ESI can be seen in the centre of the image. The vacuum cart can be wheeled in and out of the magnet bore allowing access to the ICR cell. The vacuum pumps consisting of rotary vane pumps and turbomolecular pumps (Edwards, West Sussex, UK) are contained inside the cart. The vacuum pressure displays can be seen on the right-hand-side of the vacuum cart.
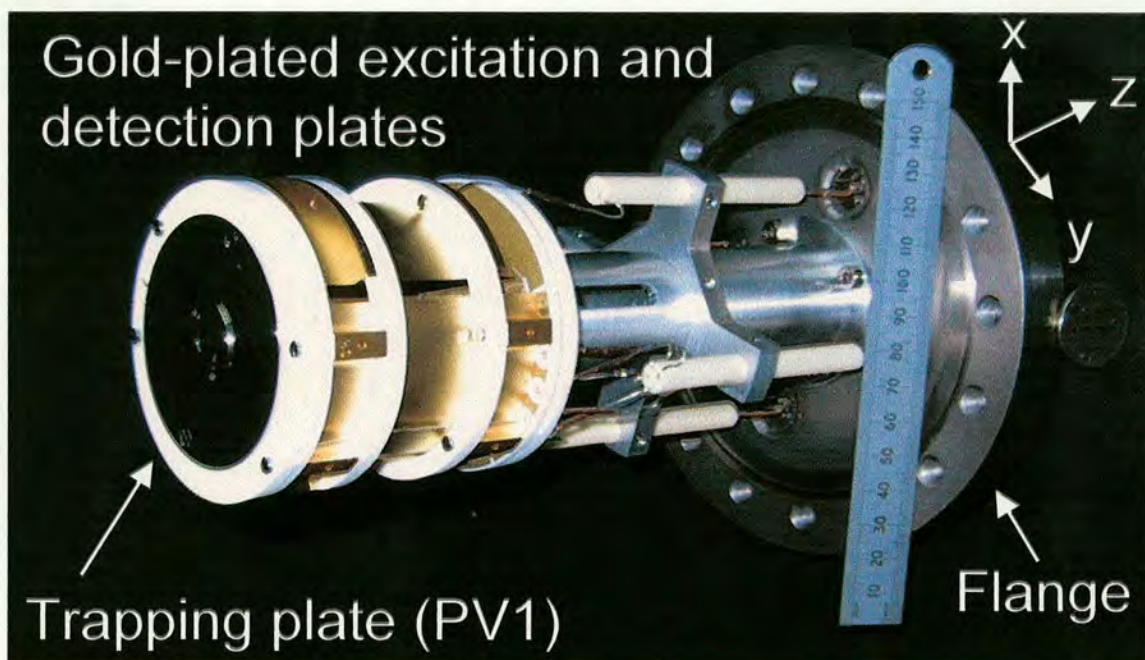
Figure 2.3 Picture of ICR Infinity™ cell.

The gold-plated ICR Infinity™ cell [9] is shown in Figure 2.3. Ions enter through the aperture of the trapping plate PV1; the second trapping plate PV2 is not visible on this picture. The cell dimensions are 60 x 60 mm (a ruler and a 5 pence coin are included in this picture to represent dimensions). By applying a DC voltage of the same sign as the ions (e.g. +1V for positive ions) the ions are trapped in the direction of the magnetic field B (z direction, see Figure 2.3). The Infinity™ cell derives its name from the fact that its electrostatic trapping plates are segmented with the intention of tricking the ions into believing to be in an infinite cell. The ICR cell is mounted on a flange on the rear side of the instrument.

2.1.2 9.4 Tesla Bruker APEX III FT-ICR MS

The 9.4 Tesla actively shielded APEX III FT-ICR MS is in many ways similar to the 3 Tesla instrument described above. Until recently 9.4 Tesla was the highest commercially available super-conducting magnet for FT-ICR, although a 14.5 Tesla magnet has been installed at the National High Magnetic Field Laboratory (NHMFL, Tallahassee, FL, USA).

This 9.4 T APEX III instrument (as opposed to the 3 T instrument) is equipped with an Apollo electrospray ionisation source (Bruker Daltonics, Billerica, MA, USA), which has superceded the Analytica ESI source. The Apollo source enables easier cleaning of the skimmer and glass capillary (length 180 mm, ID 0.2 mm). Further, switching between syringe infusion ESI, nanospray and atmospheric pressure chemical ionisation (APCI) can be achieved on the same ion source by simply exchanging the inlet door. Another ionisation source that can be fitted onto this instrument is a Scout 100 MALDI source, which includes a storage hexapole where the MALDI generated ions are cooled and ions from multiple laser shots can be accumulated prior to transfer to the ICR cell [10].

Due to a limiting digitisation rate of 10 Mhz, the highest cyclotron frequency that can be detected is 5 MHz according to the Nyquist criterion (section 1.3.2, Chapter 1). For the 9.4 T magnet instrument this means that the lowest mass-to-charge ratio (m/z) that can be detected is ~ 28.5, which compares to a lowest m/z of 9.5 on the 3 Tesla instrument. The highest m/z is theoretically unlimited, but since the mass resolution decreases linearly with m/z, even on a 9.4 T magnet ions of m/z > 15,000 are impractical to mass analyse.

2.2 Pulse sequence

The complex design of an FT-ICR instrument and the relatively long time scale of FT-ICR mass measurement (100's ms – seconds) allows for a sequential treatment of the events of the experiment. A screen-shot of the experimental sequence in the Xmass™ software for an ECD experiment is shown in Figure 2.4. A representative time-frame of a typical ECD experiment is shown in Figure 2.5, indicating the typical length of each event for the APEX III 9.4 T FT-ICR MS instrument. Each of the events is described in a paragraph hereafter.
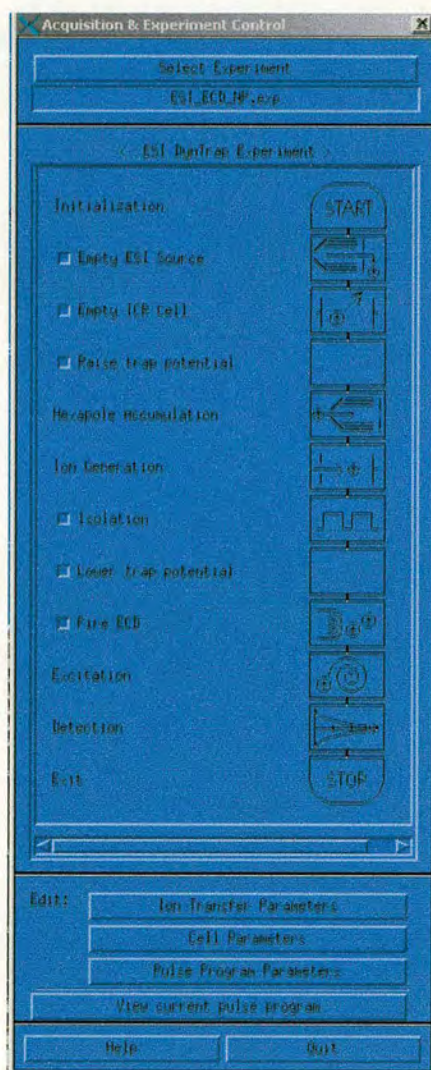
Figure 2.4 Experimental pulse sequence in Xmass™ software.
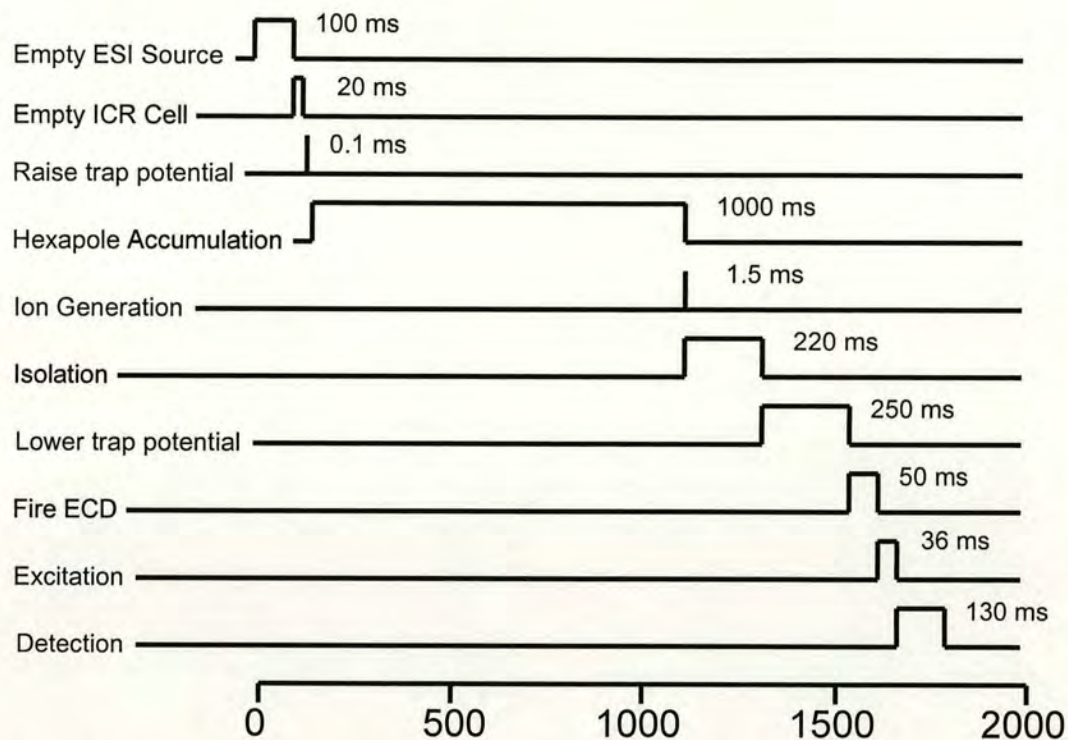
Figure 2.5 Schematic showing time frame of typical ECD experiment on APEX III 9.4 T FT-ICR MS instrument.

Empty ESI source/Empty ICR cell. After each detection pulse cycle any remaining ions are expelled in order to make way for a new batch of ions. To this effect, the trapping plates are set to +10 and −10 V respectively. The storage hexapole is purged of ions by applying a quench pulse; this is to prevent ion fragment products to form in the hexapole.

Raise trap potential. The initial ICR trapping potential are set to typically 3V on the rear trapping plate (PV2) and 2V on the front trapping plate (PV1). These settings were shown to enable efficient trapping of ions and to be more robust from day-to-day variations in the electrostatic voltages in the ion transfer optics.

Hexapole accumulation. Ions are accumulated in a storage hexapole for a duration of typically 600-1000 ms. This hexapole is suitable to store ions up to a m/z < 5000.

Ion Generation. Ions are extracted from the hexapole by a pulsed switching of the hexapole trapping plate voltage (e.g. + 20V for positive ions) to an extraction voltage (e.g. −10 V for positive ions) for a length of time (typically 1200-3000 µs).

Presumably, only 10% of the ions contained in the hexapole are extracted this way, due to the low electric field penetration of the trap/extract plate into the hexapole. Marshall and co-workers have used an octapole with DC rods which causes a constant voltage drop along the length of the multipole, thereby increasing their extraction efficiency from the octapole over 10-fold [11]. During the ion extraction time, the ion packet travels through the ion transfer optics and the fringing magnetic field into the ICR cell where trapping occurs. In order to facilitate trapping, a pulsed gas (e.g. argon, 10^{-6} mbar) can be applied to collisionally cool the z-direction velocity of the ions. The application of a cooling gas requires a pump-down delay (typically 1-5 seconds), which lengthens the experimental time frame of the experiment and is therefore not used. Alternatively, the Sidekick™ trapping method [12] can be employed, which consists in deviating the ion path laterally (xy direction, see Figure 2.4) by applying a voltage gradient between two electrode plates at the entrance of the ICR cell (EV2a and EV2b). The effect of this Sidekick™ is to decrease the z-direction velocity of the ions, effectively spreading out the ion beam into xy. Hence, Sidekick™ trapping comes at the price of increasing the magnetron radius of the ions, which can adversely affect resolution and sensitivity.

Isolation. An ion of interest can be mass isolated by applying a sweep isolation pulse on the excitation plates that does not include the frequency of the ions to be isolated. The isolation sweep generally starts at high frequency (i.e. low *m/z*) and then proceeds to low frequency at a defined frequency step (parameter XBB, normally set to 100), with a 'notch' at the frequency of the ions of interest (i.e. that should not be ejected). The dwell time for each frequency and the excitation amplitude define the shape of the excitation sweep, which can be visualised by doing a 'Wavedetect' experiment (simultaneous excitation and detection). The parameter 'ejection safety belt' (parameter JBB, typically 2000 Hz) determines how close the ejection sweep gets to the cyclotron frequency of the ions of interest. At low values of JBB (< 1000 Hz) off-resonance excitation of the ions of interest can occur, causing these ions to be taken from the centre of the cell.

Lower trap potential. The trapping plate voltages are reduced to 1V on both plates in order to increase the mass resolution that can be achieved.[2]

Fire ECD. Electron irradiation is effected by pulsing a bias voltage in order to cause spontaneous electron emission from the electron filament (3 T, APEX II) or the dispenser cathode (9.4 T, APEX III) (more detail see section 2.3.4). In order for electron capture dissociation (ECD) [13] to take place the ion cloud should be near the central axis of the ICR cell. This implies low values for Sidekick™ trapping and 'careful' isolation. Further, electron capture is maximised for low kinetic energy electrons (< 0.2 eV). A study of the electron kinetic energy distribution is shown in section 2.3.4.

Excitation. When applying a radiofrequency (rf) voltage on both excitation plates (dipolar excitation) of the same frequency than the cyclotron frequency the cyclotron radius is enhanced to a detectable radius (see Chapter 1 for theory). By applying a frequency sweep, the cyclotron radii of all ions are excited. The excitation stage is in fact very similar to the isolation stage, only this time the ions are not excited beyond the cell geometry.

Detection. As a result of the dipolar excitation, ions of a particular *m/z* cyclotron as a coherent ion packet, causing an induced current on the detection plates. Due to collisions of this ion packet with background gas de-phasing of the ion packet occurs, which results in an exponential decay of the *time-domain ICR signal*. In order to observe long transients long mean free paths are critical, which requires an ultra-high vacuum (< 10^{-9} mbar).

The induced current signal is amplified and then digitised (using an analogue-to-digital converter), which is Fourier transformed to deduce the individual frequencies that make up the *time-domain ICR signal*. The frequency spectrum is then converted to a mass-to-charge spectrum based on the fundamental equation for cyclotron frequency (see Equation 1.1, Chapter 1). Calibration of the mass spectra is

---

[2] Note that for the APEX II 3 T FT-ICR MS instrument the trapping voltages were always lowered to < 1V (typically 0.1 V) (more detail see section 2.3.2.2).

performed with the 'Cal2' procedure in the Xmass™ software, which is based on the equation developed by Ledford *et al* [14].

2.3 Preliminary results

2.3.1 Peptide fragmentation

As described in Chapter 1 (section 1.6), three distinct peptide backbone bonds exist where cleavage can occur: the C_α-C bond, the peptide bond C-N and the N-C_α bond (see Figure 1.9, Chapter 1). Of these the peptide bond is generally the weakest and in collisionally activated dissociation (CAD) [15], of which SORI CAD [16] is one example, b and y fragments are most abundant. Conversely, in electron capture dissociation c and z fragments are generally the most abundant fragments. Whereas cleavage in SORI CAD occurs due to a gradual increase in the vibrational energy levels in the molecule, the mechanism in ECD is believed to be *non-ergodic* (i.e. dissociation before energy randomisation). ECD experimental pulse sequences are configured to minimise other dissociation processes (such as vibrational excitation), with the intention of studying the mechanistic processes in ECD. Mechanistic aspects of ECD are discussed in Chapter 3.

2.3.2 Electron capture dissociation with rhenium filament

The polypeptides in this study were obtained from Sigma-Aldrich (unless specified otherwise) and were electrosprayed by syringe infusion (flow rates 80-120 µl/hr) at concentrations 10-100 µM in H_2O/MeOH/CH_3COOH (49:49:2). In order to carry out ECD experiments one requires a source of electrons, which is generally an electron filament. Alternatively, a laser could be used to emit electrons from a metal plate through the photo-electric effect. Incidentally, the first time ECD was observed it was falsely ascribed to photo-fragmentation, when actually the 193 nm laser employed caused emission of electrons, thereby inducing ECD [17]. However, the current intensities from a filament are much higher and hence it is more convenient to use. Further, most commercial FT-ICR instruments are already equipped with an electron filament; our 3 Tesla instrument is fitted with a rhenium filament located at a distance of 13 cm from the rear trapping plate (see Figure 2.6). This filament is normally used for carrying out electron ionisation (EI) inside the detection cell,

which is a good predictive tool to check that the excitation and detection work when trouble-shooting the instrument. Whereas EI experiments are performed with electrons at relatively high kinetic energy (e.g. 70 eV), ECD requires the use of low-energy electrons (i.e. < 1 eV).
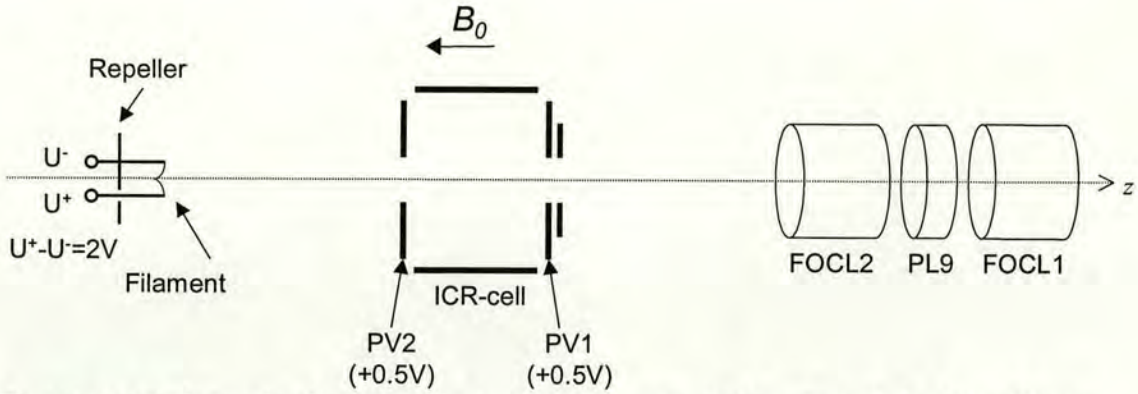

Figure 2.6 Schematic diagram of electron filament, ICR cell and the Einzel lens elements FOCL2, PL9 and FOCL1. B_0 indicates direction of magnetic field. Note that the electron gun and the ICR cell are located inside the strong magnetic field, whereas the Einzel lens elements are in the fringing field.

The rhenium filament (ca. 5 mm x 0.038 x 0.51 mm Re ribbon, resistance 0.09959 Ohm/cm) was operated at a current of 3 A, which is equivalent to a voltage drop of 2 V. When pulsing the filament to emit electrons a voltage U^- (range –0.48 to +0.65 V) was applied to one side of the filament. As a consequence of the potential drop across the filament, a voltage U^+ ($U^+ = U^- + 2$ V) was measured on the other side of the filament. The emitted electron beam passes through the central axis of the ICR cell. A more detailed account of the kinetic energy distribution of electrons emitted from this filament is given in section 2.3.3.

2.3.2.1 Static trapping

The initial set-up of the 3 Tesla instrument did not allow for the trapping voltages to be altered during the experiment. Hence these experiments were run at static trapping voltages of 0.5 V. Trapping voltages of 1 V or higher did not result in successful ECD experiments, whereas lower trapping voltages than 0.5 V do not permit trapping of the ions. Another prerequisite for ECD to occur, is that the ion cloud and the electron beam need to overlap, which is why no Sidekick™ trapping

was employed in these experiments. No pulsed gas was employed to collisionally cool the ions, to avoid CAD.
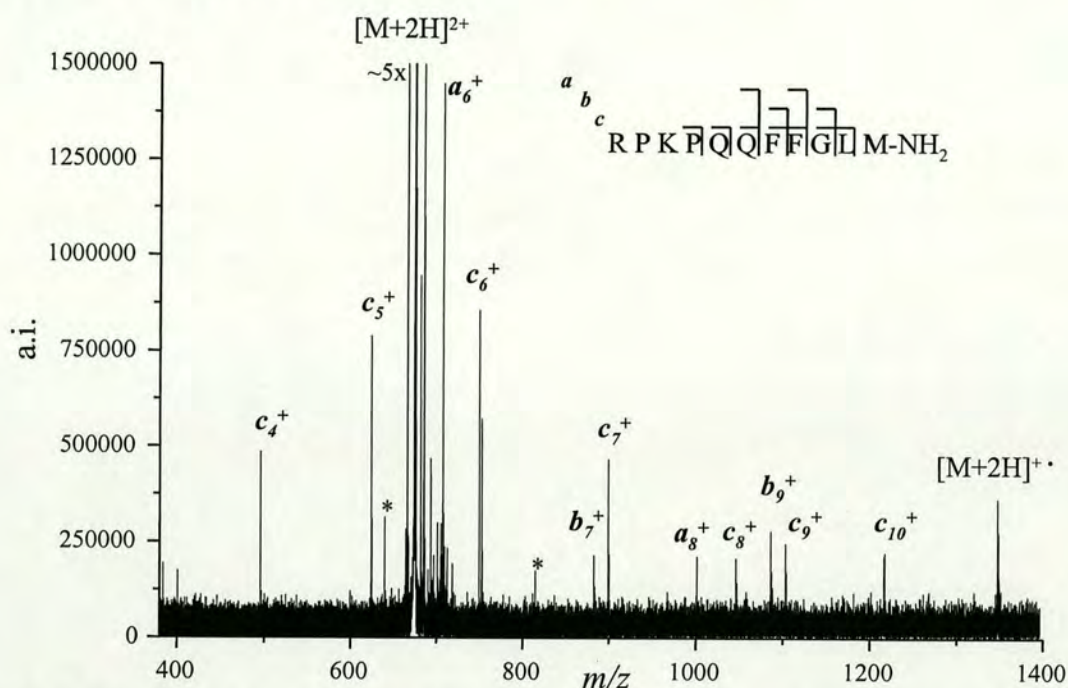

Figure 2.7 ECD mass spectrum of substance P in the 2+ charge state. Electron irradiation time 1 s; 40 time-domain transients (512 kb) were co-added to generate this mass spectrum.[3]

Electron irradiation of the mass isolated doubly protonated form $[M+2H]^{2+}$ of substance P (MW 1347.6, Sigma Aldrich) gave rise to the ECD mass spectrum shown in Figure 2.7. The capture of one electron without resulting fragmentation leads to the reduced odd electron species $[M+2H]^{+\bullet}$ at a m/z double of the isolated precursor ion. Evidently, at least doubly protonated species are necessary to perform ECD experiments on, since otherwise no charged fragments can be observed in the mass spectrum. A series of c fragments is observed in this spectrum, which can be rationalised from the fact that the charge-carrying arginine residue is located at the N-terminus of the molecule. The lowest abundance c fragment is c_8^+, with a S/N ratio of 10. Two b ions are also present in this spectrum, suggesting that other activation mechanisms took place. As a result, this spectrum does not qualify as a 'pure' ECD spectrum.

---

[3] Note that the amino acid 1 and 3-letter code is shown in Appendix 2, Table A2.1.

An attempt was made to optimise the experimental parameters of electron injection to obtain better ECD. The bias of the filament voltage U⁻, which is also applied to the repeller plate (see Figure 2.6), determines the kinetic energy of the electrons. Hence, fine tuning of U⁻ should lead to better electron capture. Figure 2.8 shows the ECD fragment efficiency for the 2+ charge state of substance P as a function of the filament voltage U⁻ [18]. A third order polynomial function gave an adequate fit to this data set and the maximum c fragment abundance is determined to be at a value of U⁻ of 0.52 V.
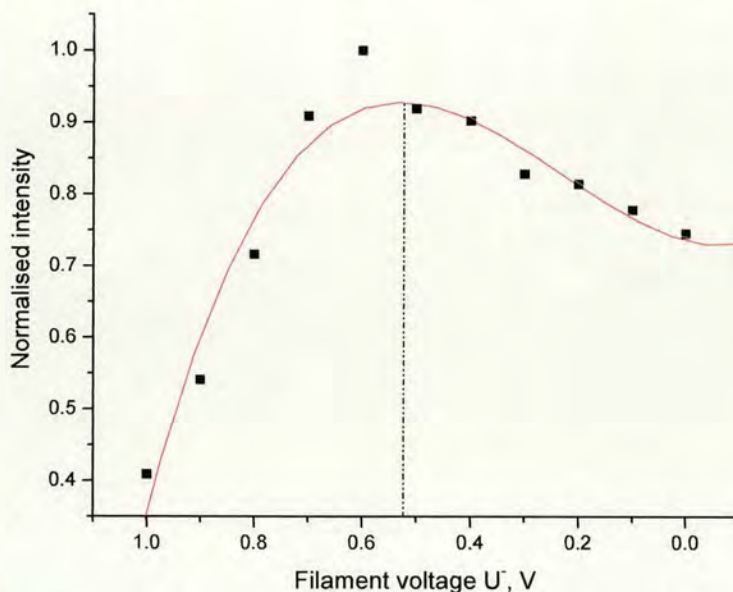
Figure 2.8 ECD fragment efficiency for substance P as a function of filament voltage U⁻. The ECD fragment efficiency is calculated by adding the ECD fragment peaks c_4^+ to c_{10}^+, and normalising with respect to the mass spectrum with the highest total abundance of c fragments.

ECD of bee venom melittin (MW 2845.5, Sigma Aldrich) was also attempted with these very similar optimised parameters (U⁻ = +0.65 V). Electron irradiation of the most abundant charge state, 4+ $[M+4H]^{4+}$, gave rise to the ECD mass spectrum in Figure 2.9. The conversion efficiency of the precursor ion $[M+4H]^{4+}$ to fragment ions is much reduced in this spectrum, compared to the ECD mass spectrum of substance P (Figure 2.7). According to Coulombic attraction of charges, the electron capture cross-section should increase by z^2, z being the charge on the electron-capturing cation, and hence one would expect better electron capture. However, the efficiency of precursor to fragment ion conversion in ECD also largely depends on

the overlap between the ion cloud and the electron beam, which may not be the same in this experiment than for ECD of substance P. For larger species the number of fragmentation channels is greater and hence the S/N ratio for individual fragment peaks is reduced. A total of 17 fragment ions are observed in this spectrum, as well as typical neutral losses expected in ECD, such as hydrogen atom desorption, ammonia and arginine side-chain losses [19,20]. The mass resolution of the reduced species $[M+4H]^{3+\bullet}$ and the fragment peaks in this spectrum is ~8000. Hence, the mass resolving power of the 3 Tesla instrument is adequate to unit mass resolve ECD fragment ions of polypeptides. ECD of the 4+ charge state of bovine insulin B chain (MW 3495.9, Sigma Aldrich) gave 13 (out of 29 possible) backbone cleavages [18]. However, irradiation times of 5 s needed to be employed and 200 time-domain transients were co-added, resulting in a total experimental time of 23 minutes. ECD of the 7+ charge state of bovine ubiquitin (MW 8564.9, Sigma Aldrich) only yielded very limited fragmentation even for 200 scans. This implies that this 3 Tesla FT-ICR MS is suitable for ECD of polypeptides of molecular weights up to 3.5 kDa, but not for top-down sequencing of proteins [21].
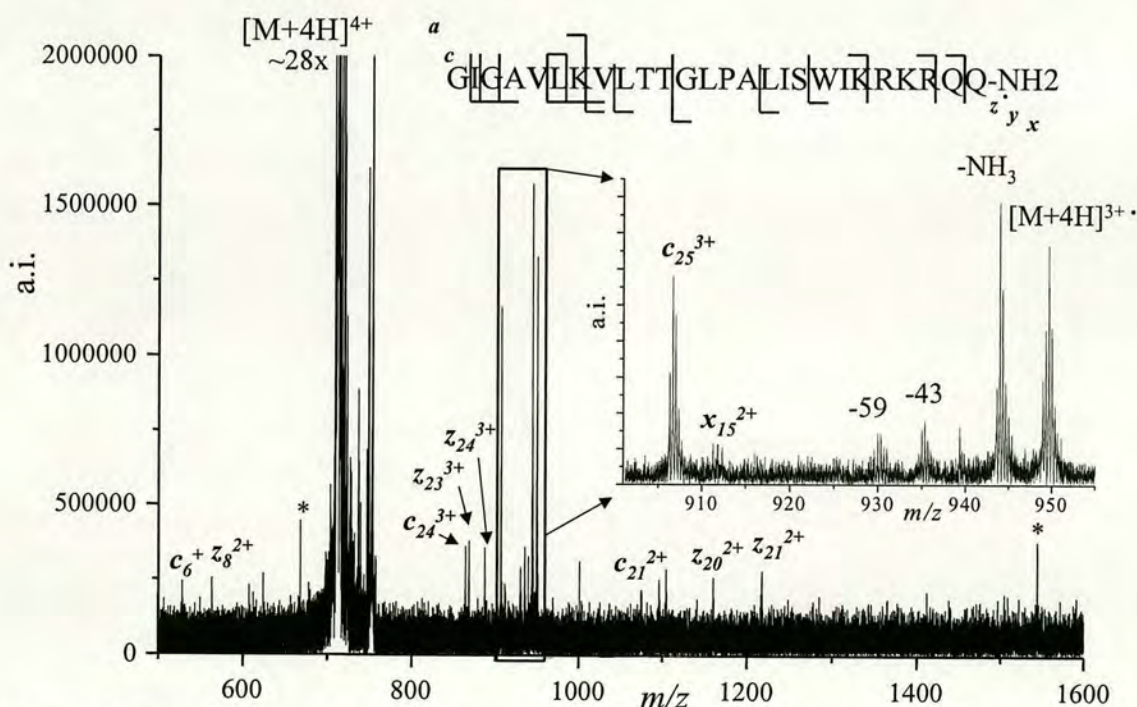

Figure 2.9 ECD mass spectrum of melittin in the 4+ charge state. Electron irradiation time was 1.5 s and 100 time-domain transients (512 kb) were co-added to generate this spectrum. The insert shows an expansion of the reduced species $[M+4H]^{3+\bullet}$. Asterisks denote artefact peaks.

2.3.2.2 Dynamic trapping

A hardware upgrade to the 3 Tesla instrument permitted ramping of the trapping plate voltages during the experiment (ICC controller). Given the experience that lower trapping plate voltages resulted in better ECD mass spectra, the attempt was made to reduce the trapping plate voltages as low as possible. By slowly (100-250 ms) ramping the trapping plate voltages from initially 3 V (PV2) and 2 V (PV1) the voltages for both trapping plates could be reduced to 0.1 V while still trapping a large number of ions. The mass isolation sweep was carried out after ramping down the voltages, followed by electron irradiation. Optimised parameters for U⁻ were again used in the range found to be optimised before.

Figure 2.10 shows the ECD mass spectrum of the 2+ charge state of a variant of the luteinizing hormone releasing hormone (LHRH, MW 1312.5). A series of z fragments (z_3^+ to z_9^+) can be seen in this spectrum. This mass spectrum exemplifies the use of ECD as a tool for *de novo* sequencing [22]. The mass differences between the z fragment peaks accounts for 6 consecutive amino acids that are sequenced without even having to know the peptide sequence. The S/N ratio of the fragment peaks is high, supporting the assumption that lower trapping plate potentials during electron irradiation lead to better ECD spectra. Further, peaks due to other fragmentation processes, such as vibrational excitation, are minimised in this spectrum. This spectrum reflects a 'purer' form of an ECD spectrum, which could be used to study the mechanistic processes in ECD.

The experimental parameters for the 3 Tesla instrument seem to be optimised well enough under these conditions to attempt more difficult problems. A study by Chan *et al.* [23] on their 4.7 Tesla FT-ICR instrument showed that they obtained good quality ECD spectra by short pulse (50 ms) electron irradiation and that employing a cooling gas improved the S/N ratio of the fragment ions. Such a strategy only worked on the 3 Tesla instrument in some isolated cases and hence it appears that different instruments may require different strategies of optimisation.
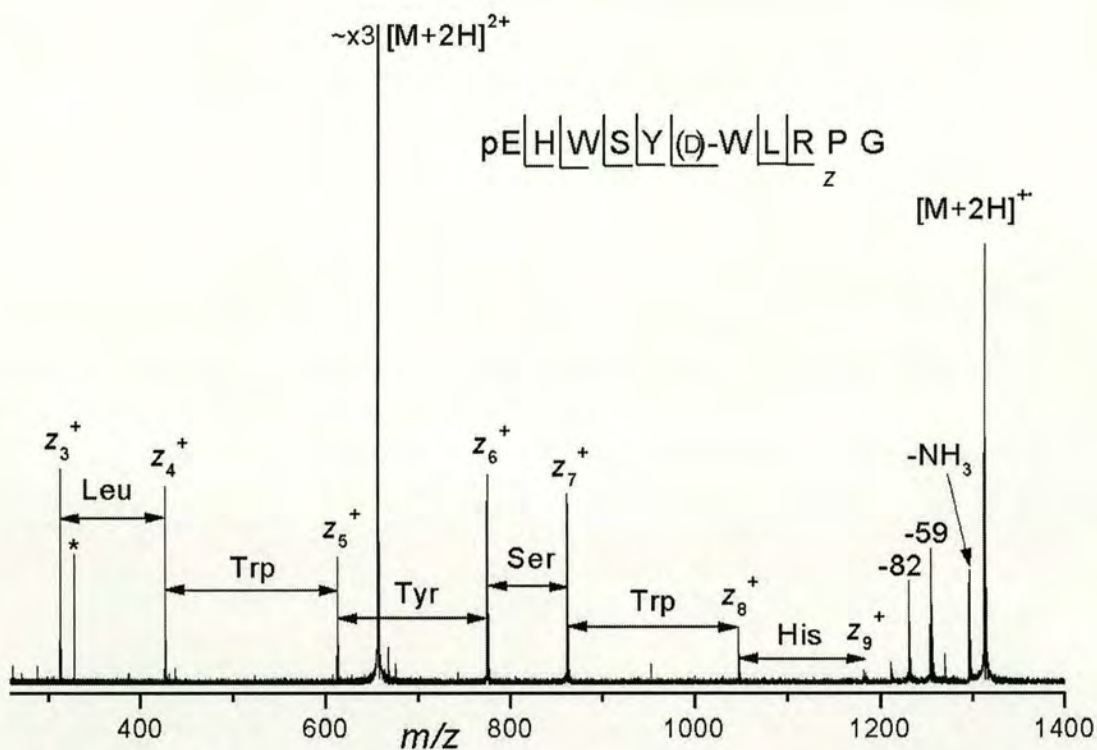
Figure 2.10 ECD mass spectrum of LHRH in the 2+ charge state. Electron irradiation time is 200 ms and 50 time-domain transients were co-added to generate this mass spectrum.

The ECD mechanism, in contrast to the *heating* techniques (e.g. SORI CAD, IRMPD, BIRD), is believed to be *non-ergodic*. This means that backbone cleavage occurs before energy randomises in the molecule, whereas fragmentation in the heating techniques takes place when the molecule is highly vibrationally excited. One would therefore assume that ECD is less likely to lose weakly bound groups in the fragmentation process. Recently, interest has grown in ECD as a tool to identify and localise post-translational modifications [24-28]. Many of these modifications, and in particular glycosylations, are more weakly bound and could hence be lost in harsher fragmentation conditions. For N-glycosylations [29] a consensus sequence exists, which allows such sites to be predicted from the protein sequence; this does not apply to O-glycosylations, which makes them a more useful target for ECD.

64

Figure 2.11 shows the ECD mass spectrum of the doubly O-glycosylated peptide fragment TAP25[4] (MW 2793.0). O-glycosylations occur on threonine (T) and serine (S) residues. Here it was possible to identify and localise the glycosylation sites, since cleavage on either side of the modified threonine residues was achieved. This mass spectrum is similar to other reported ECD spectra of modified TAP25 peptides [25].
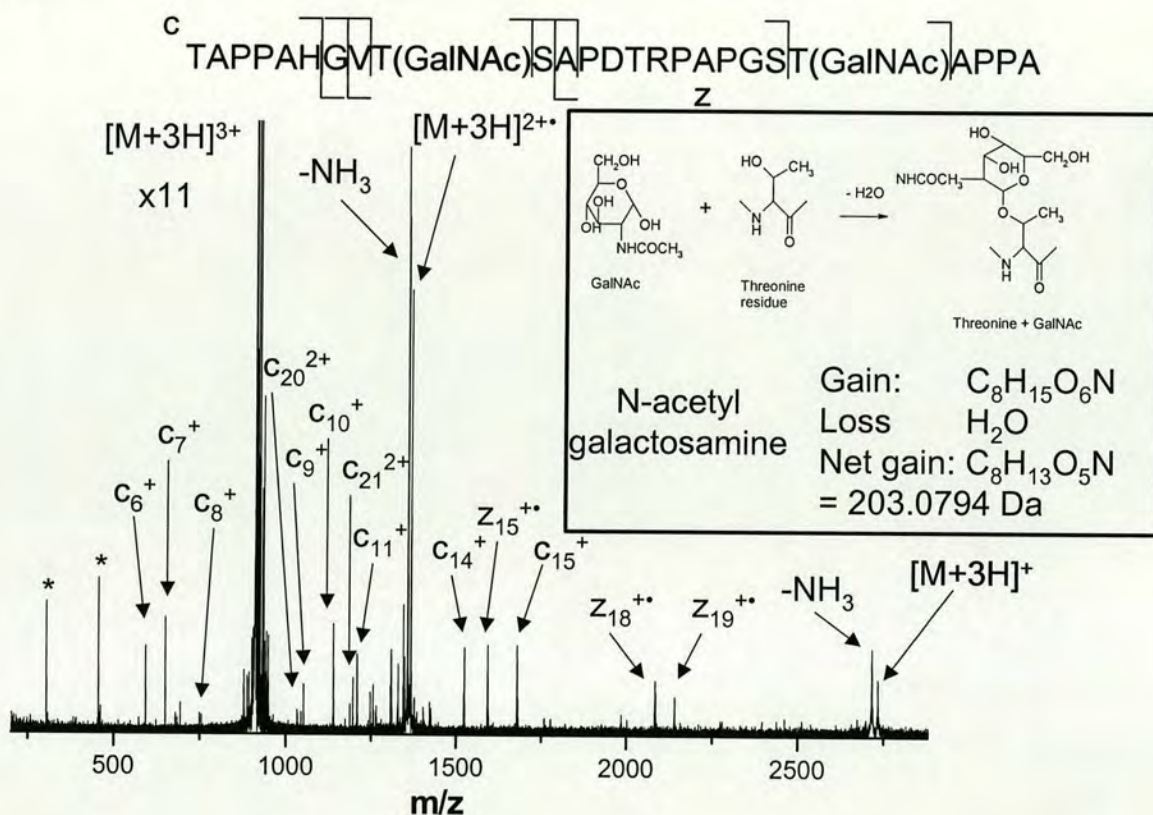

Figure 2.11 ECD mass spectrum of TAP25 carrying two GalNAc groups in the 3+ charge state. Electron irradiation time 1 s (100 scans). Insert shows addition of GalNAc group onto threonine residue and corresponding mass increase to peptide.

A combination of ECD and IRMPD, as demonstrated by Håkansson et al. [28], can be used to further elucidate the fine structure of the glycosylation. The power of the IRMPD laser can be tuned low enough to specifically break the bonds in the sugar group, without breaking backbone bonds. By using the IRMPD laser

[4] TAP25 peptide derived from the human mucin MUC1, was in vitro glycosylated with GalNAc-T3 transferase. Sample obtained from the group of Prof. R. Zubarev (Angström Laboratory, Uppsala University, Sweden) and originally made by H. Clausen (School of Dentistry, University of Copenhagen, Denmark).

slightly off-axis, it has been shown that IRMPD and ECD experiments can now be effected on the same ion population [28], which makes FT-ICR MS a very powerful technique to investigate post-translational modifications, and particularly glycosylations.

2.3.3 Electron capture dissociation with dispenser cathode

ECD experiments were attempted on the 9.4 Tesla APEX III instrument using the fitted rhenium filament, however, no successful spectra were obtained. The stronger magnetic field leads to a greater focusing of both the ion cloud and the electron beam. Hence, one would theoretically expect greater ease of ion cloud/electron beam overlap at lower magnetic fields. Similar strategies employed for the 3 Tesla instrument, such as lowering of trapping plate potentials to 0.1 V and even employing a pulsed gas (to cool or scatter the electron beam), did not result in successful ECD. It appears that when the electrons are 'born' on a particular field line, they are unlikely to move away from it. The greatest limitation of the filament set-up in performing ECD experiments is the very small diameter of the electron beam, which precludes efficient ion cloud/electron beam overlap. As McLafferty and co-workers have shown, by using a design of ICR cell consisting of an open-ended cylindrical cell with additional flat trapping plate electrodes, 'the nested cell geometry' [22], positive ions and electrons can be trapped at the same time. While this can result in better precursor ion to fragment ion conversion, the overall time scale for such experiments is very long (multiple electron irradiation and storage times resulting in single scan times of up to ~ 90 s).

Tsybin et al. [30] and Haselmann et al. [31] have used indirectly heated high-surface dispenser cathodes that have larger emitting currents and higher electron beam diameters. They have reported irradiation times as low as 1 ms, which allow on-line liquid chromatography (LC) eluted peptides to be fragmented by ECD.

A high-surface BaO coated dispenser cathode was installed on the 9.4 Tesla instrument. This has a diameter of 5 mm (HeatWave, Watsonville, CA, USA) (see Figure 2.12). This cathode was fitted in place of the rhenium filament (see Figure 2.6). In normal operating conditions, a current of 1.8 A was applied and the grid voltage was maintained at 15.2 V during the experiment, both of which were applied

using an external power supply (ElektroAutomatik, Viersen, Germany) (see Figure 2.13). This power supply was operated in current-limiting mode; for a current of 1.8 A voltage drop of 5.2 V was applied across the filament in steady-state conditions.
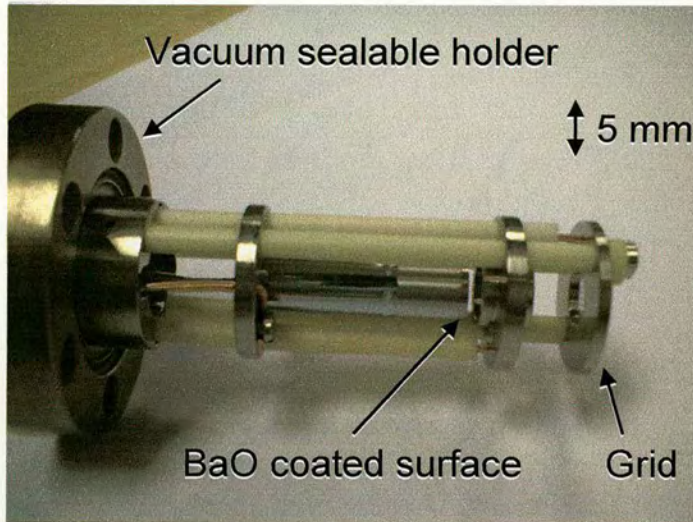

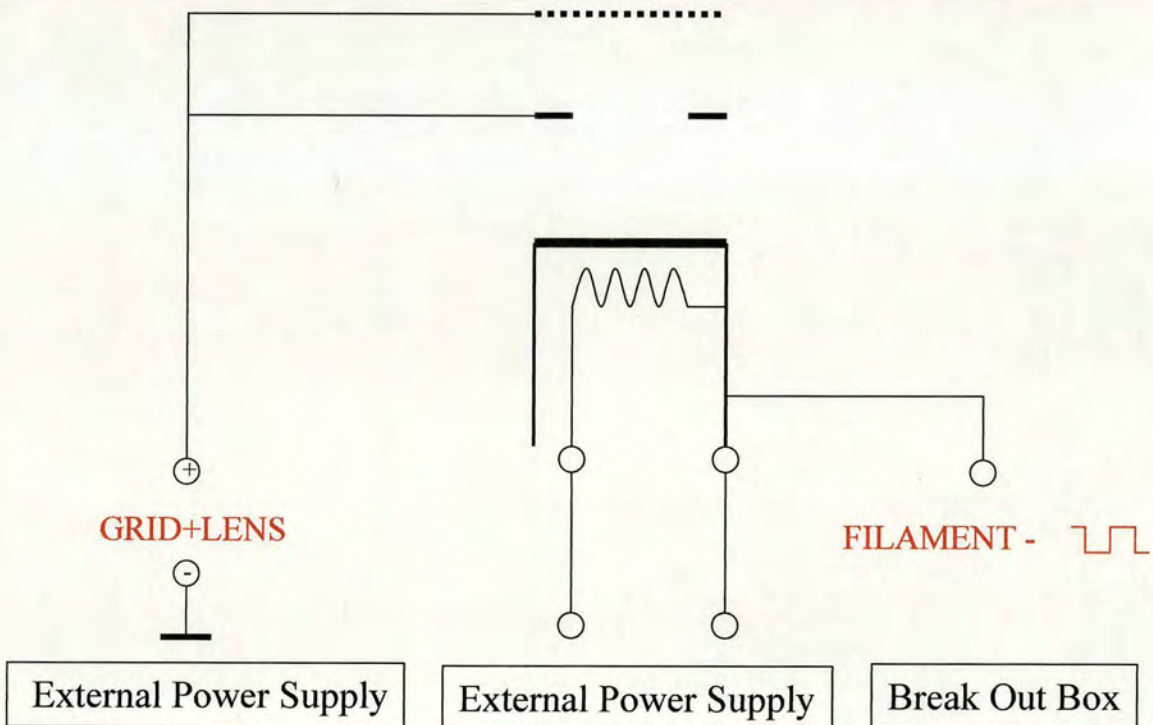**Figure 2.12** Picture of the indirectly heated dispenser cathode.

Figure 2.13 Schematic diagram of the dispenser cathode with connections.

One side of the filament is floated with the U⁻ voltage in the Xmass™ software in order to pulse electron emission. This voltage is taken from the filament (-) voltage on the break out box. In the case of the cathode, ramping the trapping voltages to lower values before electron irradiation did not seem to have a significant effect on the efficiency of precursor to fragment ion conversion.

An example ECD spectrum of substance P in the 2+ charge state is shown in Figure 2.14. Comparing this spectrum to the ECD spectrum obtained with the electron filament for static trapping potentials (Figure 2.7) one observes much better defined fragment peaks. Further, the most abundant fragment peaks are due only to the ECD process, which is critical in order to carry out mechanistic studies. Mass isolation, when performing electron irradiation with the rhenium filament, often had to be 'compromised' in order to maintain overlap between the ion cloud and the electron beam. Conversely, isolation of the 2+ charge state of substance P $[M+2H]^{2+}$ is nearly perfect, with even the sodium adduct species ejected in the isolation sweep. The ratio of species that have captured an electron to remaining $[M+2H]^{2+}$ in Figure 2.14 is 4 to 1, which suggests efficient conversion of precursor to fragment ions.
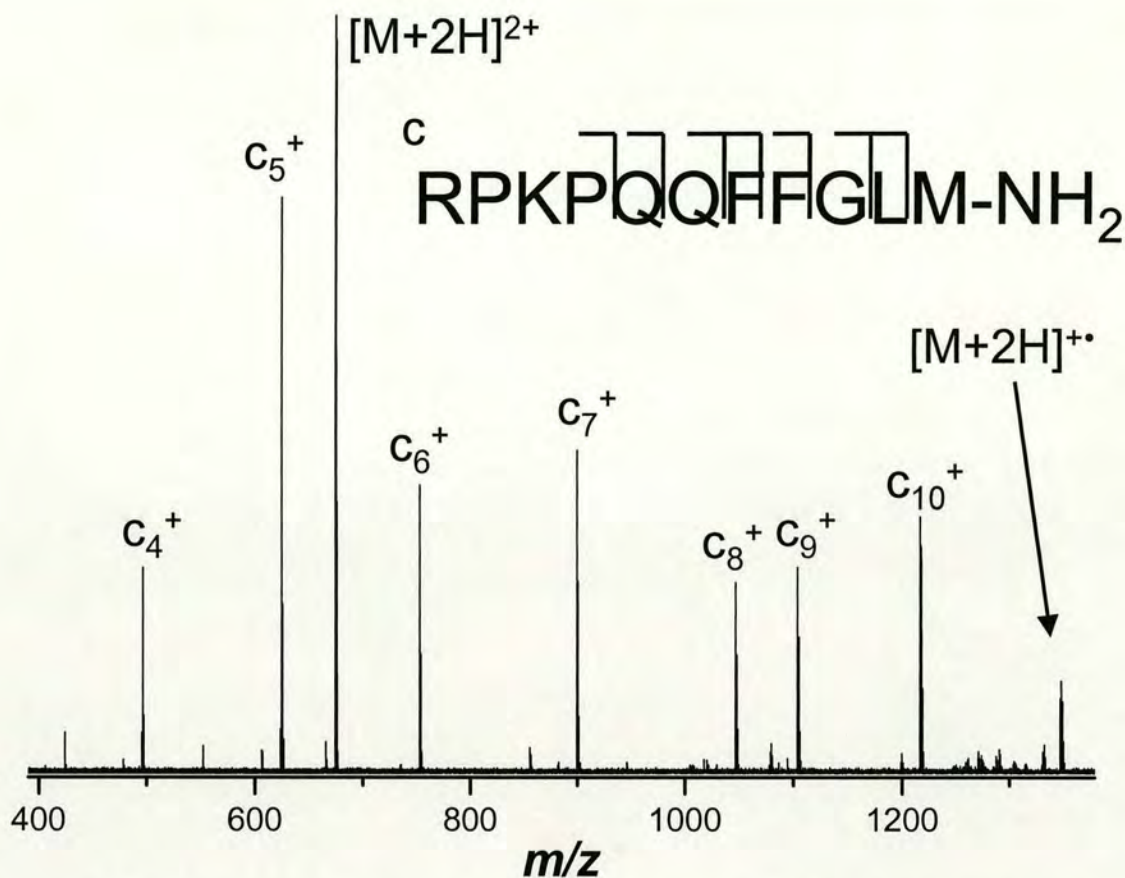
Figure 2.14 ECD mass spectrum of substance P in the 2+ charge state. Electron irradiation was 20 ms and 30 time-domain transients (512 kb) were co-added to generate this mass spectrum.

2.3.4 Characterisation of electron sources

In order to characterise the electron filament and dispenser cathodes, measurements of the emitted electron currents were made. One side of the filament and the repeller plate are held at a voltage U^-. As a consequence of the potential drop across the filament, a voltage U^+ is measured on the other side of the filament which is 2 V higher than U^- for the rhenium filament at a current of 3 A. The repeller plate is negatively biased relative to U^+, which causes pulsed emission of electrons from the filament. Some of the emitted electrons travel in the z-direction (axis of direction of magnetic field B_0) through the ICR cell to the Einzel lens components (see Figure 2.6). The emitted electron current was measured using a picoammeter (model 610C, Keithley Instruments, Munich, Germany) at the Einzel lens element closest to the ICR cell. No electron current could be detected on any of the other Einzel lens

elements. Whereas the electron gun and the ICR cell are located inside the bore of the magnet (i.e. at 3 Tesla, respectively 9.4 Tesla field strength), the Einzel lens elements are placed in the fringing magnetic field. Electrons travel along a particular field line and in the fringing magnetic field the field lines diverge, causing the electrons to impact onto the Einzel lens element FOCL2. All voltages (electrodes in the ICR cell: trapping voltages, Sidekick™ voltages, etc., and the Einzel lens elements) are set to 0 V (i.e. ground) to prevent interference with the electron path.

2.3.4.1 Rhenium filament electron emission

The emitted current intensity of the rhenium filament for a filament current of 3 A as a function of the filament bias U⁻ was measured on FOCL2 as described above and the plot is shown in Figure 2.15. The current intensity plot has a sigmoidal shape and levels off at a current of 1400 nA. The best ECD spectra (i.e. highest abundance ECD fragments in those spectra) were obtained at a value of U⁻ = +0.5 V, which is the point where the current intensity starts increasing. The measured current at that filament voltage was ~30 nA, which is equivalent to 2×10^{11} electrons. This compares to a space-charge capacity of the 3 Tesla ICR trap of a few million charges, which demonstrates that many more electrons pass through the cell than there are cationic charges present to capture electrons.
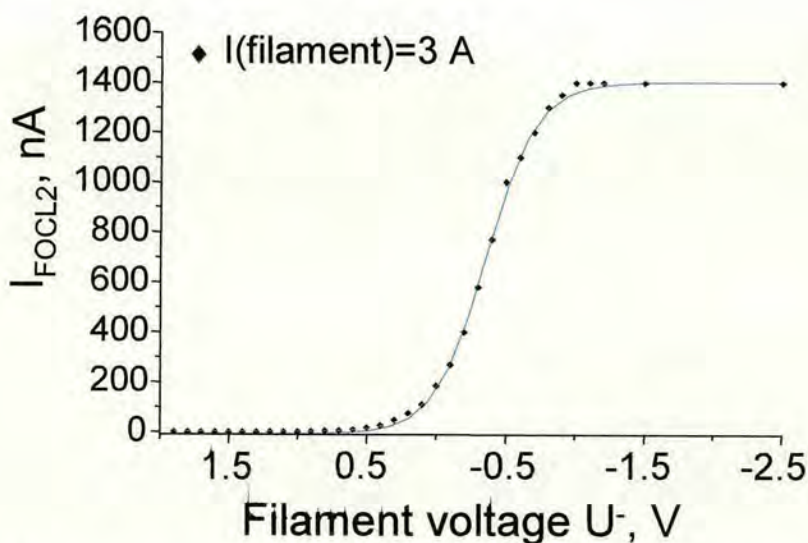

Figure 2.15 Intensity of emitted current detected on FOCL2 I_{FOCL2} as a function of the filament voltage U⁻ (filament current 3 A). Line shows best sigmoidal fit.

The question arises as to what kinetic energy (KE) the electrons need to have for *efficient* ECD to take place; ECD efficiency is defined as the conversion efficiency of precursor to fragment ions [22]. McLafferty and co-workers proposed a value of < 0.2 eV for low-energy electrons which would be involved in ECD. Measuring the KE distribution of emitted electrons from a filament is actually quite challenging, since it depends on the work function that the electrons need to overcome to enter the vacuum [32]; the work function of Re is 4.9 eV. In order to overcome this work function, the filament is heated by applying a current, which raises the energy available to electrons to escape the filament. The Re melting point is high (T_m = 3170 °C), which is why it is suitable for electron emission. Since thermionic emission from pure metals only occurs at temperatures > 2500 K, an electrostatic field is applied in the vicinity of the metal surface to lower the activation barrier [33]. Pulsing of the U⁻ voltage to a suitable value causes favourable conditions for electrons to be emitted from the filament. The mean KE distribution depends on the value of U⁻ that is applied. Nevertheless, we need not be concerned with the electron KE distribution at the filament since ECD occurs in the ICR cell. Hence, this is where the KE distribution must be known.

This renders the measurement of electron KE distribution simpler, and in fact, the FT-ICR instrument itself can be used to make the measurements. To this effect the rear trapping plate (PV2) is employed as a retarding voltage. The voltages of the other electrodes in the ICR cell are maintained at 0 V, as well as the Einzel lens elements. At a sufficiently negative value of PV2 no electron current is measured on FOCL2, since the KE of even the most energetic electrons emitted from the filament cannot overcome the potential energy barrier. Evidently, this only refers to the z-direction velocity of the electron; velocity of the electron in the xy plane would result in higher cyclotron or magnetron radii due to the focussing effect of the magnetic field. The retarding voltage where all electrons are stopped serves as the zero KE reference point. As the value of PV2 is then incrementally 'increased' (i.e. made more positive), the current detected on FOCL2 increases, until no more increase in current is detected. For a filament voltage U⁻ of −0.5 V, a voltage of −2.4 V retards the electrons and no current is observed. Since the static trapping ECD experiments are performed at voltages of +0.5 V, the resulting effective KE of these electrons

71

when entering the ICR cell is 2.9 eV [0.5-(-2.4)]. The measured current profiles on FOCL2 (as a function of PV2, shown in Figure 2.16) are differentiated and adjusted to +0.5 V trapping plate voltages employed in the static trapping ECD experiments to yield the retarding field analysis in Figure 2.17. This procedure is repeated for values of U^- of +0.5 V and +1.5 V.
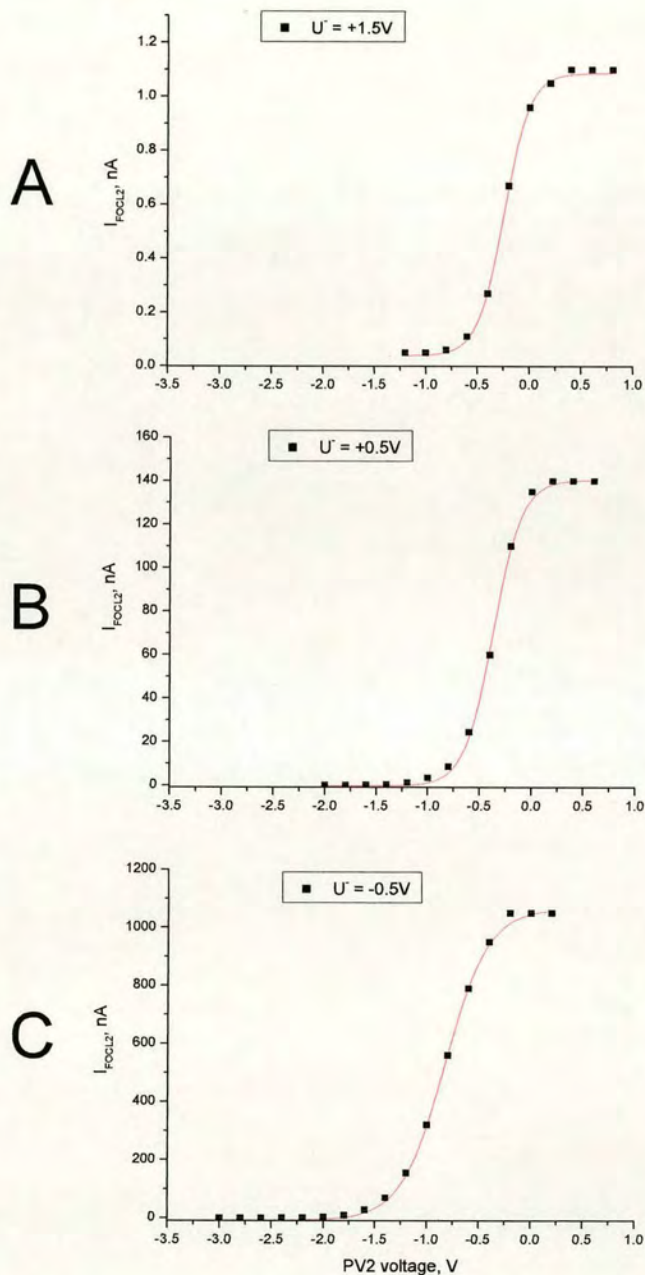

Figure 2.16 Rhenium filament emitted currents detected on FOCL2 (I_{FOCL2}) as a function of the retarding voltage on PV2. (A) $U^- = +1.5V$, (B) $U^- = +0.5V$ and (C) $U^- = -0.5V$. Red line indicates best sigmoidal fit. Note that I_{FOCL2} scales are different for each emission curve.
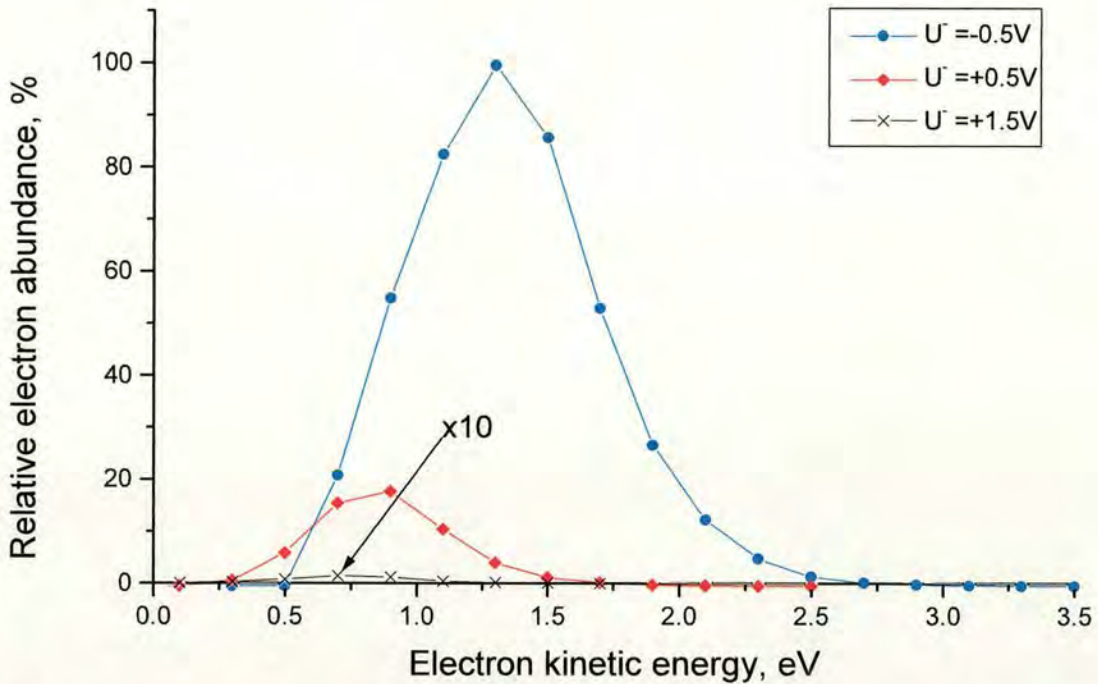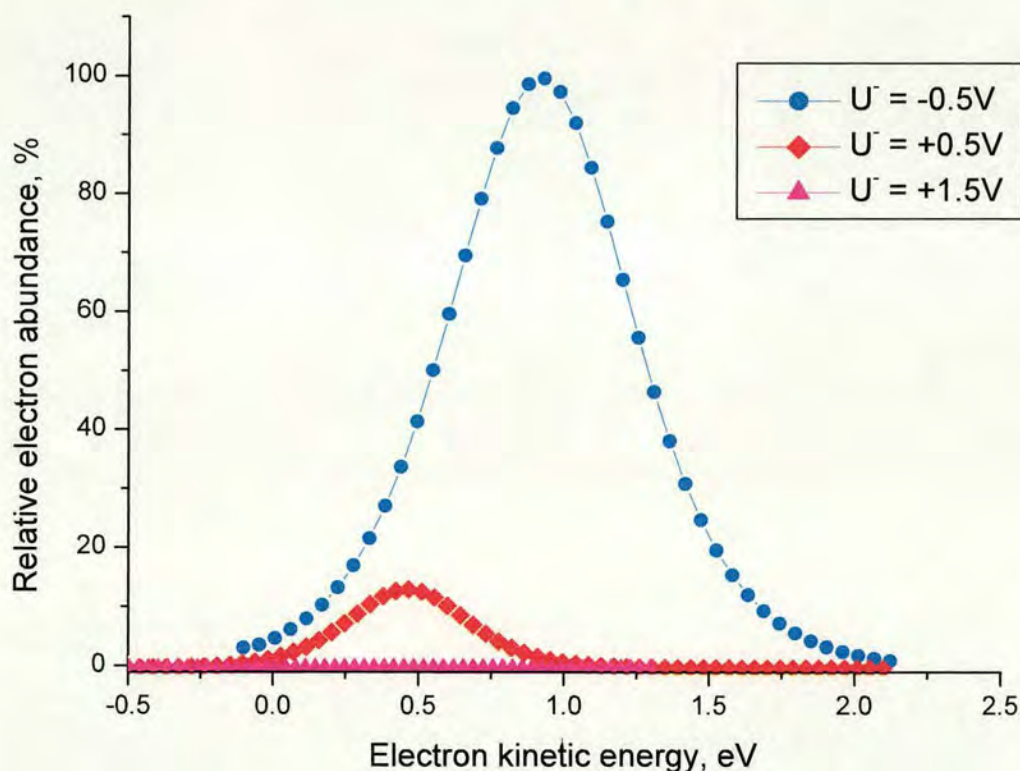
Figure 2.17 Effective kinetic energy distributions at the rear trapping plate PV2 (held at +0.5 V during static trapping ECD experiments) for different values of filament voltage U⁻. The relative electron abundance is normalised with respect to the maximum electron current emitted at a particular value of U⁻. The intensity of the relative electron abundance for U⁻ = + 1.5 V has been multiplied by a factor of 10 for clarity [18].

At a filament voltage U⁻ = -0.5 V the most abundant electrons have an effective KE of 1.3 eV, whereas at a filament voltage of U⁻ = +0.5 V, the most abundant electrons have an effective KE of 0.9 eV. The effective KE is lower for more positive values of U⁻. The most abundant electrons in the effective KE distribution of U⁻ = +1.5 V is even lower, however, the number of electrons is very low as well. The effective KE distribution for U⁻ = +0.5 V contains more low-energy electrons than is the case for either of the other two values of U⁻. Hence, it appears that a value of U⁻ = +0.5 is the best trade-off between obtaining low-energy electrons and sufficiently high electron currents to perform ECD.

The most effective approach to obtain even lower energy electrons in the ICR cell with the current set-up is to use even lower trapping plate potentials. In the dynamic trapping ECD experiments, the trapping plate voltages were lowered to +0.1 V. The retarding field data in Figure 2.16 was also analysed using a different

procedure, where the sigmoidal best fit lines in Figure 2.16 were differentiated; here the kinetic energy axis was corrected for trapping plate potentials of +0.1 V used during the dynamic trapping ECD experiments (see Figure 2.18). The advantage with differentiating the sigmoidal best fit line is that the electron KE distribution always has a Gaussian shape. Further, this procedure allows for a statistical evaluation of the mean kinetic energy of the distribution. Since only the mean KE can be determined with a degree of accuracy, whereas especially the low-energy tail of the distribution is not, differentiation of the sigmoidal fit line will be the method of choice for the remainder of this study. The mean KE of the electrons for $U^- = -0.5$ V is 0.919 ($\pm$ 0.006) eV, and in the case of $U^- = +0.5$ V the mean KE of the electrons is 0.465 ($\pm$ 0.006) eV. Considerably better ECD spectra were obtained using the lower trapping voltages, which could thus be explained. Furthermore, the fact that no ECD was observed at trapping plate voltages of 1 V or higher is consistent with this view.

Figure 2.18 Effective kinetic energy distributions at the rear trapping plate PV2 (held at +0.1 V during dynamic trapping ECD experiments) for different values of filament voltage U^-, as determined by differentiating sigmoidal best fit lines in Figure 2.16. The relative electron abundance is normalised with respect to the maximum electron current emitted at a particular value of U^-.

Many experimentalists advocate the use of a pulsed gas prior to electron irradiation to decrease the mean KE distribution of the emitted electrons. A retarding field analysis was carried out to determine if the use of a pulsed gas could really collisionally cool the electrons. A pulsed gas of argon was applied in the ultra-high vacuum (UHV) region, which caused the pressure to increase from 1×10^{-10} mbar to about 3×10^{-6} mbar, immediately prior to electron irradiation. An exact determination of the maximum pressure is difficult to ascertain, furthermore the pressure exponentially decreases and hence the pressure would not be stable over the length of electron irradiation (1 s). Once again a sigmoidal fit was performed on the plot of I_{FOCL2} vs PV2 (see Figure 2.19), and the best fit line was differentiated and the kinetic energy axis was corrected for trapping plate potentials of 0.1 V used during the ECD experiment (see Figure 2.20).
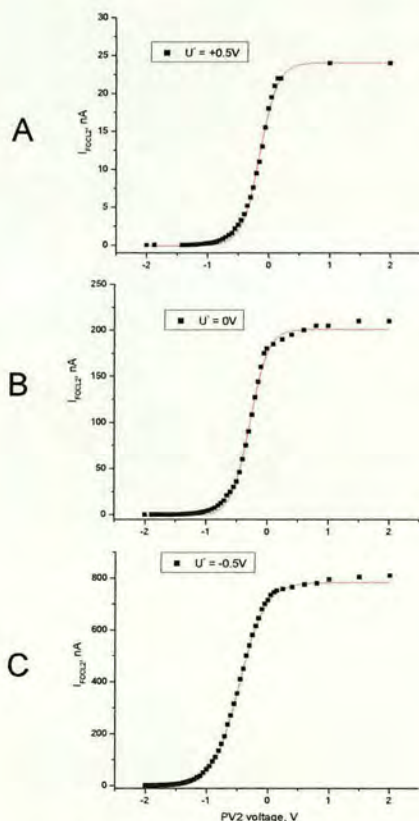
Figure 2.19 Rhenium filament emitted currents detected on FOCL2 (I_{FOCL2}) as a function of the retarding voltage on PV2, when pulsing Ar as in UHV. (A) $U^- = +0.5V$, (B) $U^- = 0V$ and (C) $U^- = -0.5V$. Red line indicates best sigmoidal fit. Note that I_{FOCL2} scales are different for each emission curve.
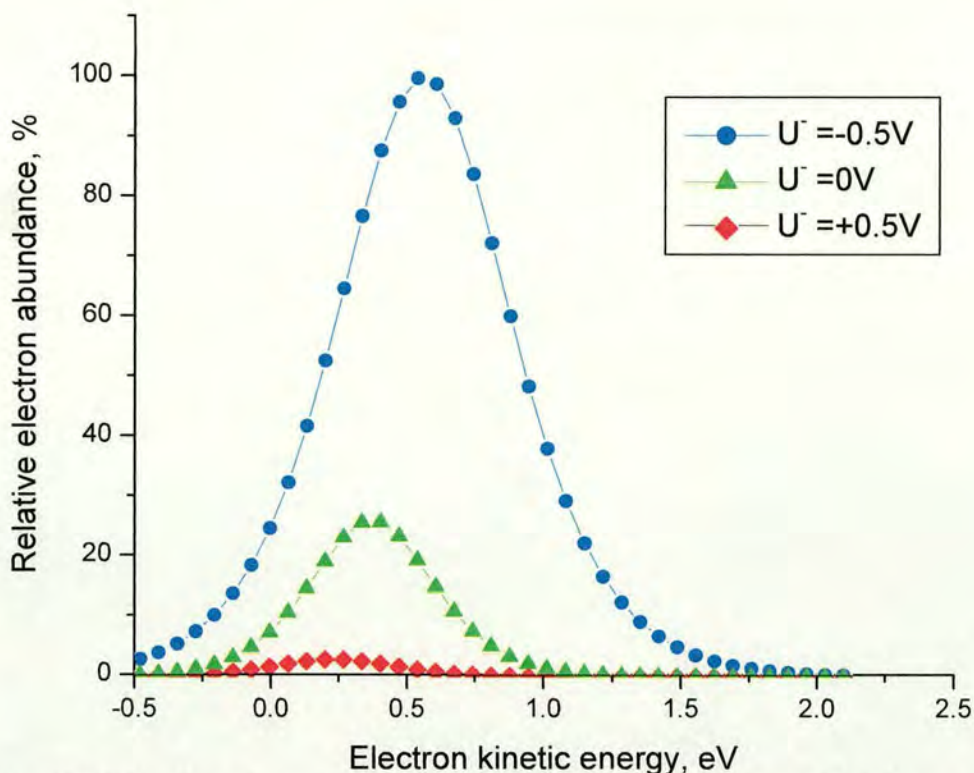
Figure 2.20 Rhenium filament effective electron kinetic energy at PV2 (for trapping voltages of 0.1 V), when employing a pulsed argon gas in the UHV.

In Figure 2.20, the mean KE energy distribution for U⁻ = -0.5 V were measured at 0.562 (± 0.002) eV, which compares to 0.919 eV measured when no pulsed gas was applied (Figure 2.18). The electron KE distributions for U⁻ = -0.5 V with and without a gas pulse are shown again in Figure 2.21, to demonstrate the effect of pulsing a gas. The maximum current was measured to be 810 nA (pulsed gas), compared to 1050 nA when no pulsed gas was used. This may suggest that loss of some electrons occurred or that the emission current was lower due to the higher pressure in the UHV chamber. The detected electron current on FOCL2 for U⁻ = +0.5 V is much reduced when applying a gas pulse, which is probably due to more significant cooling for low KE electrons. Electrons with a negative effective KE at PV2 in Figure 2.20 would not be able to overcome the retarding field and would consequently not enter the cell. The width of the KE distribution for U⁻ = -0.5 V in Figure 2.20 is ~0.75 eV, compared to ~0.74 eV for U⁻ = -0.5 V in Figure 2.18. Hence it appears that the use of a pulsed gas can decrease the mean KE of emitted electrons but does not narrow their KE distribution.
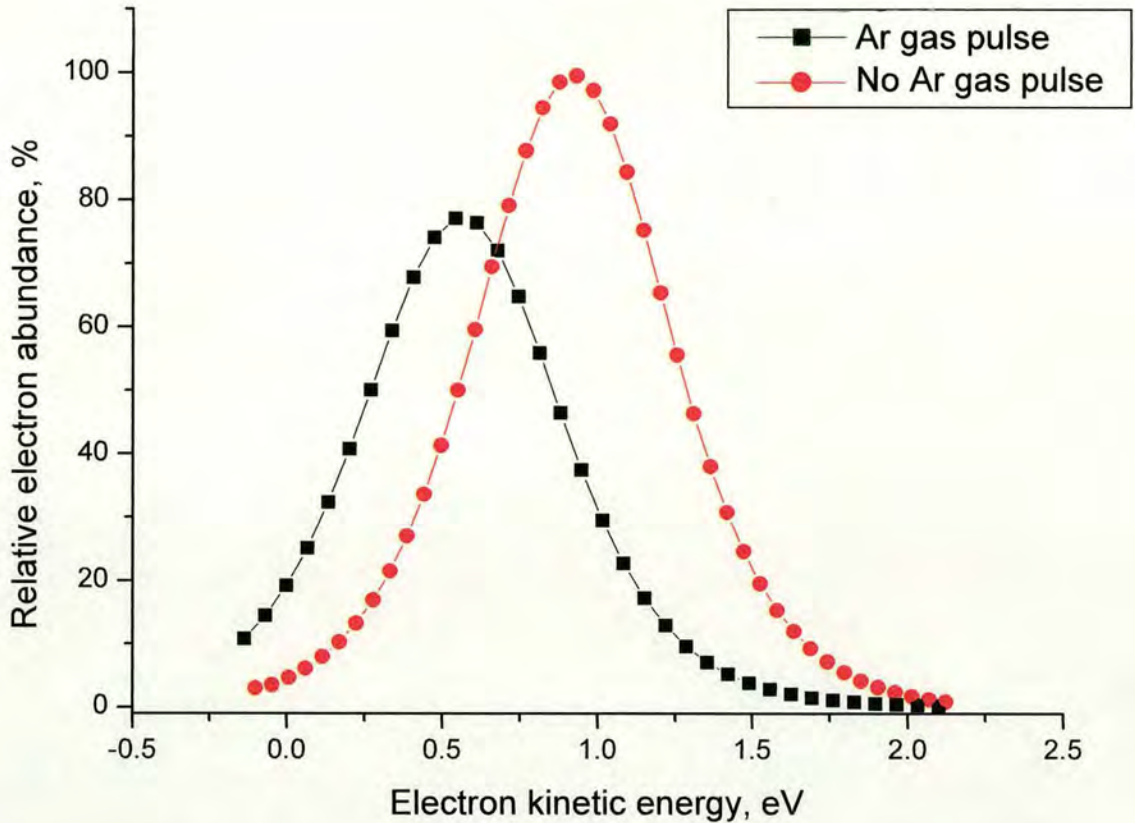
Figure 2.21 Rhenium filament effective electron kinetic energy at PV2 (for trapping voltages of +0.1 V) for U⁻ = -0.5V, with and without employing a pulsed argon gas in the UHV.

As mentioned before, there are limitations to the accuracy of this retarding field analysis especially regarding the low-energy tail of the KE distribution. In this case positive values of PV2 yielded higher current readings on FOCL2 (I_{FOCL2}), in spite of the fact that the remaining electrodes were set to 0 V, and therefore no increase should have occurred at values of PV2 higher than 0 V. One of the assumptions made when doing a retarding field analysis is that the voltage applied on PV2 has no effect on the emitted electron current from the filament. This hypothesis is probably viable since PV2 and the filament are 13 cm away in space, whereas the repeller plate that creates the local bias for pulsed emission is only 1 cm away from the filament. Nevertheless, the fact that positive values of PV2 gave higher current readings than obtained for 0 V indicates that the retarding field analysis is not accurate at values close to 0 V, which is equivalent to the low-energy tail of the electron KE distribution.

2.3.4.2 Dispenser cathode electron emission

According to the retarding field analysis, a large number of electrons can be generated with a conventional rhenium filament with sufficiently low KE to give rise to 'good' ECD. Nevertheless, the ECD efficiencies of spectra obtained with the filament are generally poor due to the limited overlap between the ion cloud and the electron beam. The emitting area on a filament is believed to be very small ($\ll$ 1 mm^2) [30], whereas the ion cloud is radially spread further. In contrast, the high-surface dispenser cathode fitted on the 9.4 Tesla instrument has an emitting surface of 19.6 mm^2 (diameter 5 mm), which is much closer to the radial displacement of the ion cloud off the central axis in the ICR cell. The emitting surface is coated with BaO which has a reduced work function (ϕ_o = 1.7 eV) [34], and hence a high number of electrons can be emitted at lower operating temperatures. The normal operating temperature of a filament is on the order of 2500 K, whereas the temperature of an oxide coated cathode is around 1200 K. Emission from the coated BaO surface occurs as a secondary electron emission process, due to electrons emitted from the heating filament impacting onto the BaO surface.

The emitted current retarding field plots in Figure 2.22 show that for more negatively biased filament voltages U⁻ higher electron currents could be emitted. The maximum current emitted for a filament voltage U⁻ = -0.5 V was 27000 nA (i.e. 27 µA), which compares to 1400 nA for similar filament voltages when using the filament. The total number of electron emitted is almost 20 times larger than for a filament and this emission occurs over a much wider area. This compares to maximum reported electron currents for similar cathodes of 2 mA [30]. The lines in Figure 2.22 indicate the best sigmoidal fits, which were used to determine the effective KE of the electrons for trapping voltages of 1 V (as used in most cathode ECD experiments). Figure 2.23 shows the resulting effective KE distributions.
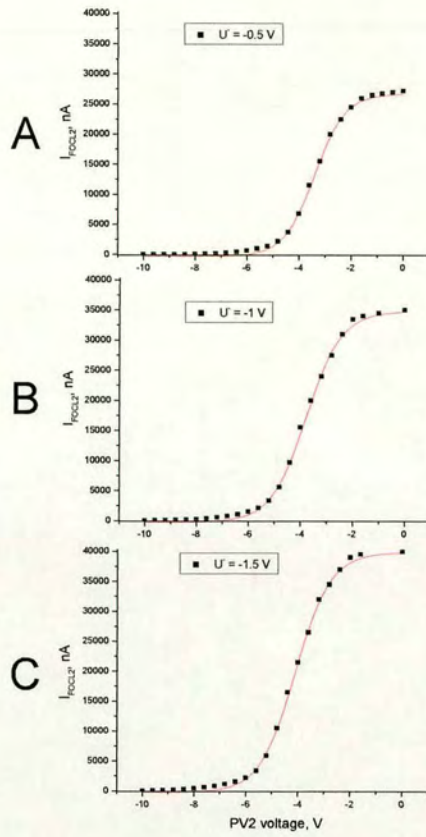
Figure 2.22 Emitted currents detected on FOCL2 (I_{FOCL2}) as a function of the retarding voltage on PV2. (A) $U^- = -0.5$ V, (B) $U^- = -1.0$ V and (C) $U^- = -1.5$ V.
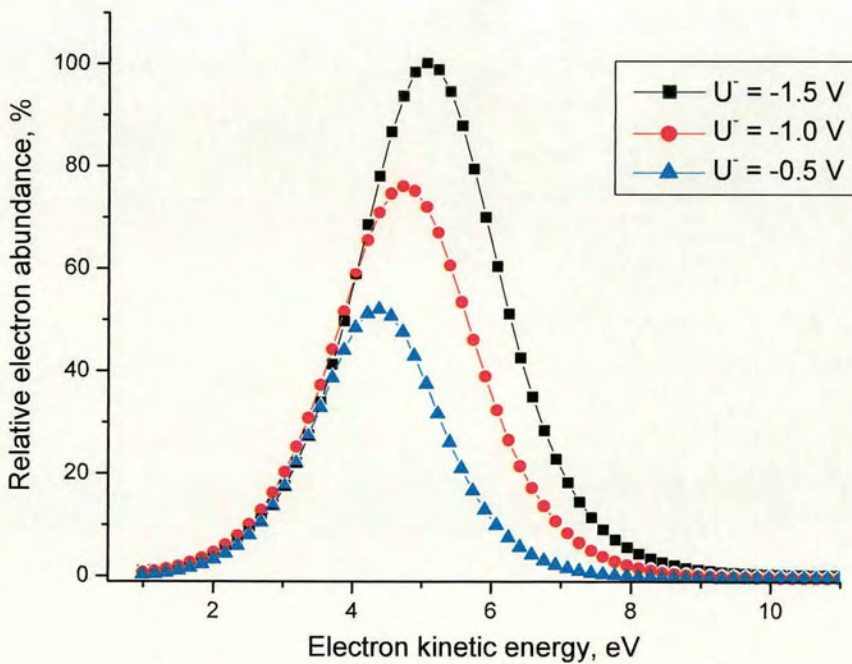

Figure 2.23 Effective kinetic energy distributions at PV2 (= 1V) for different filament voltages U^-.

The effective kinetic energy distributions at PV2 in Figure 2.23 show a decrease of the mean KE as the filament voltage U⁻ becomes more positive, as could be expected. The KE spread at half-height for $U^- = -0.5$ V is 2.27 eV, which is higher than for the filament at that value of U⁻ (0.74 eV, Figure 2.19). Generally, the KE spread measured in these retarding field experiments should be related to the Maxwell-Boltmann distribution, which in turn depends on the temperature of the emitting surface. The cathode is operated at a lower temperature than the filament ($T_{cathode} \sim 1200$ K, $T_{filament} \sim 2500$ K), and hence the emitted electron KE spread for the cathode should be narrower than for the filament. When performing ECD the filament voltage setting in the range of U⁻ -0.5 V to −1.5 V did not seem to have an effect on how good the precursor to parent ion conversion was. Given that the mean KE of the distributions in Figure 2.23 are elevated ($U^- = -0.5$ V: 4.39 eV; $U^- = -1.0$ V: 4.73 eV; $U^- = -1.5$ V: 5.07 eV) and do not seem to contain many low-energy electrons, it is somehow odd that ECD took place at all in these conditions. Chan *et al.* have approximated the electron kinetic energy for efficient ECD for electrons made by their rhodium filament to be in the range of 3-6 eV [23]. However, they have merely taken the average filament potential (i.e. average of voltages applied on either side of filament) during electron emission as the measure of electron KE. They did not take into account the work function of the filament. Further, their estimate of the electron KE does not constitute a measurement of the electron KE in the ICR cell.

In summary, the mean electron KE's as determined by this retarding field analysis for the rhenium filament seem to be in agreement with ECD efficiency measurements, which validates this approach to measure electron KE distributions. Nevertheless, for the dispenser cathode this retarding field analysis may not determine the 'real' KE spread of the electrons. This would also explain why the value of the trapping plate potentials seemed to have little effect on the ECD efficiency (trapping plate potentials at 3 V gave very similar ECD spectra to 1 V trapping voltages). In a retarding field analysis, the assumption is made that electrons in a beam behave as independent charged particles and that space-charge effects caused by electrons among one another are negligible. The emitted electron beam

from this dispenser cathode may be 5 mm wide, and hence electrons located in the centre of that beam may not 'see' the retarding field voltage. This could be due to the fact that the retarding voltage is only capable of 'peeling away' some of the layers of the electron beam, while electrons in the centre of the beam are not subjected to the retarding field voltage. Therefore, electrons in the centre of the electron beam would able to overcome a retarding field voltage in excess of their KE. This is consistent with the observation that a current of 100-200 nA was detected for the maximum retarding field voltage of −10 V. These results seem to imply that a retarding field analysis using this instrument is not suitable to determine the KE distribution of electrons emitted from a dispenser cathode.

2.3.5 Electron detachment dissociation

Electron capture dissociation is only applicable to positively charged peptides. For many polypeptides containing acidic residues (e.g. Glu, Asp) and acidic functions (e.g. C-terminal carboxylic acid, sulfation, N-acetylation, phosphorylation) it is not possible to generate a doubly protonated species and hence ECD is not feasible. The truly complementary technique to ECD would be the capture of a positron by a peptide anion. Still, positron sources are not available to the wider research community. Electron detachment dissociation (EDD) has been proposed as an analogue technique to positron capture [35]. In EDD peptide anions are irradiated with electrons at higher kinetic energies (i.e. ~21 eV), causing abstraction of one electron (i.e. oxidation) (see scheme 2.1).

$$[M-nH]^{n-} + e- \rightarrow [M-nH]^{(n-1)-\bullet*} + 2\ e^{-}$$ **Scheme 2.1**
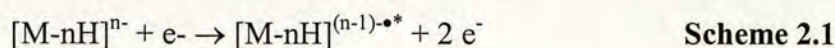
The resulting electronically excited oxidised species $[M-nH]^{(n-1)-\bullet*}$ can fragment according to an 'electron'/'hole' recombination [35]. Similarly to ECD, the precursor ion needs to be doubly charged (at least), since the charge is reduced by one in the electron detachment process.

Figure 2.24 shows the EDD mass spectrum of the doubly deprotonated species of a standard peptide. This spectrum was obtained by electron irradiation from the rhenium filament (1 s) at a filament voltage U^- of −20 V. The oxidised species $[M-2H]^{-\bullet}$ can be seen in this spectrum, as well as a loss of 44 Da from [M-

2H]⁻•. The 44 Da neutral loss constitutes an abundant fragmentation channel, which has also been observed by Budnik *et al.*, and which they have associated with the loss of CO_2 according to scheme 2.2.

$$R^-\text{-COO}^- + e^- \rightarrow R^-\text{-COO}^{\bullet-*} + 2\,e^- \rightarrow R^{-\bullet} + CO_2$$ **Scheme 2.2**
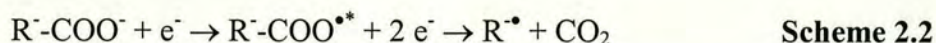
The fragments in Figure 2.24 are mostly due to the cleavage of the C_α-C bond (*a* fragments and x_5^- fragment; see Figure 1.9 Chapter 1), which may be due to differences in the mechanism compared to ECD. All fragments include the phosphorylated tyrosine residue, indicating that it is the charge carrier.
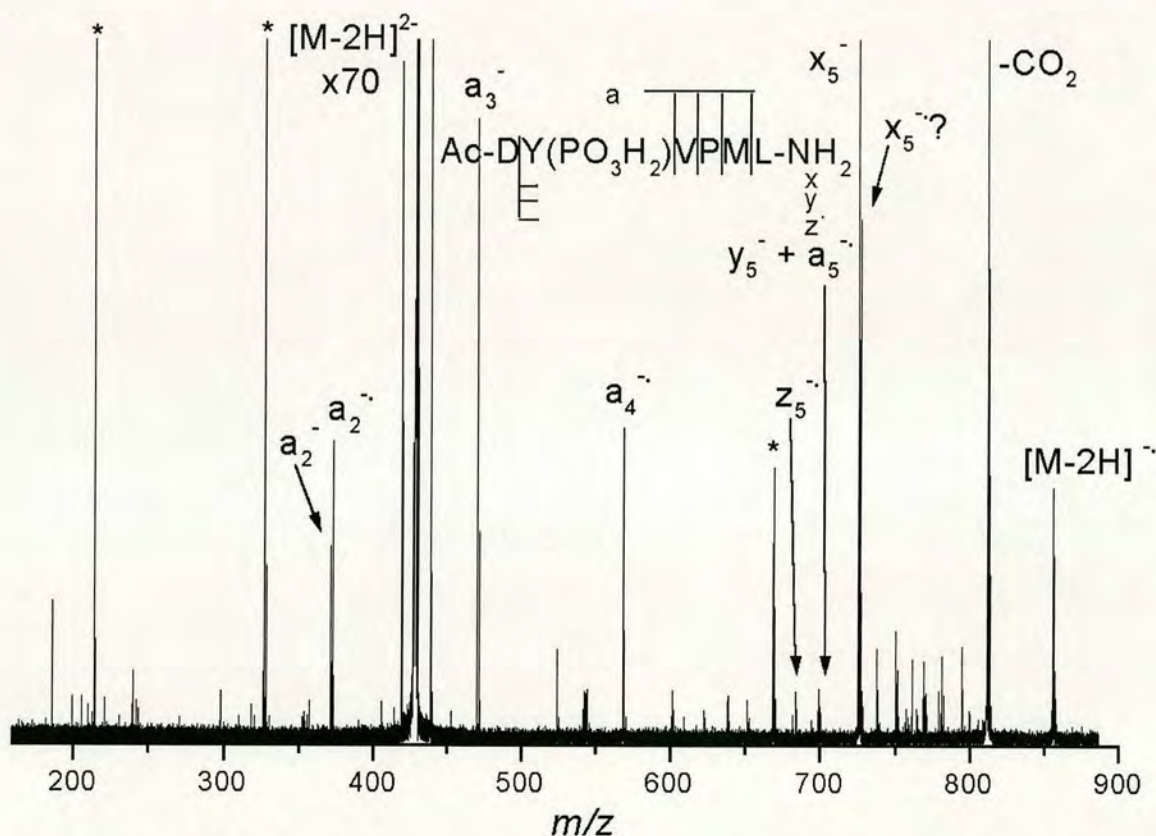
Figure 2.24 EDD mass spectrum of 2⁻ charge state of peptide with sequence Ac-DY(PO₃H₂)VPML-NH₂. Electron irradiation 1 s; 100 scans were accumulated. Asterisks indicate overtones and artefact peaks.

Figure 2.25 shows the EDD mass spectrum of the 2⁻ charge state of hirudin (sulphated). Apart from the CO_2 neutral loss, another loss of 70 Da is observed, which has been associated with the loss of SO_3. Limited fragmentation is observed

(x_8 to x_{11}). The preferred cleavage of the C_α-C bond is again confirmed in this spectrum.

The filament voltage U⁻ was observed to cause EDD if set to values of -20 V or higher (i.e. more negative). However, no fine tuning was required, and in fact, values of U⁻ = -30 or even –40 V, resulted in similar EDD spectra (not shown). It appears that in EDD the exact KE of the electrons plays less of a role, as long as the electrons have a minimum onset energy.

The cross-section for electron/ion interaction in EDD is reduced respective to ECD, since both the electron and the anion are negatively charged. This also explains why EDD efficiency on average is lower than the corresponding ECD efficiency.
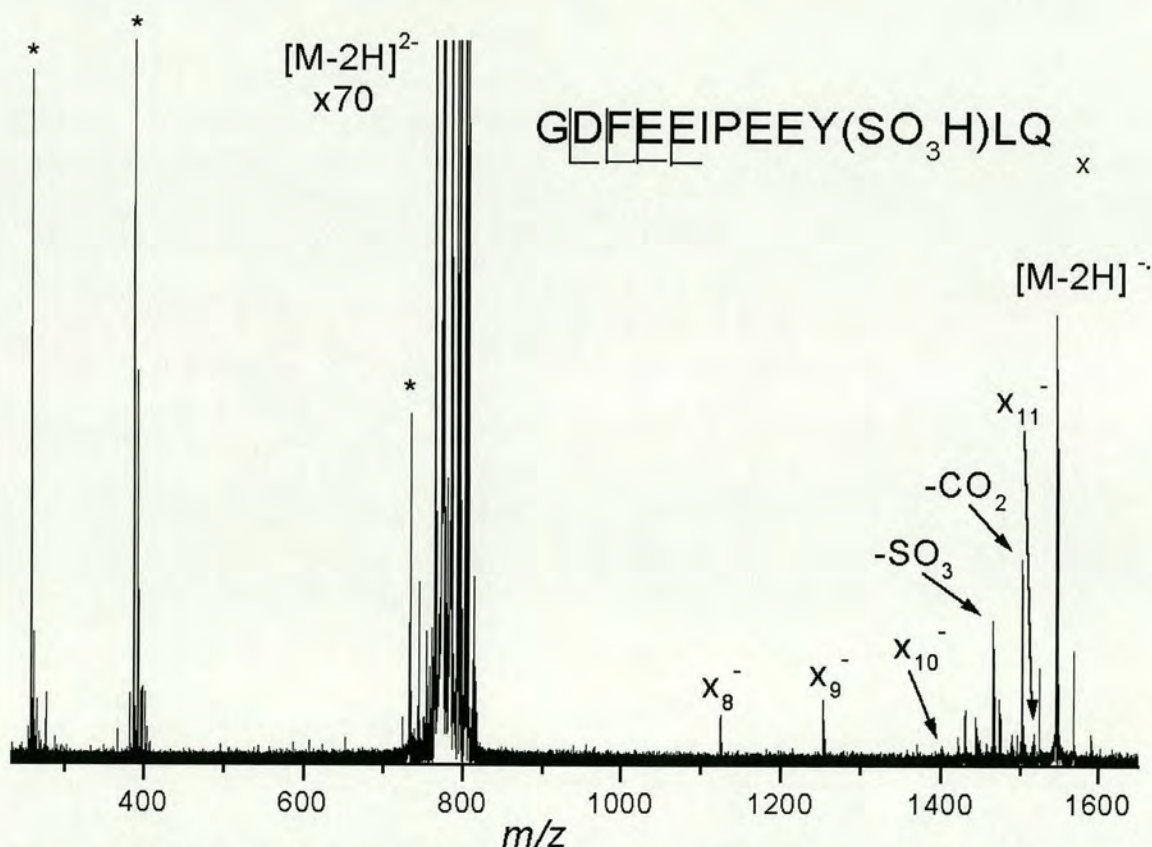
Figure 2.25 EDD mass spectrum of 2⁻ charge state of hirudin (sulphated). Electron irradiation 3 s; 200 scans were accumulated. Asterisks indicate overtones and artefact peaks.

2.4 Conclusions

Preliminary ECD results using the electron filament indicate that relatively good ECD efficiency can be obtained by using a conventional rhenium filament by lowering the trapping plate potentials to very low values (0.1 V) prior to electron irradiation. A 'retarding field analysis' of electron emission suggests that low trapping plate voltages give a higher abundance of low-energy electrons, thereby improving ECD efficiency. Nevertheless, the very small diameter of the electron beam emitted from a filament limits the extent of overlap between the ion cloud and the electron beam. Mass isolation in the filament ECD experiments often had to be compromised, in order to achieve any ion cloud/electron beam overlap. Consequently, fragments due to other processes, such as vibrational excitation, are observed in these spectra. Conversely, in the dispenser cathode ECD experiments the abundance of fragments was found to be very high and the mass isolation did not have to be compromised due to the better electron beam/ion cloud overlap. Competing fragmentation processes to ECD can be minimised by using the dispenser cathodes and almost 'pure' ECD spectra can be obtained, which enables the study of mechanistic processes in ECD (which is shown in Chapter 3). The retarding field analysis was not able to determine the real kinetic energy distribution of emitted electrons from dispenser cathodes, probably due to the fact that only some of the outer electrons in the beam are affected by the retarding voltage. Hence, the values of the trapping voltages do not seem to play a role in controlling ECD efficiency when using dispenser cathodes.

The study on EDD showed that preferential cleavage occurs at the C_α-C bond, as opposed to the N-C_α bond in ECD. EDD could be a promising electron-mediated sequencing technique for peptides where a 2+ charge state is not possible.

References

(1)   Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325-1337.

(2)   Lawrence, E. O.; Livingston, M. S. *Phys. Rev. Lett.* **1932**, *40*, 555.

(3)   Yamashita, M.; Fenn, J. B. *J. Am. Chem. Soc.* **1984**, *88*, 4451-4459.

(4)   Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Proc.* **1987**, *78*, 53-68.

(5)   Yoshida, T.; Tanaka, K. In; Patent: Japan, 1987; p 4.

(6)   Alford, J. M.; Weiss, F. D.; Laaksonen, R. T.; Smalley, R. E. *J. Phys. Chem.* **1986**, *90*, 4480-4482.

(7)   Chen, F. F. *Introduction to Plasma Physics*; Plenum Press:, 1974.

(8)   McIver, R. T. *Int. J. Mass Spectrom. Ion Proc.* **1990**, *98*, 35-50.

(9)   Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26*, 514-518.

(10)   Baykut, G.; Jertz, R.; Witt, M. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1238-1247.

(11)   Wilcox, B. E.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1304-1312.

(12)   Caravatti, P. In; Patent: USA, 1996.

(13)   Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

(14)   Ledford, E. B.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **1984**, *56*, 2744-2748.

(15)   Loo, J. A.; Udseth, H. R.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 207.

(16)   Senko, M. W.; Speir, J. P.; McLafferty Fred, W. *Anal. Chem.* **1994**, *66*, 2801-2808.

(17)   Guan, Z.; Kelleher, N. L.; O'Connor, P. B.; Aaserud, D. J.; Little, D. P.; McLafferty, F. W. *Int. J. Mass Spectrom. Ion Proc.* **1996**, *157/158*, 357-364.

(18)   Polfer, N. C.; Haselmann, K. F.; Zubarev, R. A.; Langridge-Smith, P. R. R. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 936-943.

(19)   Cooper, H. J.; Hudgins, R. R.; Hakansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 241-249.

(20)   Haselmann, K. F.; Budnik, B. A.; Kjeldsen, F.; Polfer, N. C.; Zubarev, R. A. *Eur. J. Mass Spectrom.* **2002**, *8*, 461-469.

(21)   Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty Fred, W. *J. Am. Chem. Soc.* **1999**, *121*, 806-812.

(22)   Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563-573.

(23)   Chan, T. W. D.; Ip, W. H. H. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1396-1406.

(24)   Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250-4253.

(25)   Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.

(26)   Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.

(27)   Shi, S. D. H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19-22.

(28)   Hakansson, K.; Hudgins, R. R.; Marshall, A. G.; O'Hair, R. A. J. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 23-41.

(29)   Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993**, *218*, 1-27.

(30)   Tsybin, Y. O.; Hakansson, P.; Budnik, B. A.; Haselmann, K. F.; Kjeldsen, F.; Gorshkov, M.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1849-1854.

(31)   Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Anal. Chem.* **2001**, *73*, 2998-3005.

(32)   Hutson, A. R. *Phys. Rev.* **1955**, *98*, 889.

(33)   Cosslett, V. E. In *Introduction to Electron Optics*; Clarendon Press: Oxford, 1950.

(34)   Herrmann, G.; Wagener, S. *The Oxide Coated Cathode*; Chapman and Hall Ltd.:, 1951; Vol. 1.

(35)   Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.

Chapter 3
Mechanistic studies of electron capture dissociation

3.1 Introduction

Dissociation of peptides by mass spectrometry to determine the primary structure is generally effected by collisionally activated dissociation (CAD) [1]. Computer algorithms such as Sequest[1] assist in the identification of peptide sequences [2-4], by comparing the experimental tandem mass spectrum to predicted tandem mass spectra for putative sequences. However, due to incomplete knowledge of the fragmentation patterns much of the information contained in the mass spectra is not predicted by these algorithms. Hence, Wysocki and co-workers have used SID (surface-induced dissociation) to carry out mechanistic studies in order to improve the prediction rate of automated sequence analysis by mass spectrometry [5-9]. As explained in Chapter 1, SID is a vibrational excitation technique that allows controlled deposition of energy onto the ion. An analysis of doubly protonated tryptic digest peptides dissociated by collisionally induced dissociation (CID) in an ion trap (3-dimensional Paul trap) [10] has also shown that preferential cleavage rules can be established. Electron capture dissociation is a relatively novel dissociation technique [11], which has only been successfully implemented on FT-ICR instruments, hence far fewer mechanistic studies have been undertaken [12,13].

This chapter highlights mechanistic aspects of ECD which can yield further primary sequence information: neutral side-chain losses and ease of H atom transfer. The ease of hydrogen atom transfer is assumed to be dependent on the ease of abstraction from the protonated site and ease of acceptance of the backbone carbonyl, resulting in backbone cleavage. Further, a 'reverse' backbone alpha-cleavage is proposed to be in competition with 'normal' alpha-cleavage. Simple polypeptides were chosen with predictable protonation sites, and by modifying one amino acid at a

---

[1] http://fields.scripps.edu/sequest/index.html

time the effect on the fragmentation pattern can be observed. Results are discussed in terms of the *non-ergodic* mechanism in ECD and the proposed *hot hydrogen re-arrangement* mechanism. These results will also be discussed in terms of the secondary structure of the peptide ion in more detail in Chapter 4.

All ECD mass spectra shown in this Chapter were generated on the 9.4 Tesla APEX III instrument (Bruker Daltonics) using a dispenser cathode (for procedure see Chapter 2, section 2.2), unless specified otherwise. The peptides included in this study are summarised in Table 3.1.

Abbreviation	Description	Peptide sequence
[LHRH_HR]$^{2+\#}$	Luteinizing hormone releasing hormone	pEHWSYGLRPG-OH
[LHRH_RH]$^{2+*}$	idem	pERWSYWLHPG-OH
[LHRH_HH]$^{2+*}$	idem	pEHWSYWLHPG-OH
[LHRH_RR]$^{2+*}$	idem	pERWSYWLRPG-OH
[LHRH_HR_NH$_2$]$^{2+\#}$	idem	pEHWSYGLRPG-NH$_2$
[LHRH_(l)W]$^{2+\#}$	idem	pEHWSYWLRPG-NH$_2$
[LHRH_(d)W]$^{2+\#}$	idem	pEHWSY(d)WLRPG-NH$_2$
[BOMB]$^{2+\#}$	Bombesin	pEQRLGNQWAVGHLM-NH$_2$
[BK]$^{2+\#}$	Bradykinin	RPPGFSPFR-OH
[Ac_BK]$^{2+\#}$	N-acetylated bradykinin	Ac-RPPGFSPFR-OH

Table 3.1 Abbreviations used for doubly protonated peptide ions with corresponding sequences. For amino acid single letter code, see Table A2.2, Appendix 2. The most likely sites of protonation are the side-chains of the amino acids arginine and histidine which are red colour-coded.[2] Note that pEHWSYGLRPG-NH$_2$ refers to the mammalian form of LHRH.[3] (# refers to peptides obtained from Sigma Aldrich, * refers to peptides made by Dr. Kim Haselmann from the University of Southern Denmark by solid phase synthesis).

3.2 Neutral losses

The losses of amino acid side-chains as a diagnostic tool for the presence of amino acids in peptides was first proposed by Cooper *et al* [14], in analogy to immonium ions in high-energy collision-activated dissociation (CAD) [15].

---

[2] The gas-phase basicities of arginine (237 kcal mol^{-1}) and histidine (223.7 kcal mol^{-1}) are highest (see Table A2.2, Appendix 2). N-terminus protected in case of pyro-glutamic acid and therefore not a possible site of protonation. For [LHRH_HH]$^{2+}$, side-chain of tryptophan{6} could be a competing site of protonation, as discussed in section 3.2.

[3] Note: pE refers to pyro-glutamic acid, which is made from N-terminal glutamic acid by cyclisation of the NH$_2$ terminus with glutamic acid side-chain through elimination of water. For structure, see Figure A2.1, Appendix 2.

Haselmann *et al* have named this region the (M• –X) region (1 Da ≤ X ≤ 130 Da) [16]. Neutral losses are in general 1 Da in mass higher than the expected mass of the amino acid side-chains, implying hydrogen atom rearrangement and/or incorporation in the cleavage. Generally these losses are observed for basic amino acids such as arginine, histidine and lysine, which would also be expected to be protonated (see gas-phase basicities Table A2.2, Appendix 2). Since work has already been done on neutral losses [14,16] these will only be briefly discussed in this Chapter. The (M• – X) regions for luteinizing hormone releasing hormone (LHRH) variants [LHRH_(d)W]$^{2+}$ (Figure 3.1), [LHRH_HR_NH$_2$]$^{2+}$ (Figure 3.2) and [LHRH_HH]$^{2+}$ (Figure 3.3) are shown in the following Figures. Neutral losses for [LHRH_HR]$^{2+}$ (ECD mass spectrum in Figure 3.4) and [LHRH_RR]$^{2+}$ (ECD spectrum in Figure 3.9), for which the (M• –X) region is not specifically shown, will also be discussed. Losses are colour-coded in order to link them to particular amino acids or functional groups.
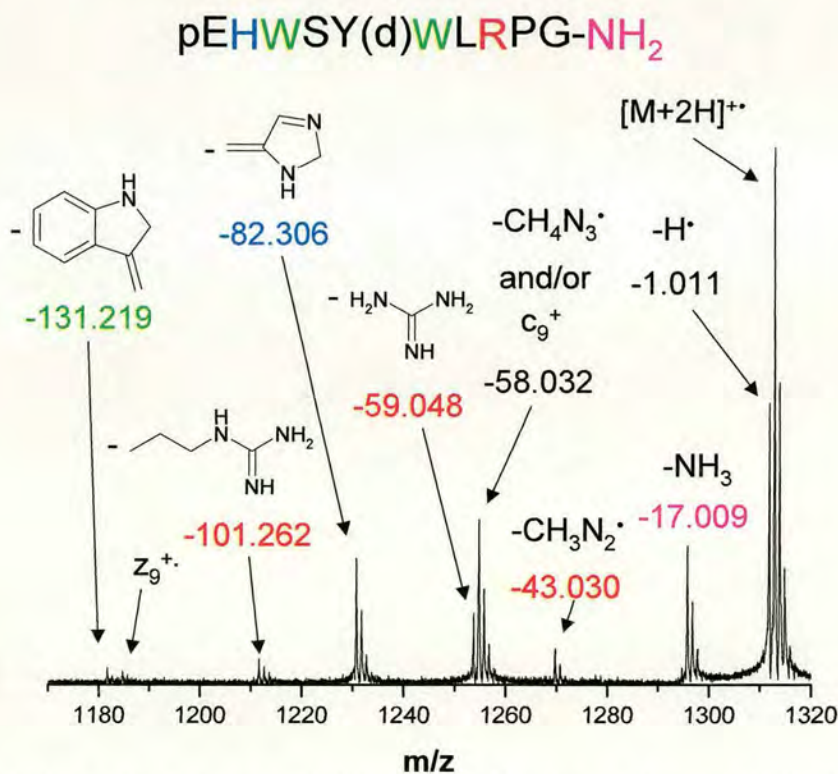
Figure 3.1 (M• –X) region for [LHRH_(d)W]$^{2+}$. The peaks are labelled with mass losses, the proposed chemical formulae of the losses and colour-coded with the associated amino acid residues. (electron irradiation: 35 ms)

89
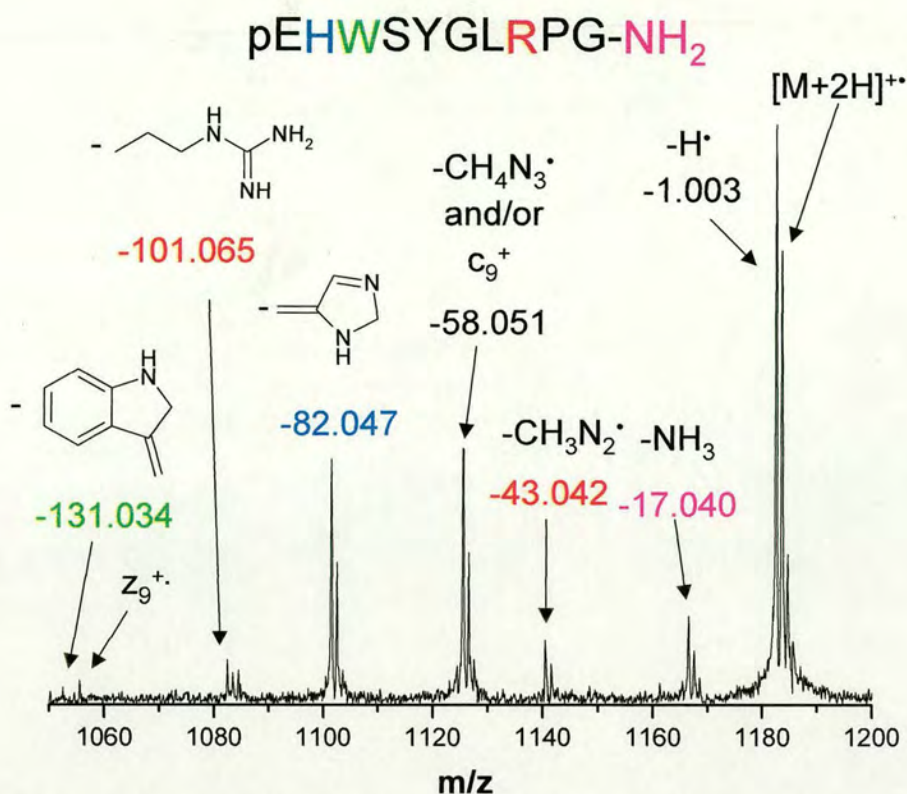
Figure 3.2 (M• –X) region for [LHRH_HR_NH$_2$]$^{2+}$. (electron irradiation: 30 ms)
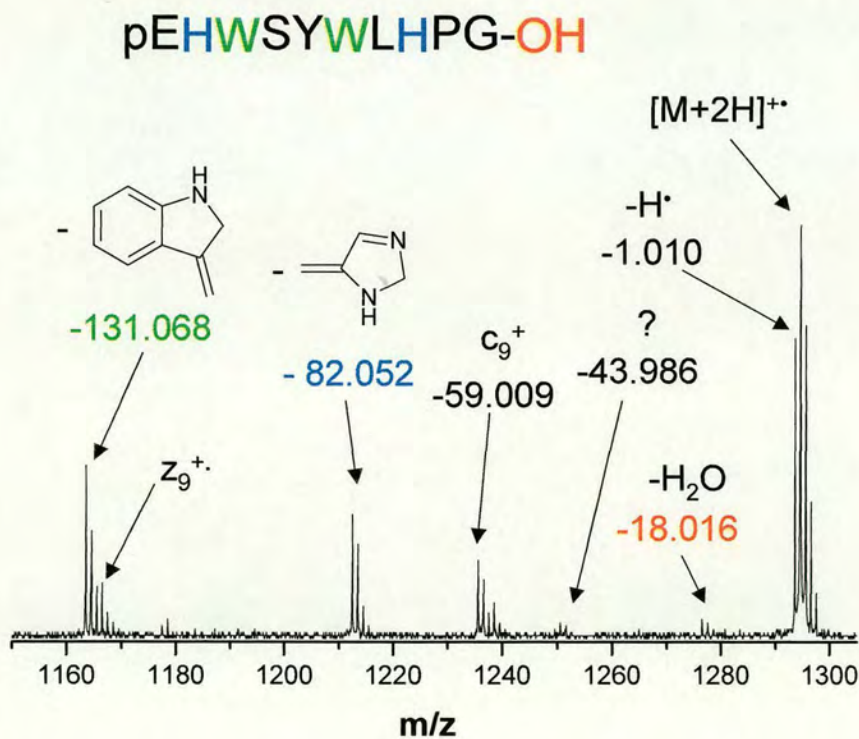
Figure 3.3 (M• –X) region for [LHRH_HH]$^{2+}$. Unidentified peaks are labelled with a question mark (?). (electron irradiation: 35 ms).

The masses of the proposed fragment losses in Figures 3.1 to 3.3 are summarised in Table 3.2. For most experimentally observed losses and the masses of the predicted fragments there is reasonable agreement.
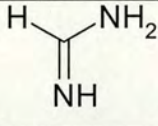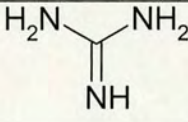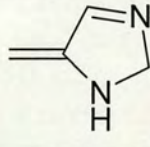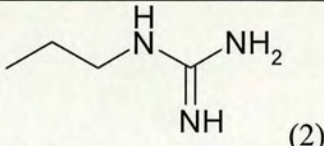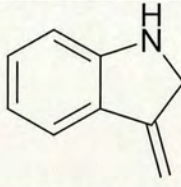
Fragment formula	Mass (Da)
H$^\bullet$ (1)	1.008
NH$_3$ (2)	17.027
H$_2$O (2)	18.011
CH$_3$N$_2$$^\bullet$ (2)	43.029
(2)	44.037
CH$_4$N$_3$$^\bullet$ (3)	58.041
(2)	59.048
(4)	82.053
(2)	101.095
(4)	131.073

Table 3.2 Masses of the predicted fragment losses in (M$\bullet$ –X) region. Structures of fragment ions first suggested by (1) [11], (2) [14], (3) this author and (4) [16].

Table 3.3 summarises the neutral losses observed in Figures 3.1 to 3.3. Each of these neutral losses is discussed with respect to its specificity and sensitivity for the presence of a particular amino acid or group.

Peptide sequence	17	18	43	59	82	101	131
pEHWSY(d)WLRPG-NH$_2$	Y	N	Y	Y	Y	Y	Y
pEHWSYGLRPG-NH$_2$	Y	N	Y	Y	Y	Y	Y
pEHWSYWLHPG-OH	N	Y	N	Y	Y	N	Y
pEHWSYGLRPG-OH	Y	N	Y	N	Y	Y	Y
pERWSYWLRPG-OH	Y	Y	N	Y	N	N	N

Table 3.3 Summary of neutral losses for Figures 3.1 to 3.3, 3.4 and 3.9.

1 Da loss. The hydrogen atom desorption, in contrast to the other neutral losses, cannot be associated with the presence of particular amino acids. Nevertheless, it can be a measure of how globular the structure of the peptide ion is [17]. Table 3.4 summarises the relative abundances of hydrogen atom desorption and the 'reduced' species [M+2H]$^{+\bullet}$. Hydrogen atom desorption is more pronounced for [LHRH_HR_NH$_2$]$^{2+}$, which may be due to more extended structures than for the other peptide ions.

Peptide sequence	[M+H]$^+$	[M+2H]$^{+\bullet}$
pEHWSY(d)WLRPG-NH$_2$	75 %	25 %
pEHWSYGLRPG-OH	75 %	25 %
pEHWSYGLRPG-NH$_2$	85 %	15 %
pEHWSYWLHPG-OH	45 %	55 %
pERWSYWLRPG-OH	40 %	60 %

Table 3.4 Relative abundances of hydrogen atom desorption ([M+H]$^+$) relative to 'reduced' species ([M+2H]$^{+\bullet}$). Abundances calculated by taking heavier isotope peak of [M+H]$^+$ into account.

17 Da loss. In spite of the fact that this loss is much more abundant in the case of amidated peptides, it is not specific to the presence of that group since it is also observed in the case of [LHRH_HR]$^{2+}$. The ammonia (NH$_3$) loss could also originate from the arginine side-chain and this loss may be isobaric with the loss of $^\bullet$OH (17.0027 Da) when the mass accuracy is insufficient.

18 Da loss. This loss is linked to water loss and is observed for the free acid C-terminal peptide. Nevertheless, this loss is of low abundance and therefore of lesser diagnostic value.

43 Da loss. The 43 Da loss is generally a good indication for the presence of an arginine residue. The loss of the radical CH$_3$N$_2$$^\bullet$ group from the arginine side-

chain, without incorporation of a hydrogen atom, seems to be specific to the LHRH peptides, Cooper *et al.* observed a 44 Da loss (see Table 3.2) for other arginine-containing peptides they studied.

59 Da loss. Due to the C-terminal amino acid being a glycine the 59 Da loss (loss of $(CH_2)_2C=NH$) is isobaric with the c_9^+ fragment ion in the case of free acid C-terminal peptides and therefore both cannot be distinguished. The 59 Da loss is not seen in the (M• −X) region of pEHWSYGLRPG-NH$_2$ (Figure 3.2), despite the fact that an arginine residue is present. Instead, a 58 Da loss is seen in the (M• −X) region of pEHWSY(d)WLRPG-NH$_2$ (Figure 3.1), which can be associated with the loss of $CH_4N_3^•$ from the arginine side-chain. This is in analogy to the 43 Da loss observed for LHRH peptides, for which no hydrogen atom incorporation into the neutral loss fragment occurs. The 58 Da loss seems to be specific for the presence of an arginine residue in LHRH peptides.

82 Da loss. This loss is associated with the loss of the histidine side-chain. It is abundantly observed every time that histidine is present and is not observed when histidine is not present. Hence, this 82 Da loss constitutes a true positive and negative check for the presence of histidine in the amino acid sequence. Unfortunately, the relative loss of 82 Da is not increased for the two histidine-containing LHRH peptide (i.e. pEHWSYWLHPG-OH) over any of the other peptides that only contain one histidine. While this loss is a qualitative indication for the presence of histidine, it does not seem to give a quantitative assessment of how many histidines are present.

101 Da loss. The 101 Da loss, similarly to the 43 Da loss, is a good predictive tool for the presence of the arginine residue. Importantly, it does not give a false positive result in the case of pEHWSYWLHPG-OH (see Figure 3.3). This loss, in conjunction with the 43 Da loss (or 44 Da loss in other peptides than LHRH), can be used as a diagnostic tool for the presence of the arginine residue.

131 Da loss. This loss gives a good indication for the presence of the tryptophan residue, since it cannot be confused with other neutral losses. However, for most of the studied peptides the abundance of the loss is quite low, which decreases its value as a predictive tool. The peptide pEHWSYWLHPG-OH (Figure

3.3) shows a very abundant 131 Da loss. In contrast to the other LHRH variants, this peptide has no arginine residue and hence the tryptophan side-chains are competing sites of protonation to the histidine side-chains.[4] This would also explain why the 82 Da loss due to the histidine side-chain is not more abundant for pEHWSYWLHPG-OH.

In summary, neutral losses in the (M• −X) region for the studied LHRH variants are shown to provide information for the presence of particular amino acids. This especially applies to the 82 Da loss linked to the loss of the histidine side-chain. The 43 and 101 Da losses are useful indicators for the presence of the arginine residue. Conversely, the 59 Da loss is slightly more ambiguous, especially due to the fact that for these peptides it can be confused with the c_9^+ fragment ion and similar ambiguities may arise in other sequences. The 131 Da loss, while generally of low abundance, can serve as a good indication for the presence of tryptophan. Further, a high abundance of this loss could be due to protonation of the tryptophan side-chain. On the other hand, the 17 Da loss is not specific enough and the 18 Da loss is of low abundance.

---

[4] The gas-phase basicity of tryptophan (216.1 kcal mol^{-1}) is relatively close to the gas-phase basicity of histidine (223.7 kcal mol^{-1}).

3.3 Ease of hydrogen atom transfer from protonated site

As explained in Chapter 1, the hypothesis in the *hot hydrogen re-arrangement* mechanism is that a slow electron (< 0.2 eV) is initially captured at the protonated site (i.e. H^+). The ensuing hydrogen radical ($H^\bullet$) is then transferred to a high affinity site for hydrogen atoms. The following study investigates the relationship between the gas-phase basicity of the protonated site and the ease of hydrogen atom transfer from that site.

3.3.1 ECD of $[LHRH\_HR]^{2+}$

The ECD mass spectrum of $[LHRH\_HR]^{2+}$ is shown in Figure 3.4. The most likely sites of protonation for this peptide ion are the side-chains of arginine and histidine, given their higher gas-phase basicities. The ECD process gives rise to the reduced species $[M+2H]^{+\bullet}$ and neutral losses, as discussed in section 3.2. Further, a series of z fragment ions (z_3^+ to z_9^+) can be seen in this spectrum as well as the c_9^+ fragment. All of these fragment ions include the arginine residue, which in terms of the *hot hydrogen re-arrangement* mechanism means that only electron capture at the histidine side-chain causes backbone cleavage. Considering that the gas-phase basicity of arginine is higher, this may rationalise why electron capture at that site is less likely to give rise to a 'mobile' hydrogen atom, in analogy to the *mobile proton* model proposed by Wysocki and co-workers [9]. Electron capture occurring on the arginine side-chain would then only result in the typical neutral losses associated with the arginine side-chain are observed (i.e. 43, 59 and 101 Da).

The insert in Figure 3.4 shows the detailed peak structure of the labelled z_5^+ fragment ions. This peak is in fact made up of two different peaks. One peak corresponds to the mass calculated for the $z_5^{+\bullet}$ fragment, however, the labelled $z_5^{+'}$ peak is not the ^{13}C isotope peak of $z_5^{+\bullet}$. The mass difference between these peaks is 1.00795 Da, corresponding to a hydrogen atom difference. Therefore, $z_5^{+'}$ is an even-electron z fragment ion, which will be discussed in detail in section 3.4. The appearance of even-electron z fragments has been documented in the literature [18-22], and the notation as proposed by Kjeldsen *et al* [23] is employed in this thesis.
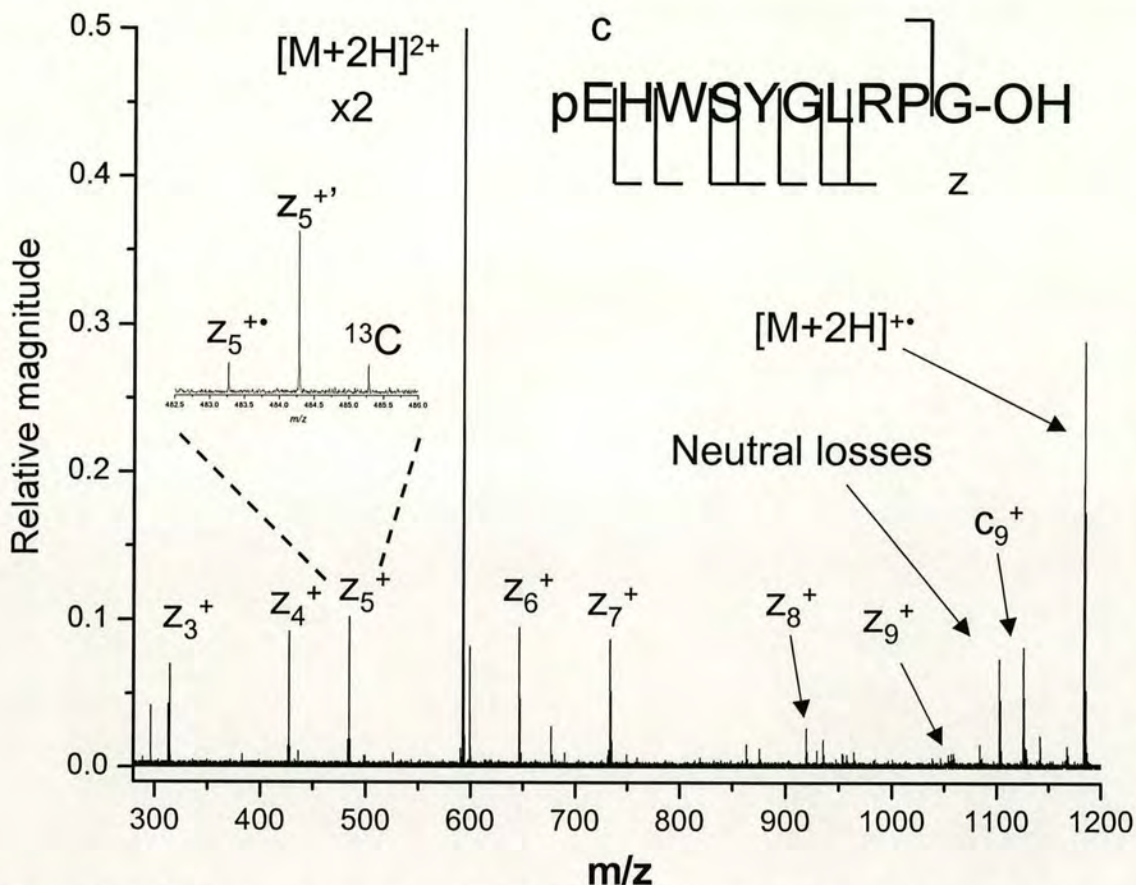
Figure 3.4 ECD mass spectrum of [LHRH_HR]$^{2+}$ showing z and c fragment ions. (100 time-domain transients) Insert shows detail of z_5^+ fragment ion. (electron irradiation 35 ms).

3.3.2 ECD of [LHRH_RH]$^{2+}$

To test the assumption that electron capture at the arginine side-chain gives rise to very inefficient backbone cleavage compared to electron capture at the histidine side-chain, electron irradiation was effected on [LHRH_RH]$^{2+}$. This peptide has reversed positions for the arginine and histidine residues compared to [LHRH_HR]$^{2+}$.[5] The ECD mass spectrum of [LHRH_RH]$^{2+}$ (see Figure 3.5) confirms that only charged ECD fragment ions are observed that also contain the arginine residue.

The second site of protonation for [LHRH_RH]$^{2+}$ could be either on the histidine side-chain in position 8 or the tryptophan side-chain in position 6.

---

[5] [LHRH_RH]$^{2+}$ also contains a tryptophan residue in position 6 instead of a glycine residue in the case of [LHRH_HR]$^{2+}$. This should, however, not affect the preferential protonation of the arginine side-chain.

Protonation of the tryptophan side-chain in position 3 is less likely due to Coulombic repulsion with the protonated arginine side-chain. Neutral side-chain losses for [LHRH_RH]$^{2+}$ are very weak; an 82 Da loss, linked to the histidine side-chain, is at a signal-to-noise ratio = 3, but no 131 Da loss (associated with the tryptophan side-chain) is seen. Since the gas-phase basicity of histidine is higher than for tryptophan and the distance between arginine and histidine (5 residues) is higher than for arginine to tryptophan in position 6 (3 residues), the histidine side-chain is likely to be the 2nd most preferred site of protonation in this molecule.
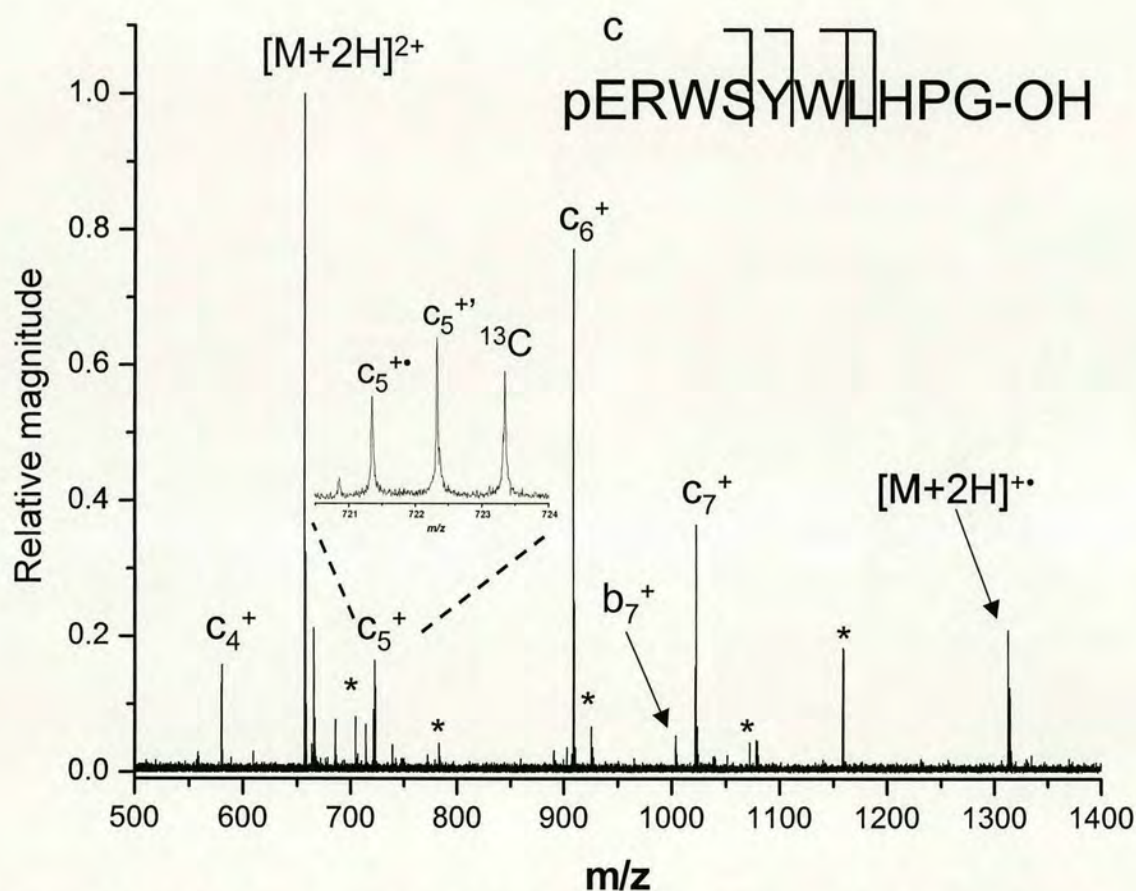
Figure 3.5 ECD mass spectrum of [LHRH_RH]$^{2+}$. (100 time-domain transients) Insert shows detail of c_5^+. (electron irradiation 35 ms) Starred (*) peaks due to instrument artefacts.

The c fragments in this mass spectrum also consist of two different types of fragments. The insert in Figure 3.5 shows the $c_5^{+\prime}$ fragment and a fragment which is 1.017 Da lower in mass, corresponding to the mass of a hydrogen atom, labelled as

$c_5^{+\bullet}$. Therefore, $c_5^{+\bullet}$ is an odd-electron species, which will be discussed in section 3.4.

3.3.3 SORI CAD of [LHRH_HR]$^{2+}$

In contrast, collisionally activated dissociation (CAD) of [LHRH_HR]$^{2+}$ (Figure 3.6) does not result in a higher abundance of C versus N-terminal fragment ions. According to the *mobile proton* model one would expect the proton on the arginine side-chain to be more fixed than the proton on the histidine side-chain [9]. Hence, the proton from the histidine side-chain would be expected to preferentially induce backbone cleavage. The resulting b and y fragments are in competition as to which one carries the proton depending on their proton affinities, as suggested by Nold *et al.* [24]. For b ions to be abundantly observed in this case, the proton needs to preferentially remain with that fragment. The complementary y fragment would be charged in this case due to the proton on the arginine side-chain. If all the protons (from the histidine side-chain) that induce backbone cleavage remain with the b fragment, one should observe equal fragment abundances for complementary fragments. However, the N-terminal fragment abundances are higher in some cases (e.g. b_2^+ vs. y_8^+, b_3^+ vs. y_7^+ and b_4^+ vs. y_6^+). This implies that either a proton is lost from some fragments (and hence they would not be observed), or that larger fragments further dissociate. The latter hypothesis is more likely, since smaller fragments are more abundant in the mass spectrum than larger fragments. Collisionally activated dissociation of these peptides does not seem to give information where the most basic amino acid in the peptide is located.

Conversely, in ECD of these peptides all the fragment ions include the most basic amino acid. In terms of the *hot hydrogen re-arrangement* mechanism this could be explained by the relative ease of hydrogen atom abstraction from the protonated site. Thus, the ease of hydrogen atom transfer from the protonated site would be inversely linked to the gas-phase basicity of the protonated site. In analogy with the *mobile proton* model, the hydrogen atom on the arginine side-chain is more localised.
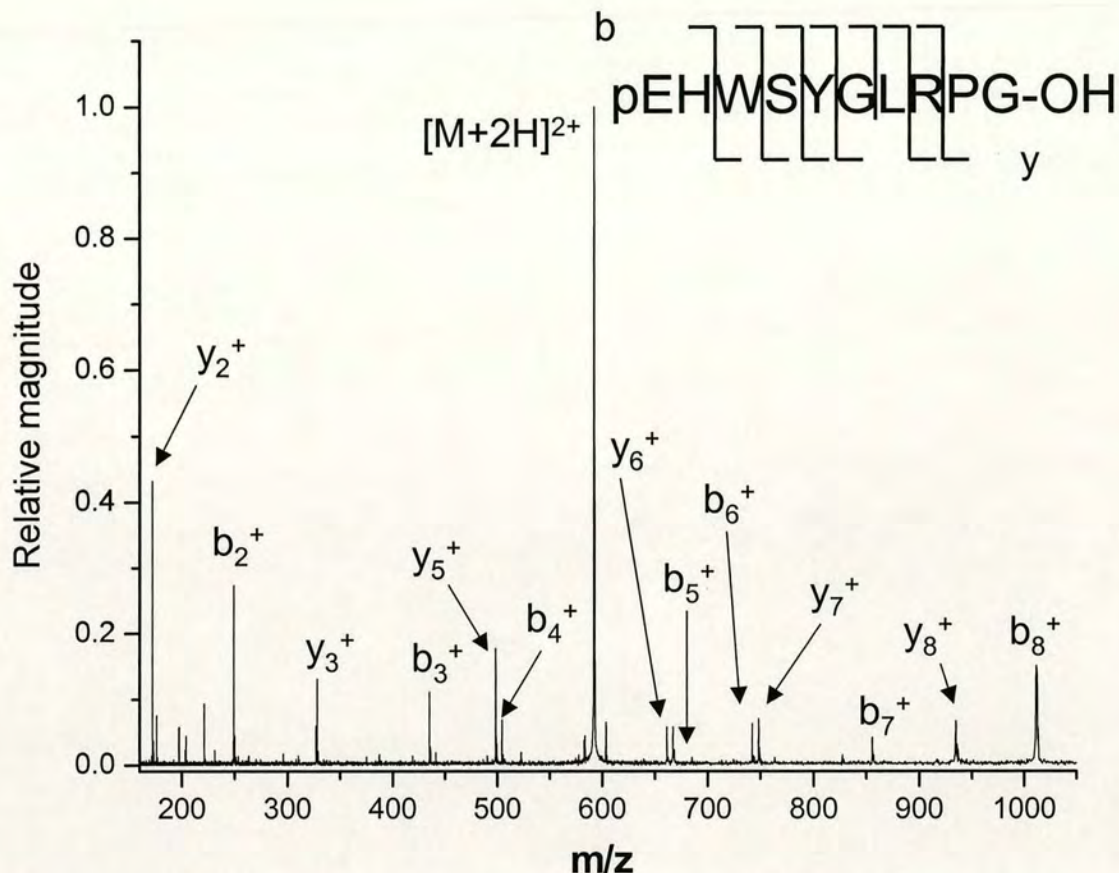
Figure 3.6 SORI CAD mass spectrum of [LHRH_HR]$^{2+}$. (100 time-domain transients, hexapole accumulation = 0.5 s)

3.3.4 ECD of [BOMB]$^{2+}$

To test this assumption further, the peptide ion [BOMB]$^{2+}$ was irradiated with electrons (see Figure 3.7). The most likely sites of protonation in the doubly charged peptide are the side-chains of arginine and histidine. Once again only ECD fragment ions are observed which include the arginine residue, thus confirming the hypothesis that electron capture at the histidine side-chain is more likely to result in backbone cleavage than capture at the arginine side-chain. This mass spectrum compares favourably to the ECD mass spectrum of [BOMB]$^{2+}$ shown by Chan et al. [25], where only c fragments were observed.
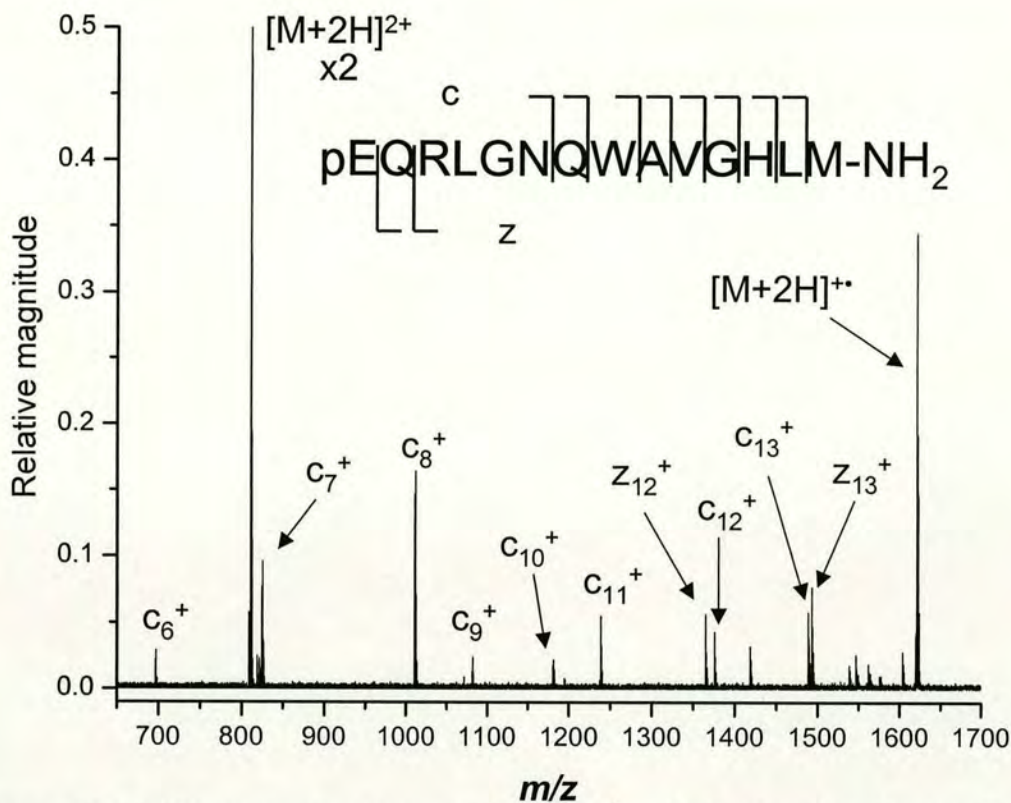
Figure 3.7 ECD mass spectrum of $[BOMB]^{2+}$. (100 time-domain transients) (electron irradiation = 50 ms).

3.3.5 ECD of $[LHRH\_HH]^{2+}$ and $[LHRH\_RR]^{2+}$

If the relative gas-phase basicities of the protonated amino acids in a peptide with basic residues on either end of molecule determine whether one observes primarily N or C-terminal fragment ions, the presence of two amino acids of the same gas-phase basicity should result in an even amount of N or C-terminal fragment ions. The ECD mass spectra of $[LHRH\_HH]^{2+}$ (Figure 3.8) and $[LHRH\_RR]^{2+}$ (Figure 3.9) both show c and z fragments. For $[LHRH\_HH]^{2+}$ the sites of protonation are more ambiguous as for any of the other peptides and hence the original sites of protonation could be on the histidine or tryptophan side-chains or even a mixture thereof. This means that a multitude of fragmentation channels are available.
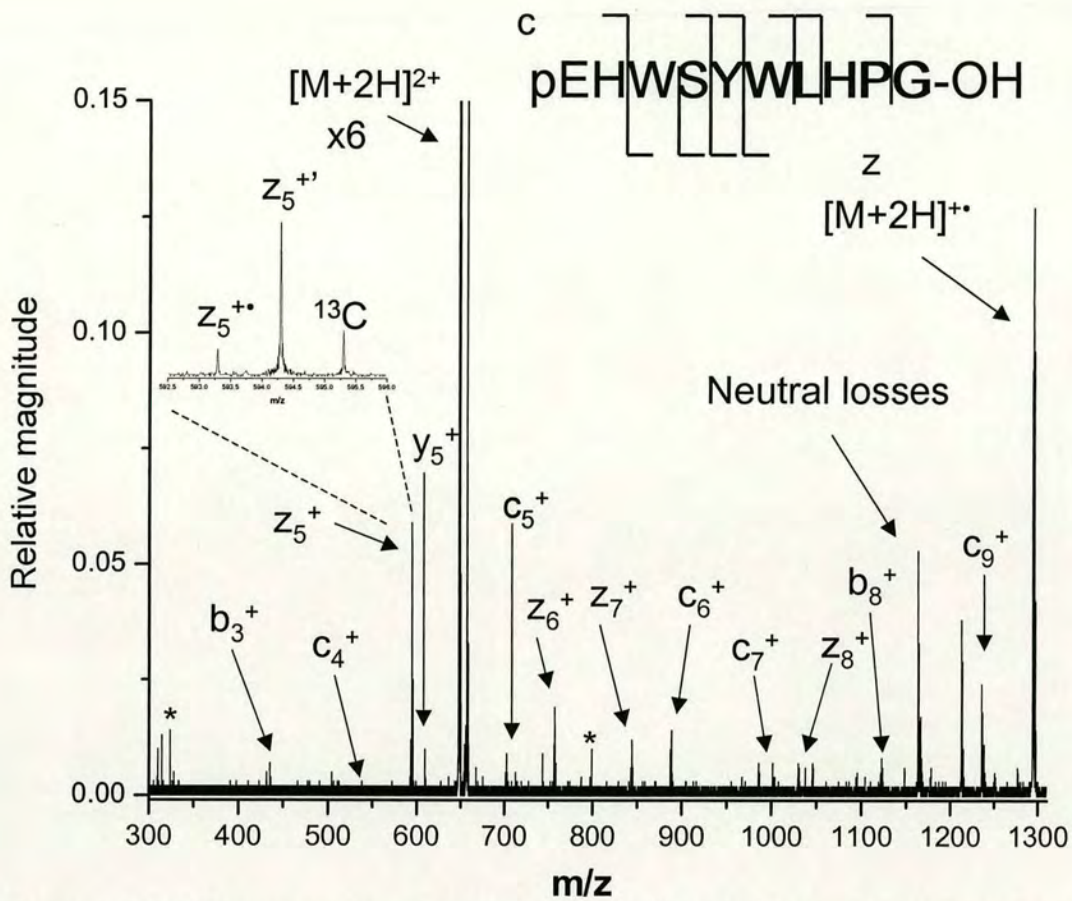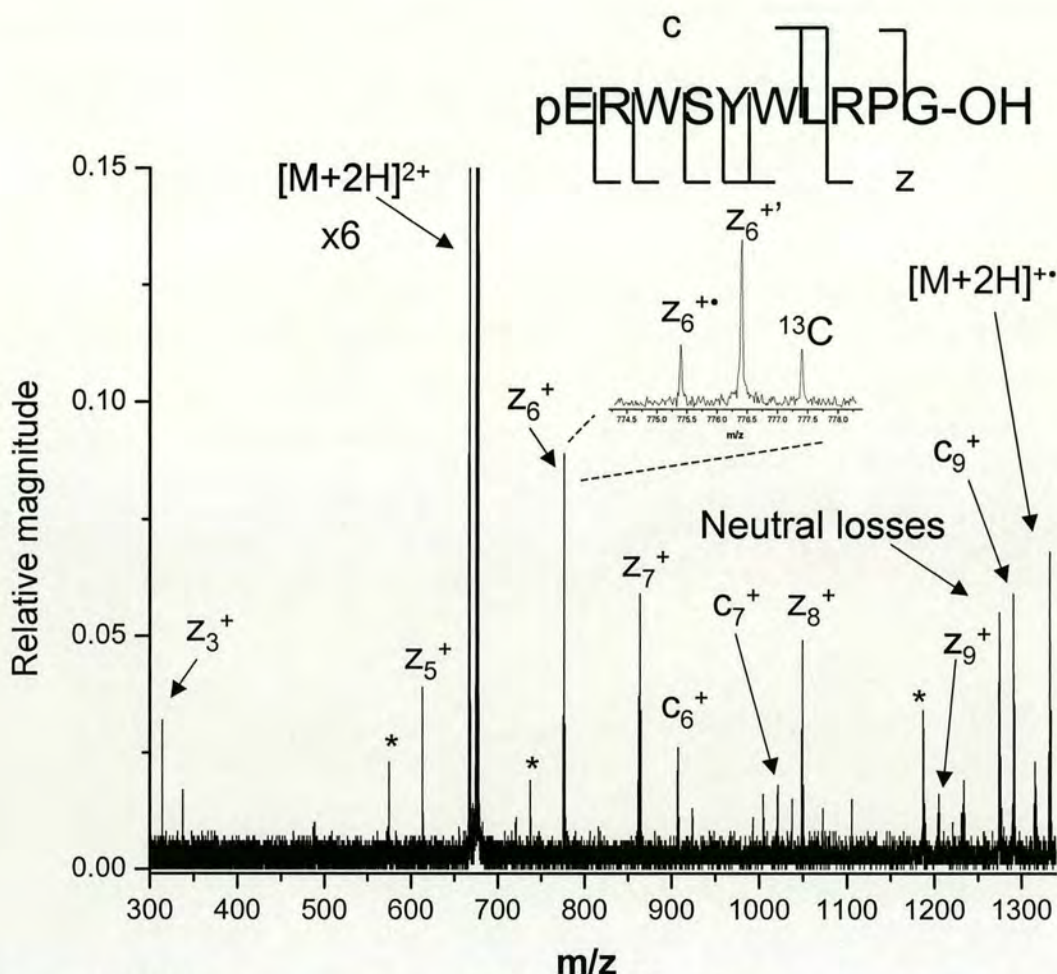
Figure 3.8 ECD mass spectrum of [LHRH_HH]$^{2+}$. (100 time-domain transients) Insert shows detail of z_5^+. (electron irradiation = 35 ms).

The low ECD efficiency (i.e. precursor to fragment ions conversion) for [LHRH_RR]$^{2+}$ (see Figure 3.9) could be due to the low efficiency of hydrogen atom transfer from the arginine side-chain to the backbone. Nevertheless, the ECD efficiency actually depends on a host of factors including ion cloud/electron beam overlap and the electron kinetic energy distribution and hence one needs to be careful when comparing different mass spectra. The experiment was repeated a number of times yielding no improvement in the precursor to fragment ion conversion.

Figure 3.9 ECD mass spectrum of [LHRH_RR]$^{2+}$. (300 time-domain transients) Insert shows detail of z_6^+. (electron irradiation = 35 ms).

3.3.6 ECD of [BK]$^{2+}$

In order to establish whether electron capture at the arginine side-chain always gives rise to inefficient backbone cleavage in ECD, the peptide bradykinin [BK]$^{2+}$ was irradiated with electrons to give the mass spectrum in Figure 3.10.

Figure 3.10 shows an efficient conversion of precursor to fragment ions. The N-terminal fragments in the spectrum are much more abundant than C-terminal fragments. These results are analogous to results shown by Tsybin *et al* [26] on a doubly protonated y fragment ion of a tryptic decapeptide from the signal recognition particle (SRP) of Saccharomyces cerevisae; ECD of the doubly protonated peptide fragment with sequence REYPLLIR only gave rise to N-terminally charged fragment ions.
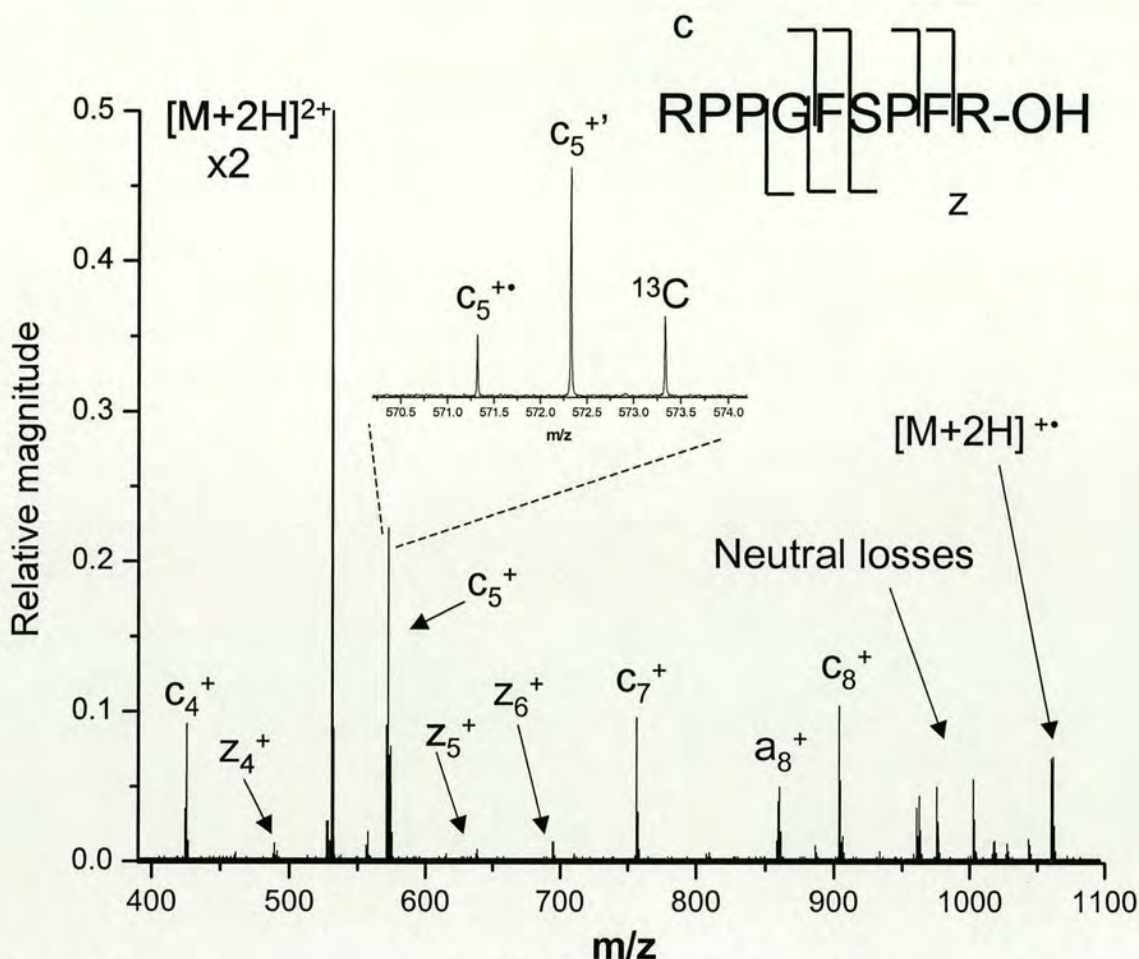
Figure 3.10 ECD mass spectrum of $[BK]^{2+}$. (100 time-domain transients) Insert shows detail of c_5^+ fragment. (electron irradiation $= 35$ ms).

In the light of the previous discussion, this could be due to two (among other) reasons: either the arginine on the C-terminus is less basic and therefore electron capture at that site results in more efficient hydrogen atom transfer to the backbone explaining more extensive backbone cleavage, or another site of protonation is present in the molecule. Since the N-terminus in this molecule is unprotected, it is a possible site of protonation. In the case of $[BK]^{2+}$ this would suggest a Zwitterionic structure (Figure 3.11), where the N-terminus is protonated and the C-terminus is de-protonated. This structure would explain why the N-terminal ECD fragments carry one proton, whereas the C-terminal fragments are neutral and therefore are not detected. The Zwitterionic structure of $[BK]^{2+}$ has also been proposed by Freitas and Marshall based on H/D exchange studies with D_2O [27]. H/D exchange studies of

bradykinin in an ion trap by Mao and Douglas have also indicated that $[BK]^{2+}$ can exist in the Zwitterionic form [28]. *Ab initio* calculations have shown that the Zwitterionic structures of arginine and arginine aggregates in the gas-phase can be stable [29-31], although the conclusions for theoretical calculations on bradykinin are more ambiguous [32]. Conversely, a direct probe of neutral arginine in the gas-phase by infrared cavity ringdown laser absorption spectroscopy (IR-CRLAS) [33] indicated that arginine exists predominantly in the neutral configuration in the gas-phase [34]. Since addition of water is known to stabilise Zwitterionic structures, the H/D exchange studies may in fact induce a Zwitterionic structure.
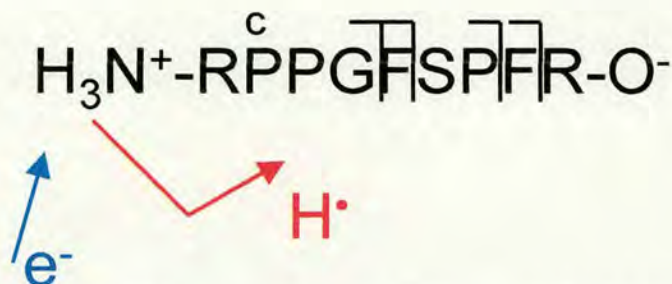
Figure 3.11 Schematic of electron capture at Zwitterionic bradykinin $[BK]^{2+}$, where the side-chains of arginine (R) are protonated.

3.3.7 ECD of $[Ac\_BK]^{2+}$

ECD of N-terminal protected bradykinin $[Ac\_BK]^{2+}$ (see Figure 3.12) gives rise to a similar fragmentation pattern than for $[BK]^{2+}$. In this case no free amine group is available at the N-terminus as a possible site of protonation. The two most likely sites of protonation are on the side-chains of the arginines. Nevertheless, the C-terminus may be deprotonated, while an extra proton is carried on another basic site (e.g. backbone carbonyl). The N-terminal arginine would be expected to have a higher gas-phase basicity than the C-terminal arginine residue in $[BK]^{2+}$ due to the free amine group on the N-terminus.

Even if $[BK]^{2+}$ cannot adopt the Zwitterionic structure, the presence of the free amine group on the N-terminus confers a higher basicity to the N-terminal arginine residue over the C-terminal arginine residue. Replacement of the free amine group in $[BK]^{2+}$ with an acetylated N-terminus in $[Ac\_BK]^{2+}$ results in a decrease of

the gas-phase basicity of the N-terminal arginine group, which manifests itself in Figure 3.12 in the form of a higher proportion of C-terminal fragments relative to Figure 3.10. It therefore appears that the relative abundance of N vs C-terminal ECD fragments can give information about the location and relative gas-phase basicities of the protonated amino acids in doubly protonated peptides.

ECD of singly protonated acetylated bradykinin $[Ac\_BK]^+$ (sequence: Ac-RPPGFSPFR) was attempted, but did not give rise to any (charged) fragments. A Zwitterionic structure of $[Ac\_BK]^+$ (i.e. deprotonated C-terminus and two protonated arginine side-chains) could yield charged c fragments if electron capture on the C-terminal arginine side-chain gives rise to backbone cleavage and the resulting c_n^+ and z_m^- fragments are able to break apart. Hence, this result shows that either electron capture on the arginine side-chain does not result in efficient backbone cleavage, and/or that the resulting fragments oppositely charged fragments are held together by Coulombic interactions.
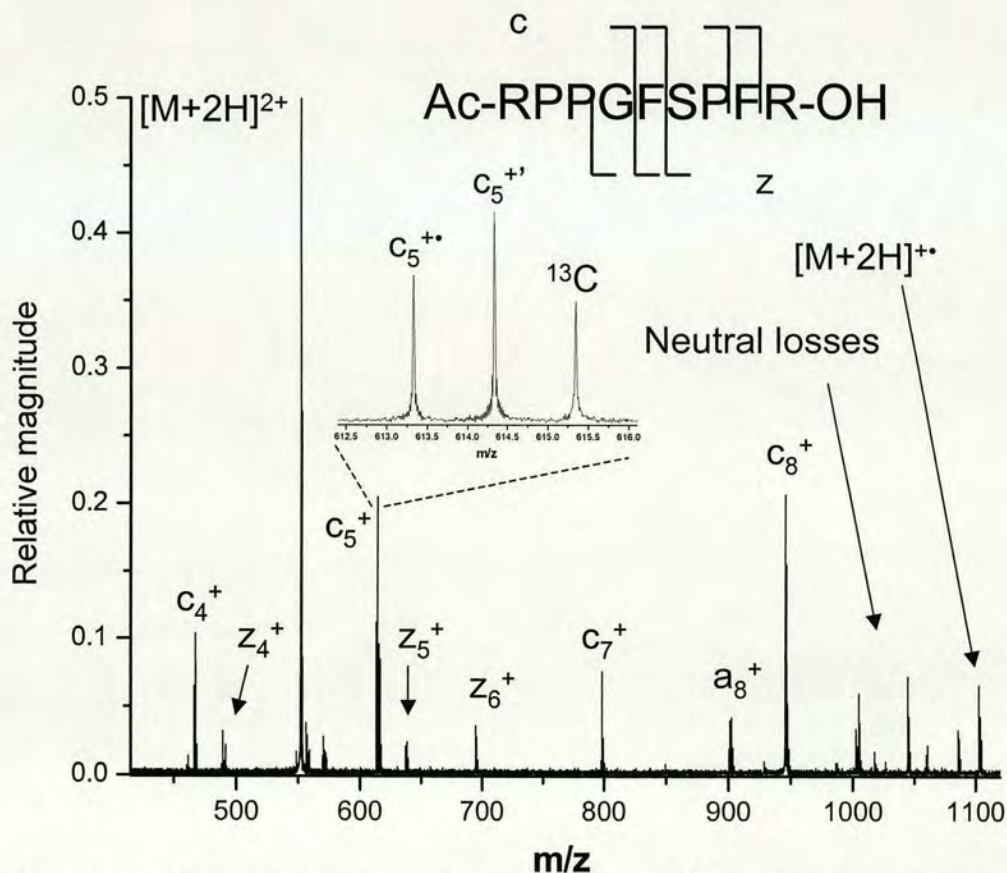
Figure 3.12 ECD mass spectrum of $[Ac\_BK]^{2+}$. (100 time-domain transients) Insert shows detail of c_5^+ fragment. (electron irradiation = 35 ms)

3.3.8 Conclusions

The observation that in the ECD mass spectra of [LHRH_HR]$^{2+}$, [LHRH_RH]$^{2+}$ and [BOMB]$^{2+}$ all charged ECD fragments include the arginine residue implies a correlation between the higher gas-phase basicity of that site and the fragmentation pattern. Various groups have shown that ECD of doubly protonated peptides preferentially gives rise to charged fragments that include an arginine residue [35] [36] [22] [25]. Håkansson *et al* [22] have linked the higher gas-phase basicity of arginine with the observation that charged fragments include this residue.

One way to explain these results, in analogy to the *mobile proton* model [9] and in terms of the *hot hydrogen re-arrangement* mechanism [18], is that hydrogen atom transfer from the arginine side-chain to the backbone is less efficient than from a less basic site. This may also be due to the fact that the recombination energy of electron capture at a less basic site is higher than for a more basic site (see Figure 1.13, Chapter 1). In support of this assumption, ions at higher internal energies have been shown to give rise to a higher proportion of charged fragments that do not include an arginine residue. Haselmann *et al.* [36] have shown that pulsing a gas during electron irradiation resulted in a higher loss of labile groups and a higher proportion of charged fragment ions that do not include an arginine residue. Similarly, Tsybin *et al.* [26] have shown that infrared irradiation of ions prior to electron irradiation caused the appearance of fragment ions without an arginine residue, which were not observed for electron irradiation only. This shows that 'proton mobility' and 'hydrogen atom mobility' in a peptide are related and that the relative ECD fragmentation pattern is related to the internal energy of the molecule.

ECD of [LHRH_RR]$^{2+}$ appears to confirm the assumption that hydrogen atom transfer from the arginine side-chain is less efficient. In this light, the abundant ECD fragmentation of [Ac_BK]$^{2+}$ is surprising, but could be explained by deprotonation of the C-terminus and a protonation site elsewhere on backbone. As mentioned before, arginine may exist as a Zwitterion in the gas-phase [30].

An alternative explanation for the observation that the most basic amino acid residue is preferentially included in the charged fragment is that the ECD mechanism takes place on a much slower time-scale and that backbone cleavage and proton transfer take place on similar time-scales. The time-scale of the ECD process cannot be measured directly. In fact, the time-scale of an ECD experiment in an FT-ICR instrument is very long (i.e. 100's ms) and hence it is not possible to establish whether the reaction is *non-ergodic* (i.e. < ps) or at a longer time-scale of ns, μs or even ms! *Ab initio* calculations by Uggerud and co-workers have indicated that ECD is a *non-ergodic* mechanism for a minority of peptide ions (~ 10%), while for the majority of ions backbone cleavage occurs on a slower time-scale (~ 10^{-5} s) [37]. *Ab initio* calculations by Tureček have also shown that the mechanism may not have to be *non-ergodic* in order to account for the observed N-C_α backbone cleavage [38].

Further, Tureček and co-workers [39] have proposed that the electron and proton move separately, as first suggested by Siegbahn *et al* [40], effectively resulting in hydrogen atom transfer from the protonated site to the backbone carbonyl (see Chapter 1, section 1.6.3.2). A separate movement of the electron and proton could explain why proton transfer occurs mainly from the less basic site in the molecule (This idea will be re-visited in Chapter 6).

McLafferty and co-workers believe that the ECD mechanism involves capture of an electron in a Rydberg state, in analogy to dissociative recombination [41], followed by fast intersystem crossing to result in *non-ergodic* fragmentation [42]. Neverthess, the molecule still somehow 'knows' where the most basic site in the molecule is. Charge delocalisation at the guanidine group is particularly favoured (thus explaining the high gas-phase basicity of that site), hence we can assume that charged fragments including the guanidine group should be energetically favoured to charged fragments that instead include a site of less favourable charge delocalisation (e.g. imidazole ring of histidine side-chain). This would imply that ECD preferentially results in fragments that are more energetically favourable. Since ions at higher internal energies give rise to fragments that are less energetically favoured, the ECD process appears to behave analogous to a thermodynamically driven reaction. This also implies that the cleavage pattern is not 'reducible' (i.e. cleavage

efficiency related to immediate chemical vicinity, not rest of molecule) [13], since the cleavage pattern clearly does depend on where the protonated sites are present.

In spite of the ambiguity surrounding the mechanism in ECD, the fact that the charged ECD fragments in this study preferentially include the most basic amino acid gives ECD an advantage over vibrational excitation techniques: if the most basic amino acid is located at one of the termini one would expect predominantly N or C-terminal fragment ions for doubly protonated peptides. This renders *de novo* sequencing much easier. Since most tryptic digests have a C-terminal lysine or arginine this strategy is applicable to many compounds.

3.4 Reverse cleavage

The occurrence of $c^•$ (odd-electron) and z' (even-electron) fragments was first reported in the literature by Zubarev *et al.* [18], and the fragments were rationalised by hydrogen atom transfer to an $C_α$. An alternative explanation consists in the attraction of the highly polarisable hydrogen atom by the electric field gradient for smaller z fragments [43], resulting in z' and $c^•$ ions.

The insert in Figure 3.4 shows the $z_5^{+•}$ and $z_5^{+'}$ fragments, which are separated by a mass difference of 1.00795 Da, corresponding to a hydrogen atom difference. Conversely, the mass difference between the $z_5^{+'}$ peak and the ^{13}C labelled peak is 1.0033 Da, which is consistent with a heavier isotope peak. By analogy, Figure 3.5 shows the $c_5^{+'}$ fragment and a fragment, labelled as $c_5^{+•}$, which is 1.017 Da lower in mass, corresponding to a hydrogen atom difference. The fragments can be explained in terms of the *hot hydrogen re-arrangement* mechanism due to a 'reverse' alpha-cleavage taking place (see Figure 3.13, scheme B). When a 'reverse' alpha-cleavage takes place the hydrogen atom would be associated with the C-terminal fragment z', thus causing this fragment to be 1 Da higher in mass higher than generally expected in ECD, whereas the N-terminal fragment $c^•$ would be 1 Da lower in mass than generally predicted. The 'reverse' mechanism has been put forward by Kim Haselmann [44].

The observation of radical, as well as even-electron c and z fragments supports the view that a *hydrogen atom re-arrangement* takes place in ECD.

However, Tureček and co-workers believe that the appearance of c• and z' fragments can be explained by the fact that <u>after</u> cleavage of the N-C_α bond the c' and z• fragments are hydrogen-bonded for long enough to account for an exothermic hydrogen atom abstraction from C_α of the c' fragment to C_α on the z• fragment [39], resulting in c• and z' fragments.

The observation by O'Connor and co-workers that one electron capture can result in two backbone cleavages [45] opens the possibility that odd-electron fragment ions can further fragment. Hence, one would expect the most stable odd-electron fragments to be observed preferentially.

It is possible that multiple mechanisms are in competition with one another (see Chapter 1, section 1.6.2). In this study, the assumption will be made that electron capture in these studied peptides <u>predominantly</u> results in one backbone alpha-cleavage which can proceed either in the 'normal' direction (Figure 3.13 A) or in the 'reverse' direction (Figure 3.13 B). The results in this section will be presented in the light of a competition between these two processes.
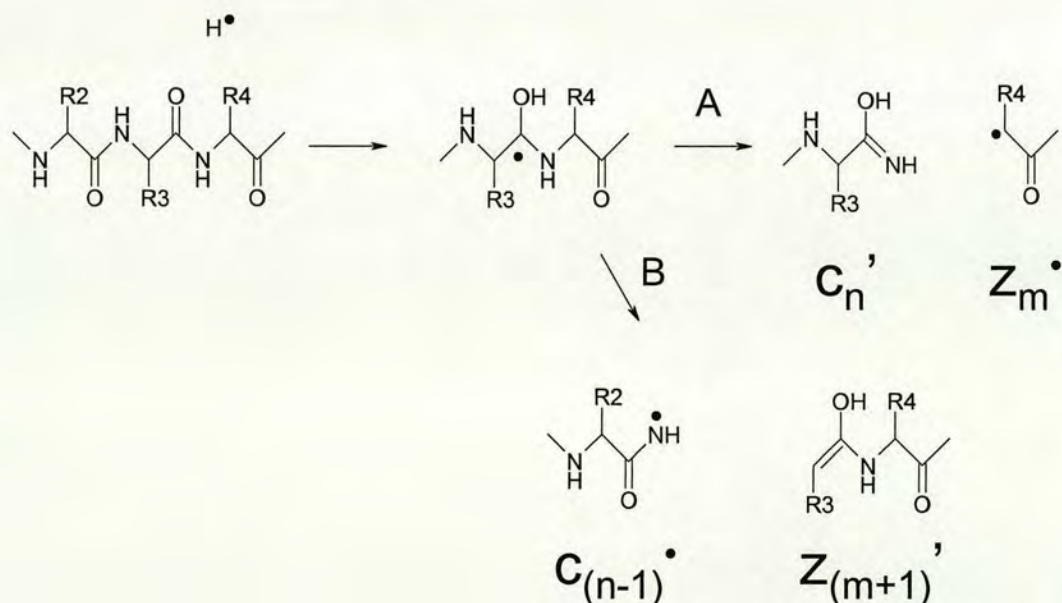
Figure 3.13 (A) Schematic of 'normal' alpha-cleavage on peptide backbone as proposed by McLafferty and co-workers [18] resulting in c_n' and z_m• fragments (see Chapter 1, section 1.6.3.1). (B) 'Reverse' alpha-cleavage on peptide backbone, resulting in $c_{(n-1)}$• and $z_{(m+1)}'$ fragments, where n is the number of residues from N-terminus, m the number of residues from the C-terminus and (n+m) is the total number of residues in the peptide.

From Figure 3.13 it can be seen that fragments $z_m{}^\bullet$ and $z_{(m+1)}'$ are due to the hydrogen atom being transferred to the same backbone carbonyl. By analogy, fragments c_n' and $c_{(n-1)}{}^\bullet$ also result from a hydrogen atom being transferred to the same backbone carbonyl. If one assumes that the 'normal' and 'reverse' alpha-cleavage are in competition with one another, the relative appearance of either alpha-cleavage mechanism determines the relative abundance of $c_n'/\,z_m{}^\bullet$ and $c_{(n-1)}{}^\bullet/\,z_{(m+1)}'$. This section is concerned with investigating the trends giving rise to either mechanism. It is possible that $c_{(n-1)}{}^\bullet$ would further dissociate to $a_{(n-1)}{}^\bullet$ fragments [38], however, no such fragments are observed in the present study, with the exception of $[BK]^{2+}$ and $[Ac\_BK]^{2+}$.

3.4.1 ECD of $[LHRH\_HR]^{2+}$

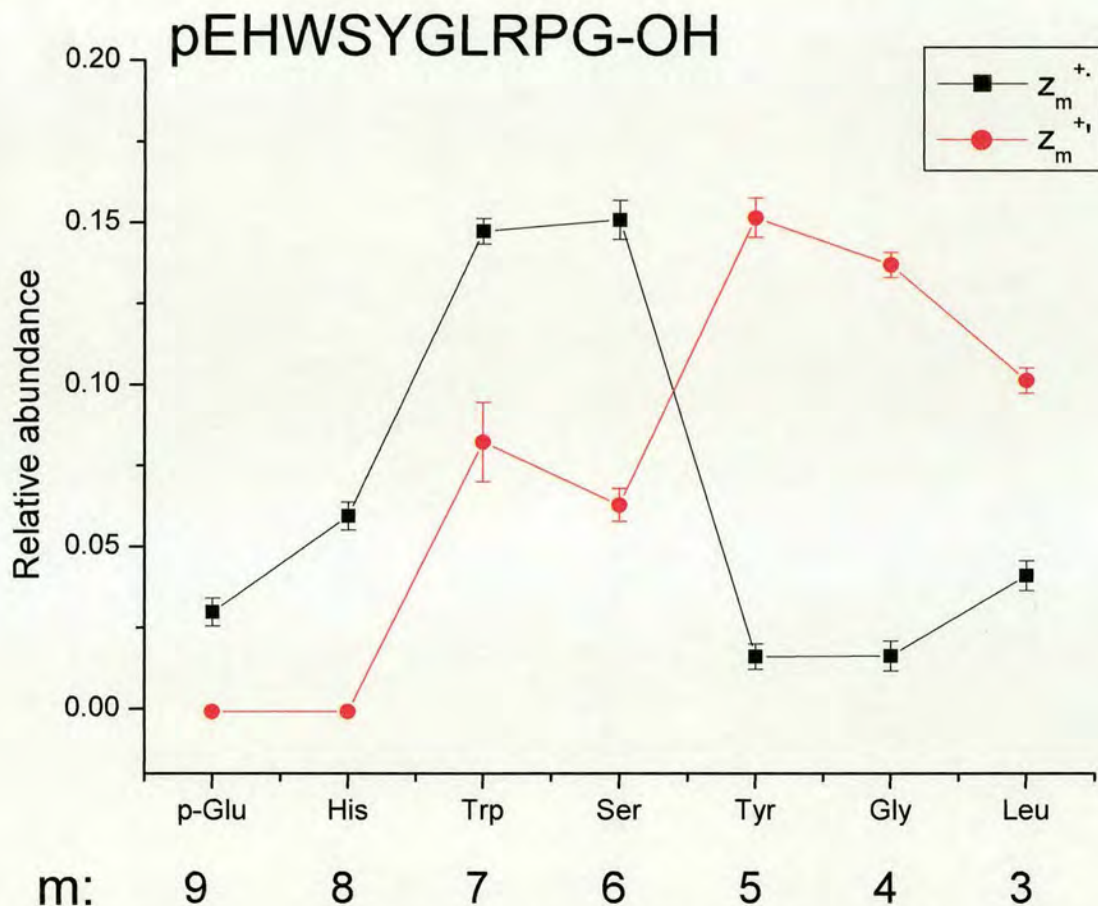Figure 3.14 shows the relative abundances of the $z_m{}^{+\bullet}$ and $z_m{}^{+'}$ fragment ions for ECD of $[LHRH\_HR]^{2+}$, as observed in the mass spectra. This Figure was made by the following procedure: 3 mass spectra (consisting of 100 co-averaged time-domain transients) were taken at a particular hexapole accumulation time. This was repeated for two other hexapole accumulation times, yielding 9 mass spectra in total. The relative intensity of each peak in the mass spectrum was calculated by dividing its peak intensity by the sum of the intensities of all the z fragments ($z_3{}^{+\bullet}$ to $z_9{}^{+\bullet}$, and $z_3{}^{+'}$ to $z_9{}^{+'}$)[6], taking into account the heavier isotope peaks of $z_m{}^{+\bullet}$. For the 9 spectra the mean and standard deviation of each relative fragment abundance was determined, which is summarised in Figure 3.14.

A general observation in Figure 3.14 is that even-electron z' fragments are preferentially observed for smaller fragments, while larger z fragments only exist as radical $z^\bullet$ fragments ($z_8{}^{+'}$ and $z_9{}^{+'}$ not observed). This has also been reported by Zubarev and co-workers [43]. Nevertheless, the trend in Figure 3.14 shows that the relative abundance of $z_m{}^{+'}$ ions is not simply inversely proportional to size, but rather shows a maximum at $z_5{}^{+'}$.

---

[6] $(z+1)_{10}{}^+$ not possible due to cyclic structure of pyro-glutamic acid and no smaller z fragment than $z_3{}^{+\bullet}$ observed.

An alternative way to represent these fragment abundances is in terms of a competition between 'normal' and 'reverse' alpha-cleavage. Figure 3.15 shows the relative z fragment abundances in terms of this competition and in terms of which backbone carbonyl the hydrogen would have to be transferred to for that fragment to be observed. Note that in Figure 3.15 the fragment abundance of $z_3^{+'}$ is not represented, since the corresponding 'normal' alpha-cleavage fragment, $z_2^{+\bullet}$, is not observed. Therefore, the abundance of $z_3^{+'}$ is not included in calculating the relative fragment abundance in Figure 3.15.

Figure 3.14 Relative abundances of $z_m^{\bullet+}$ and $z_m^{'+}$ fragment ions for ECD of [LHRH_HR]$^{2+}$ as observed in the ECD mass spectra. Error bars represent standard deviation obtained for 3 spectra (of 100 time domain transients each) taken at hexapole accumulation times of 0.7, 0.8 and 0.9 s. (30 ms electron irradiation).

The different hexapole accumulation times should account for different heating of the ions in the hexapole. A high number of ions in the hexapole causes the ions to be further away from the centre of the hexapole, thereby resulting in higher

kinetic energy collisions and therefore more significant heating of the ions [46]. Longer accumulation times (> 1.5 s) can give rise to peptide fragmentation (data not shown). Figure 3.15 shows that the relative abundance of the fragments does not vary significantly over the different hexapole accumulation times, this indicates that the temperature of the ions is probably similar over this range of hexapole accumulation times.
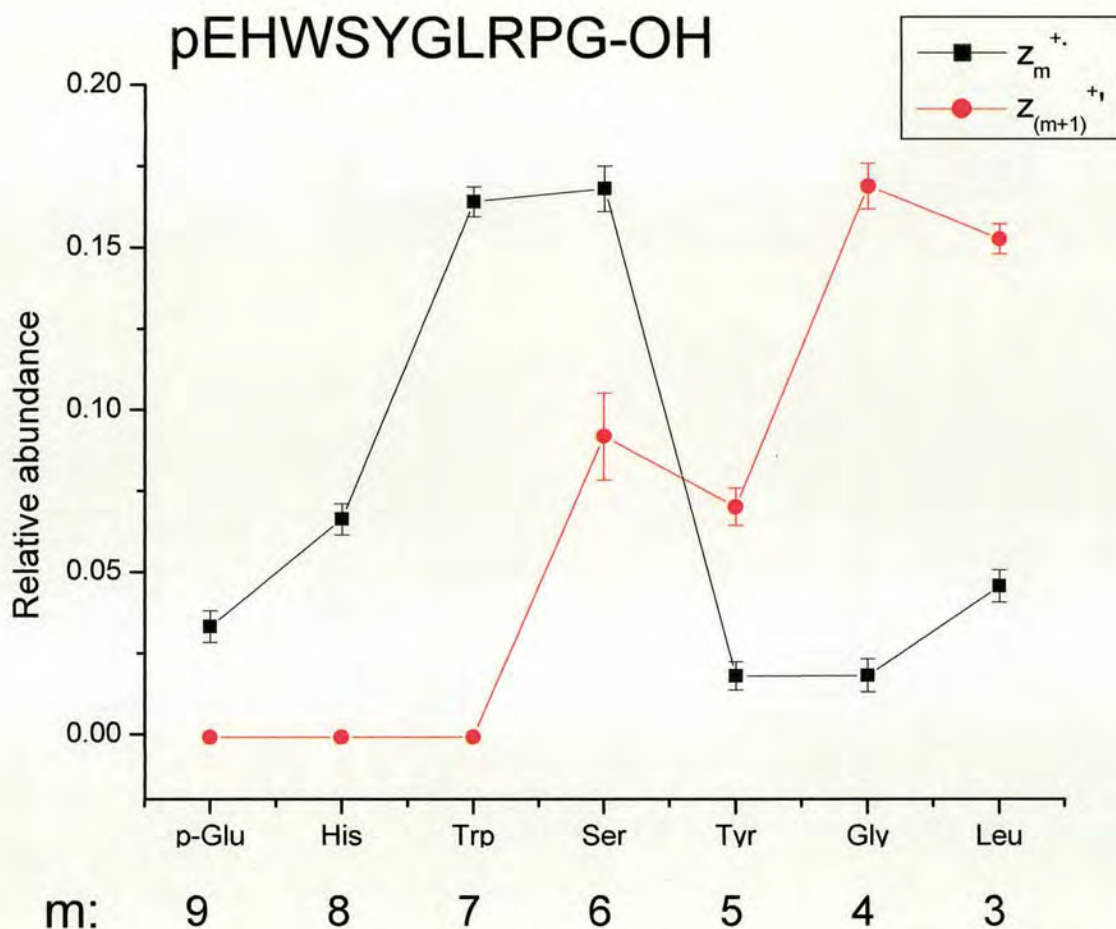
Figure 3.15 Relative intensity of $z_m^{\bullet+}$ and $z_{(m+1)}^{+,}$ fragment ions for ECD of $[LHRH\_HR]^{2+}$ showing amino acid backbone carbonyl where hydrogen atom would have to be transferred to for that fragment to appear. Error bars represent standard deviation obtained for 3 spectra (of 100 time domain transients each) taken at hexapole accumulation times of 0.7, 0.8 and 0.9 s. (30 ms electron irradiation).

Electron irradiation was also carried out on $[LHRH\_HR]^{2+}$ using the 3 Tesla APEX II instrument equipped with the rhenium filament. The relative z fragment abundances from these experiments are shown in Figure 3.16 (for raw ECD

spectrum, see Appendix 3, Figure A3.1). The procedure to calculate the relative abundances of the fragments in this Figure is the same as for Figure 3.15. The relative fragmentation pattern in Figure 3.16 is similar, but not identical, to the one in Figure 3.15. The intensity of fragment $z_6^{+\bullet}$, in particular, is higher in Figure 3.15 than in Figure 3.16. In spite of the fact that the hexapole accumulation times (0.7, 0.8 and 0.9 s) were the same in both experiments and the analyte concentration was the same (20 μM pEHWSYGLRPG-OH, MeOH:H$_2$O:CH$_3$COOH, 0.495:0.495:0.01) other instrumental features are different in both instruments:

- The 3 T APEX II Analytica source probably has a higher transmission of ions into the hexapole due to a higher internal diameter (ID) of the glass capillary (ID 0.3 mm vs 0.2 mm) (see Chapter 2, section 2.1.1 and 2.1.2). A higher number of ions in the hexapole results in a more significant heating of the ions.

- The ion transmission in the 3 T instrument is probably better due to a lower magnetic mirror effect compared to the higher field 9.4 T instrument. This is consistent with other findings.

- The electron sources in both instruments are different (see Chapter 2). The electron capture efficiency in the 3 T instrument is generally poor due to the small diameter of the electron beam from the rhenium filament.[7] Since only a small proportion of ions capture an electron, this also results in higher statistical variation, which can be seen in the higher standard deviation in Figure 3.15 compared to Figure 3.14.

- The electron irradiation time in both instruments varies significantly (30 ms for high-surface cathode as opposed to 500 ms for electron filament). It is conceivable that the gas-phase structure could vary over such long experiment times, hence resulting in a different fragmentation pattern.

---

[7] Intensity of [M+2H]$^{2+}$ for [LHRH_HR]$^{2+}$ remaining after 500 ms electron irradiation using rhenium filament compared to sum of intensities of [M+2H]$^{+\bullet}$ and [M+H]$^+$ 11 : 1 (see Appendix 3, Figure A3.1). Compare this to high-surface cathode results where intensity of [M+2H]$^{2+}$ for [LHRH_HR]$^{2+}$ remaining after 30 ms electron irradiation relative to sum of intensities of [M+2H]$^{+\bullet}$ and [M+H]$^+$ 3.5 : 1 (see Figure 3.4).
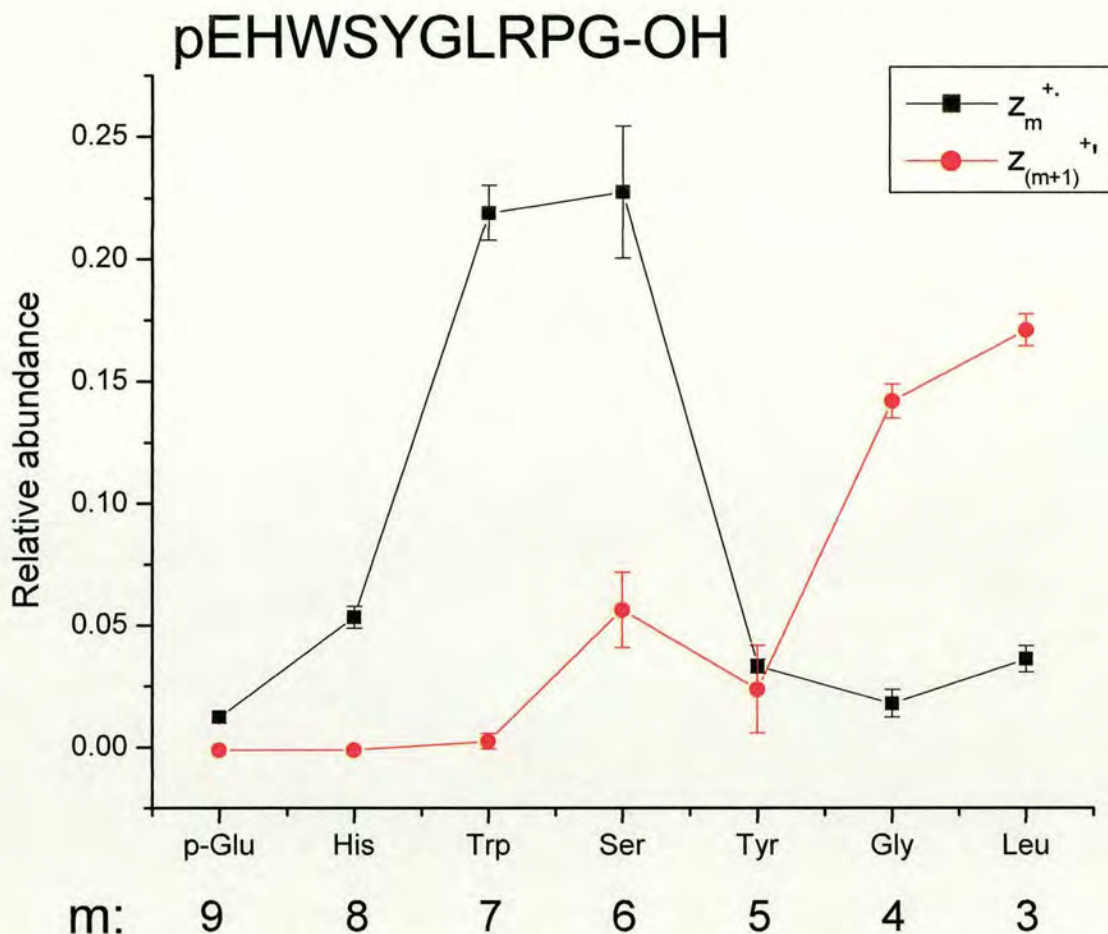
pEHWSYGLRPG-OH

Figure 3.16 Relative intensity of $z_m^{\bullet+}$ and $z_{(m+1)}^{+'}$ fragment ions for ECD of [LHRH_HR]$^{2+}$ showing amino acid backbone carbonyl where hydrogen atom would have to be transferred to for that fragment to be observed. Error bars represent standard deviation obtained for 3 spectra (of 100 time domain transients each) taken at hexapole accumulation times of 0.7, 0.8 and 0.9 s. (3 T APEX II, rhenium filament, 500 ms electron irradiation).

In Figures 3.15 and 3.16, the relative occurrence of the 'reverse' alpha-cleavage decreases as a function of the fragment size. The 'normal' alpha-cleavage results in odd-electron $z^{+\bullet}$ fragments, whereas the 'reverse' alpha-cleavage results in even-electron charged $z^{+'}$ fragments. It appears that the odd-electron charged species are disfavoured relative to even-electron charged fragments in the vicinity of the protonated arginine side-chain. In Figure 3.15 the $z_4^{+\bullet}$ fragment is particularly disfavoured compared to $z_5^{+'}$. Figure 3.17 shows the proposed structures for these fragments. The amino acid side-chain adjacent to the radical carbon may play a role in stabilising the odd-electron species. Table 3.4 shows the correlation between the

amino acid side-chain adjacent to the radical carbon and the preferential appearance of the $z_m^{+\bullet}$ fragment, as opposed to $z_{(m+1)}^{+}$', based on Figure 3.15. The charge on the $z_m^{+\bullet}$ fragment would be expected to be located on the arginine amino acid side-chain. It appears from Table 3.4 that in particular aromatic amino acid side-chains adjacent to the radical C favour the appearance of that $z_m^{+\bullet}$ fragment.

A

B
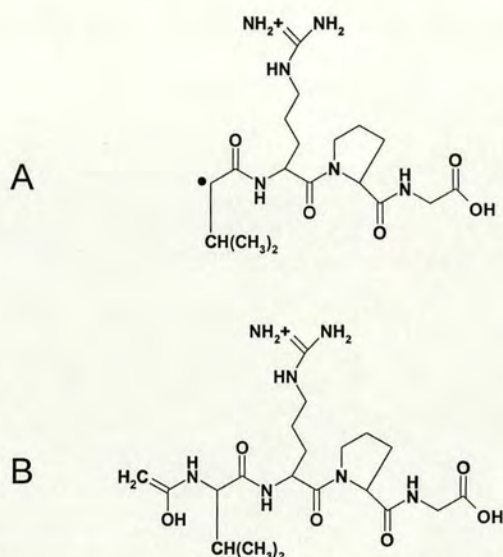
Figure 3.17 Proposed structures (stick structure) of $z_4^{+\bullet}$ and z_5^{+}' fragment ions for [LHRH_HR]$^{2+}$.

Side-chain	Chemical formula	Appearance of $z_m^{+\bullet}$
Arg	$CH_2CH_2CH_2NHC(NH_2)_2$	N
Leu	$CH_2CH(CH_3)_2$	N
Gly	H	N
Tyr	$CH_2C_6H_6OH$	Y
Ser	CH_2OH	Y
Trp	$CH_2C_7H_6N$	Y

Table 3.4 Correlation between amino acid side-chain and preferential appearance of $z_m^{+\bullet}$ fragment for Figure 3.17. If $z_m^{+\bullet}$ fragment is preferred (Y), if corresponding 'reverse' alpha-cleavage $z_{(m+1)}^{+}$' fragment is preferred (N).

3.4.2 ECD of [LHRH_HR_NH$_2$]$^{2+}$

The relative z fragmentation pattern of the doubly charged peptide [LHRH_HR_NH$_2$]$^{2+}$ (mammalian type LHRH) is shown in Figure 3.18. The raw ECD mass spectrum is shown in Appendix 3, Figure A3.2. The ECD mass spectrum of [LHRH_HR_NH$_2$]$^{2+}$ presented by Håkansson *et al.* [22] shows a series of z fragments (z_3^+ to z_7^+) and c_9^+. The relative fragment abundance is similar, although they did not observe z_8^+ and z_9^+. While in Figure A3.2 (see insert, Appendix 2) the peak intensity of $z_6^{+\bullet}$ is much higher than for z_6^{+}, the results presented by Håkansson *et al.* showed equal peak intensities for both [22]. The different experimental parameters employed by Håkansson *et al.* (8 s accumulation in multipole, 30 s irradiation with electron filament) are very different to the ones employed in this experiment (< 1 s accumulation, 30 ms electron irradiation with dispenser cathode). This shows that different experimental conditions can yield different fragment abundances, which could be rationalised by a change in the gas-phase structure.

[LHRH_HR_NH$_2$]$^{2+}$ and pEHWSYGLRPG-OH (i.e. [LHRH_HR]$^{2+}$) only differ by their C-termini and therefore one would expect the relative fragmentation patterns to be similar. As far as the relative abundances of $z_m^{+\bullet}$ and $z_{(m+1)}^{+}$ fragments are concerned, this seems to be the case. For pEHWSYGLRPG-NH$_2$ the amidated C-terminus is a competing site of protonation which could account for the slight differences in fragmentation compared to [LHRH_HR]$^{2+}$, and this may also affect the gas-phase structure (see Chapter 4, section 4.4).

pEHWSYGLRPG-NH$_2$
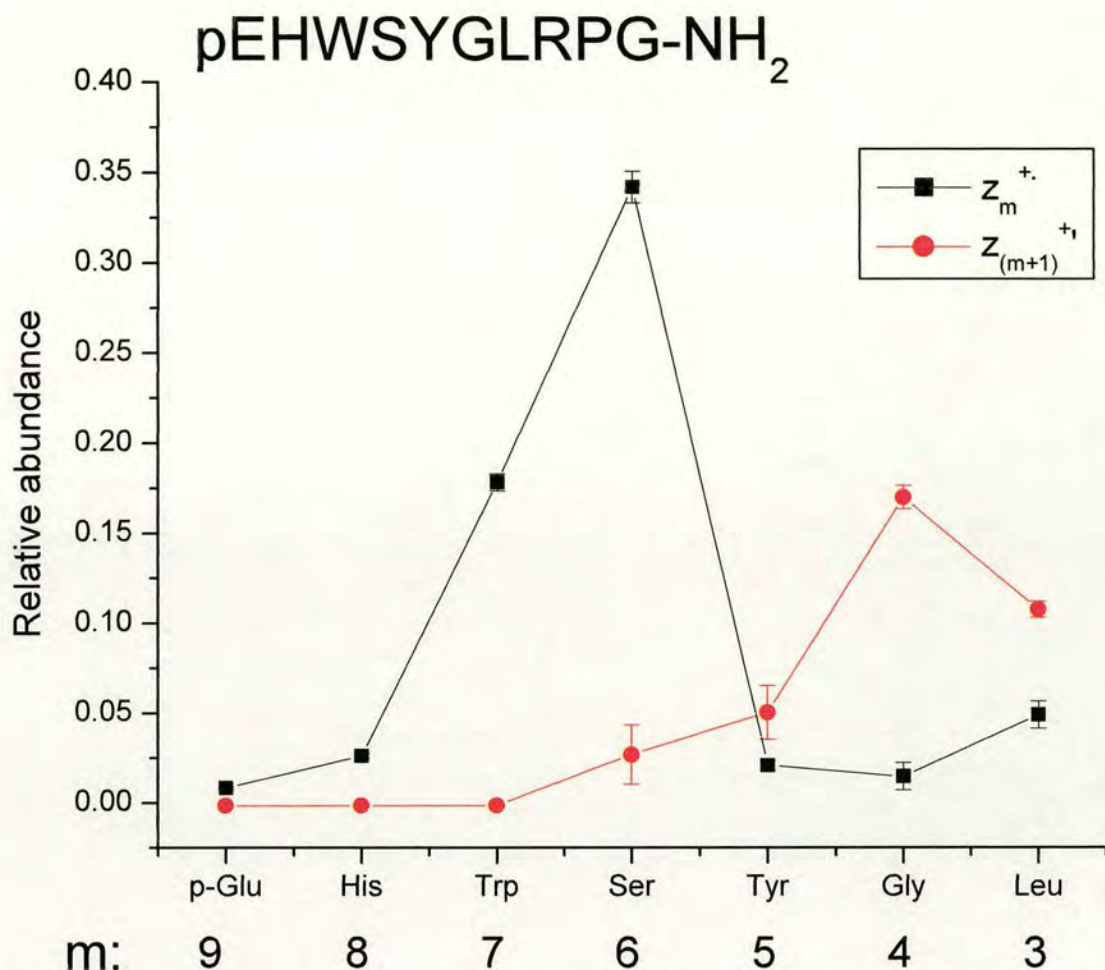
Figure 3.18 Relative abundance of $z_m^{+\bullet}$ and $z_{(m+1)}^{+'}$ fragment ions for ECD of [LHRH_HR_NH$_2$]$^{2+}$ showing amino acid backbone carbonyl where hydrogen atom would have to be transferred to for that fragment to be observed. Error bars represent standard deviation obtained for 3 spectra (of 100 time domain transients each) taken at hexapole accumulation times of 0.7, 0.8 and 0.9 s. (30 ms electron irradiation).
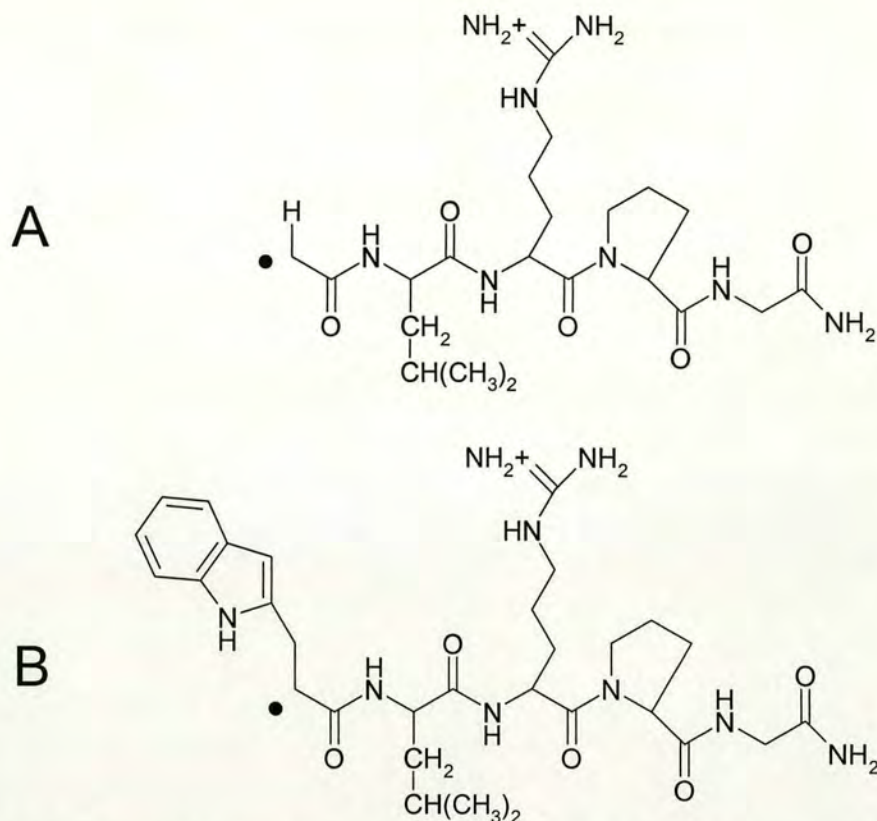
3.4.3 ECD of [LHRH_(l)W]$^{2+}$ and [LHRH_(d)W]$^{2+}$

The relative z fragmentation pattern of [LHRH_(l)W]$^{2+}$ is shown in Figure 3.19. The raw ECD mass spectrum for this peptide is shown in Appendix 3, Figure A3.3. The relative occurrence of fragments $z_5^{+\bullet}$ and $z_6^{+'}$ is reversed in this case compared to Figures 3.17 and 3.15. Since in this peptide the glycine residue in position 6 is replaced by a tryptophan residue, the structure of the fragment ion is different. The structures of $z_5^{+\bullet}$ for pEHWSYWLRPG-NH$_2$ and pEHWSYGLRPG-

117

NH_2 are shown in Figure 3.20. This appears to confirm the assumption that an aromatic side-chain group adjacent to the radical carbon stabilises $z_m^{+\bullet}$ fragments.
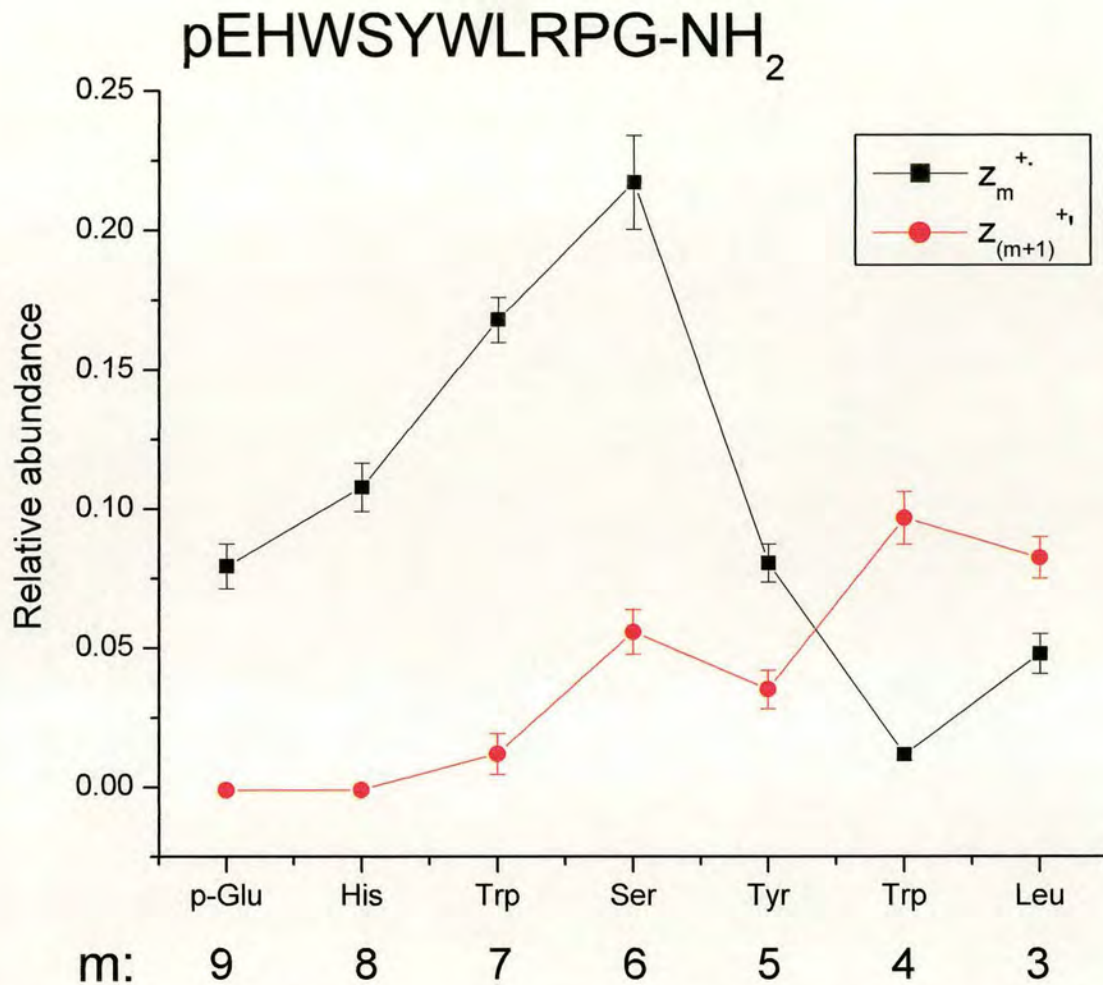
Figure 3.19 Relative intensity of $z_m^{+\bullet}$ and $z_{(m+1)}^{+'}$ fragment ions for ECD of $[LHRH\_(l)W]^{2+}$ showing amino acid backbone carbonyl where hydrogen atom would have to be transferred to for that fragment to be observed. Error bars represent standard deviation obtained for 3 spectra (of 100 time domain transients each) taken at hexapole accumulation times of 0.7, 0.8 and 0.9 s. (35 ms electron irradiation).

Figure 3.20 Proposed structures (stick structures) for $z_5^{+\bullet}$ fragment for (A) $[LHRH\_HR\_NH_2]^{2+}$ and (B) $[LHRH\_(l)W]^{2+}$. The site of protonation is assumed to be on the arginine side-chain.

The relative z fragmentation pattern of the doubly charged peptide $[LHRH\_(d)W]^{2+}$ is shown in Figure 3.21. The raw ECD mass spectrum for this species is shown in Appendix 3, Figure A3.4. This peptide varies compared to $[LHRH\_(l)W]^{2+}$ only by the fact that the tryptophan residue in position 6 is a (d) instead of an (l) stereoisomer. In the case of $[LHRH\_(l)W]^{2+}$ the $z_5^{+\bullet}$ fragment is more abundant than the corresponding reverse cleavage $z_6^{+'}$ fragment. Whereas the ECD mass spectra of the $[LHRH\_(l)W]^{2+}$ (Figure A3.3) and $[LHRH\_(d)W]^{2+}$ (Figure A3.4) look very similar, the inserts show the relative abundances of the $z_6^{+'}$ fragment, which is higher for $[LHRH\_(d)W]^{2+}$. This finding contradicts the previous finding that an aromatic side-chain adjacent to the radical carbon favours odd-electron $z_m^{+\bullet}$ ions. The substitution of an (l)W by a (d)W may result in a different gas-phase conformation and hence the difference in fragmentation may be related to the structure of the peptide ion. This will be re-visited in Chapter 4, section 4.5.
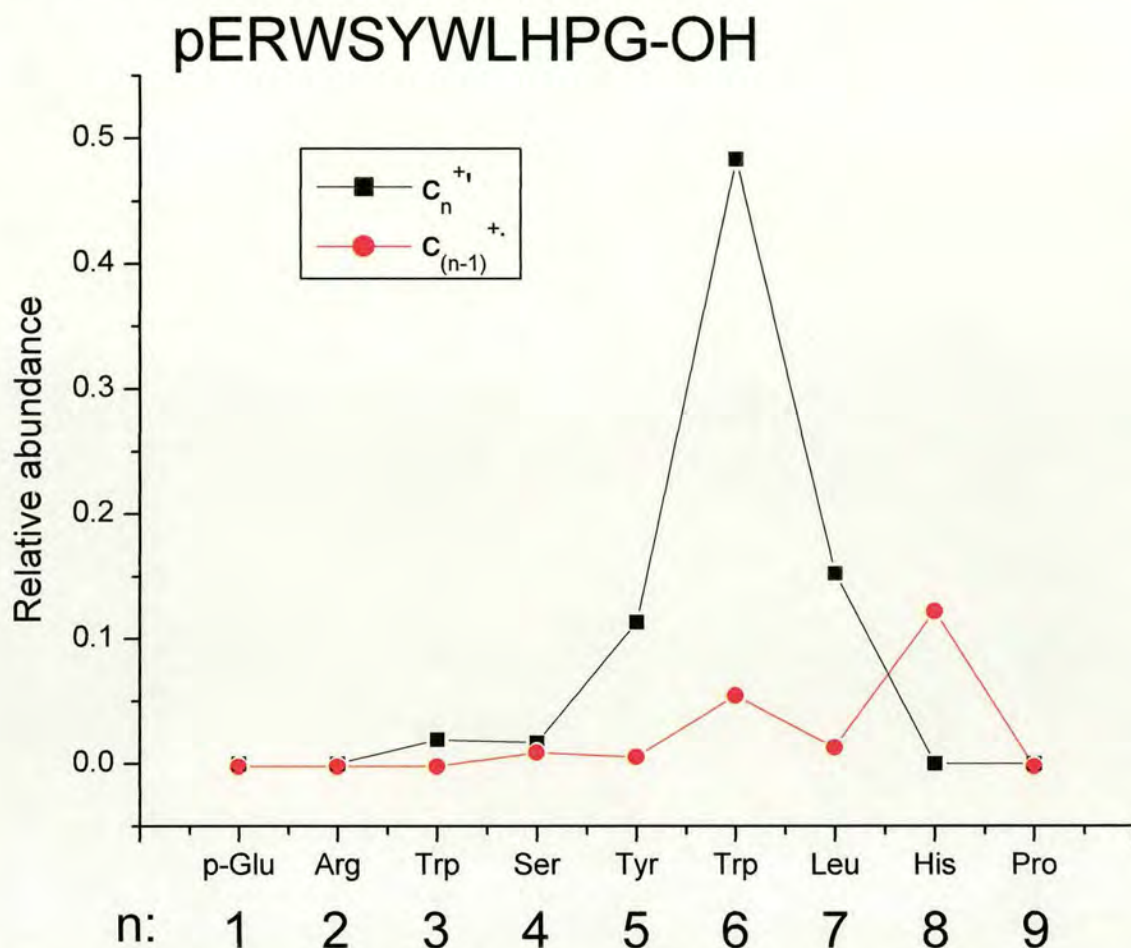
pEHWSY(d)WLRPG-NH$_2$
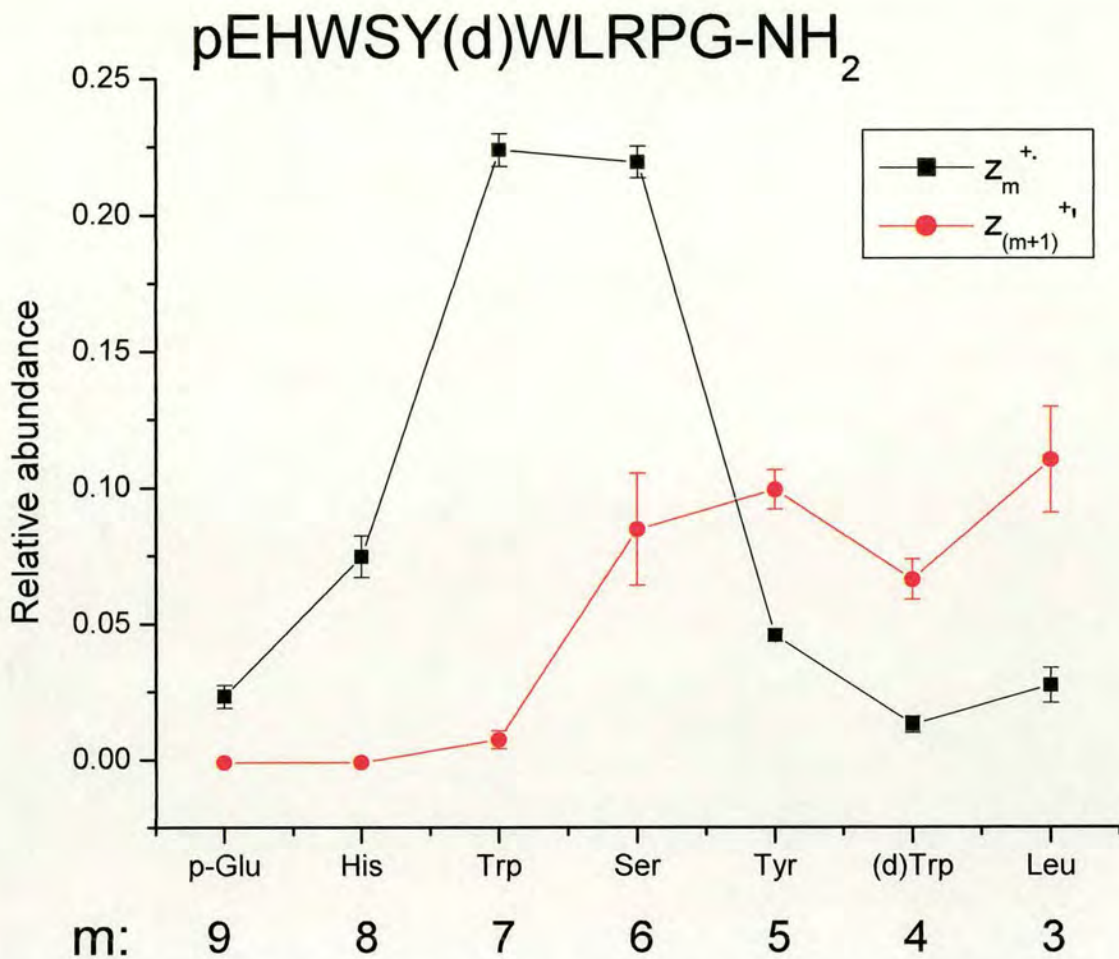
Figure 3.21 Relative abundance of $z_m^{\bullet+}$ and $z_{(m+1)}^{+'}$ fragment ions for ECD of doubly charged [LHRH_(d)W]$^{2+}$ peptide showing amino acid backbone carbonyl where hydrogen atom would have to be transferred to for that fragment to be observed. Error bars represent standard deviation obtained for 3 spectra (of 100 time domain transients each) taken at hexapole accumulation times of 0.7, 0.8 and 0.9 s. (35 ms electron irradiation).

3.4.4 ECD of [BOMB]$^{2+}$

The relative c fragmentation pattern for [BOMB]$^{2+}$ is shown in Figure 3.22. In this case the 'normal' alpha-cleavage results in even-electron $c_n^{+'}$ fragments and the 'reverse' cleavage results in corresponding odd-electron $c_{(n-1)}^{+\bullet}$ fragments. The 'reverse' cleavage is particularly favoured for hydrogen atom transfer to the alanine{9} backbone carbonyl, yielding a much higher abundance $c_8^{+\bullet}$ than the corresponding 'normal' alpha-cleavage $c_9^{+'}$ fragment. Figure 3.23 shows the proposed structures for both fragments. The vicinity of the tryptophan side-chain in

relation to the radical N atom could play a role in stabilising $c_8^{+\bullet}$, by forming a cyclic structure (as shown in Figure 3.23 B), which is resonantly stabilised.
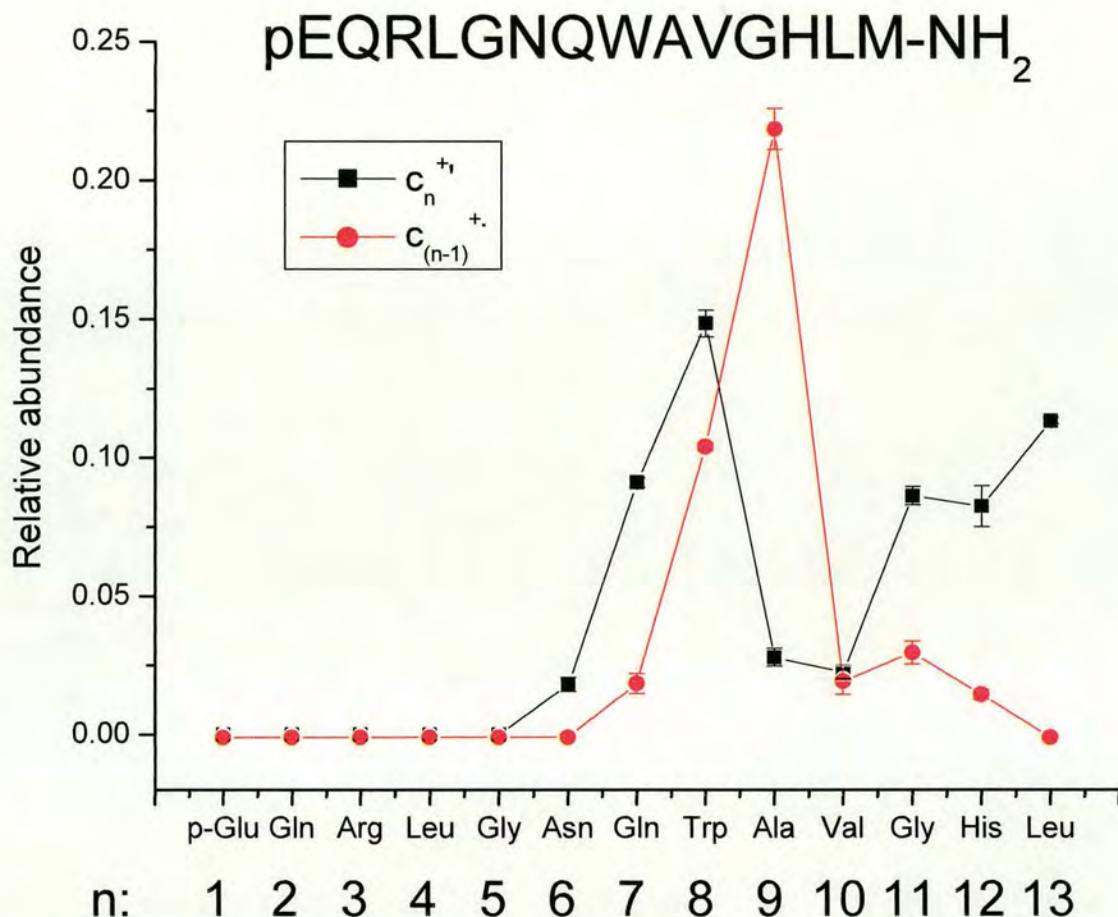
Figure 3.22 Relative intensity of c_n^+ and $c_{(n-1)}^{+\bullet}$ fragment ions for ECD of $[BOMB]^{2+}$. Error bars represent standard deviation obtained for 3 spectra (of 100 time domain transients each) taken at hexapole accumulation times of 0.7, 0.8 and 0.9 s. (50 ms electron irradiation).

121
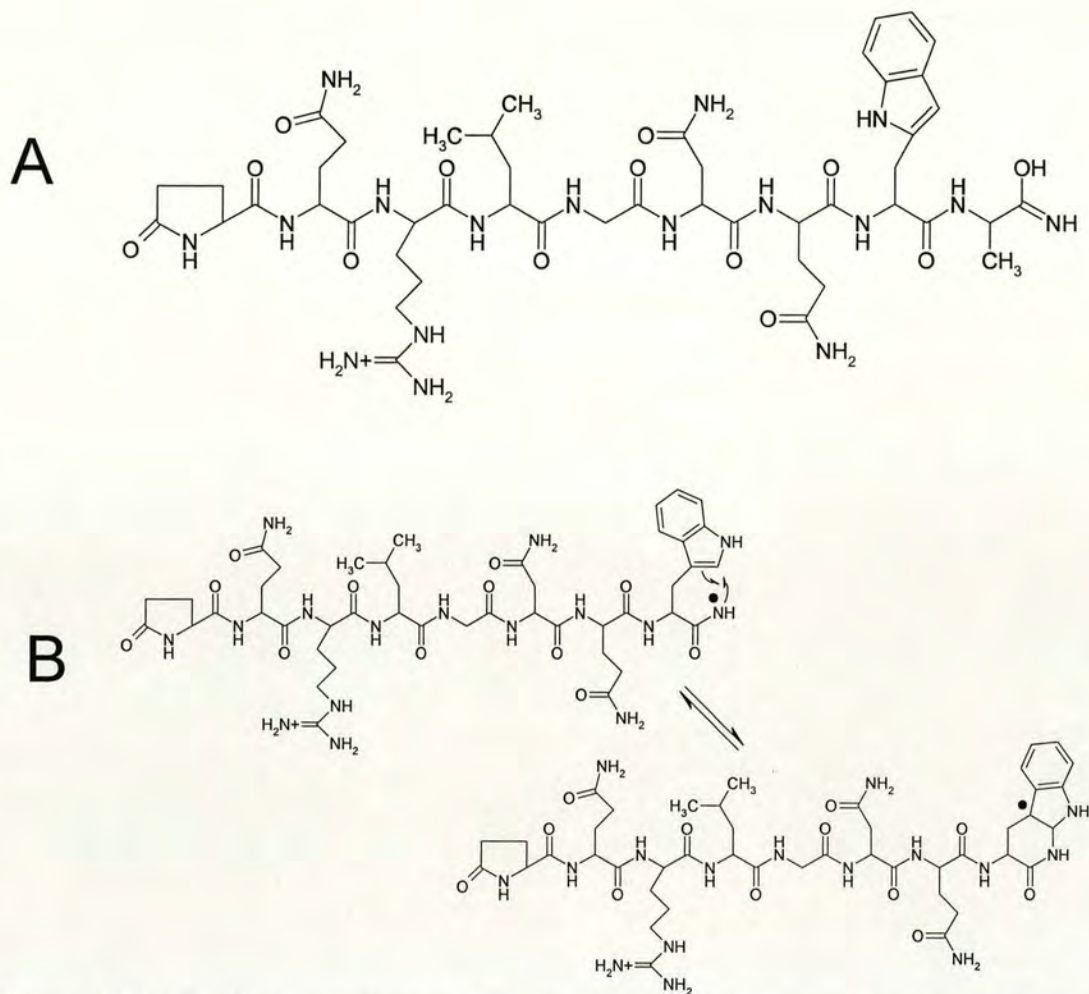
Figure 3.23 Proposed structures (stick structures) for hydrogen atom transfer to alanine{9} backbone carbonyl giving rise to 'normal' alpha-cleavage (A) $c_9^{+\prime}$, and 'reverse' alpha-cleavage (B) $c_8^{+\bullet}$, which may be able to form a resonantly stabilised cyclic structure.

3.4.5 ECD of [LHRH_RH]$^{2+}$

The relative fragmentation pattern for ECD of [LHRH_RH]$^{2+}$ is shown in Figure 3.24. The fragmentation pattern is dominated by $c_6^{+\prime}$, which implies N-C$_\alpha$ cleavage on the C-terminal side of tryptophan{6}. One 'reverse' cleavage is seen for hydrogen atom transfer to the backbone carbonyl of histidine{8}, yielding abundant $c_7^{+\bullet}$. However, the corresponding 'normal' alpha-cleavage $c_8^{+\prime}$ cannot be seen due to the presence of proline{9} C-terminal to tryptophan{8}. Hence, the competition between the 'normal' and 'reverse' cleavages cannot be assessed. The low

122

abundance of $c_6^{+\bullet}$ is somehow surprising, given that the presence of a tryptophan side-chain appears to favour $c_8^{+\bullet}$ in the case of $[BOMB]^{2+}$.

pERWSYWLHPG-OH

Figure 3.24 Relative abundance of $c_n^{+\prime}$ and $c_{(n-1)}^{+\bullet}$ fragment ions for ECD of doubly charged $[LHRH\_RH]^{2+}$ peptide showing amino acid backbone carbonyl where hydrogen atom would have to be transferred to for that fragment to be observed. No error bars are present, since peak intensities only taken from Figure 3.5.

3.4.6 ECD of $[LHRH\_HH]^{2+}$

The relative c and z fragmentation patterns of $[LHRH\_HH]^{2+}$ are shown in Figures 3.25 and 3.26. Since in this case complementary fragmentation is seen (i.e. N and C-terminal fragments), the trends in both can be observed for the same ion. Electron capture at the side-chain of histidine{2} would result in C-terminal fragments, whereas electron capture at the side-chain of histidine{8} would result in N-terminal fragment ions according to the *hot hydrogen re-arrangement* mechanism.

The higher abundances of fragments $c_6^{+\bullet}$ (Figure 3.28) and $z_6^{+'}$ (Figure 3.29) as opposed to their corresponding 'normal' cleavage fragments supports the earlier findings that the 'reverse' cleavage takes place in the vicinity of a tryptophan residue. The higher abundance of $c_2^{+\bullet}$ as opposed to $c_3^{+'}$ (Figure 3.27) is odd at first sight, since this constitutes a small odd-electron c fragment. Nevertheless, it is conceivable that the histidine side-chain could stabilise the electron on the amide nitrogen by forming a resonantly stabilised cyclic ring (see Figure 3.27 B), in analogy to the tryptophan side-chain in [BOMB]$^{2+}$ (Figure 3.23 B). In fact, a similar ring stabilisation involving the histidine side-chain has also been proposed by Wysocki and co-workers to explain the high occurrence of peptide bond cleavage C-terminal to histidine to yield b fragments [9,10].
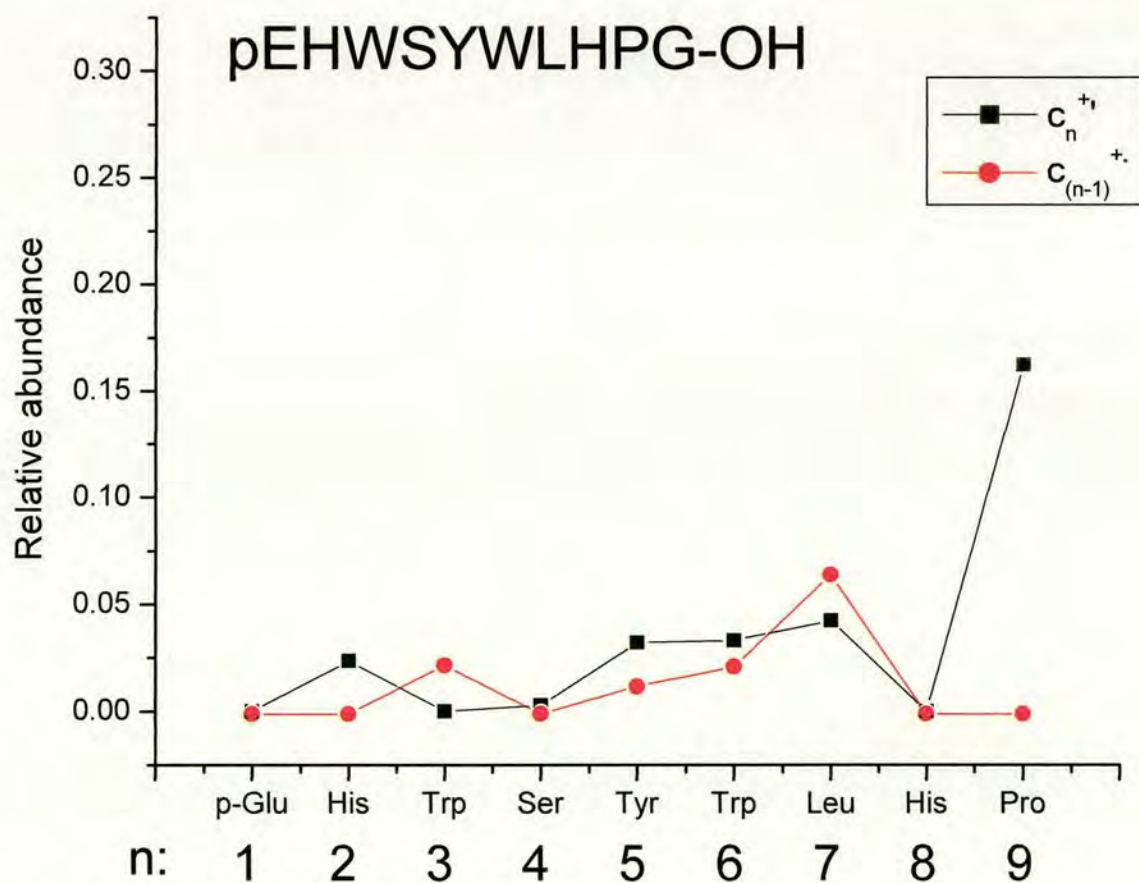
Figure 3.25 Relative abundances of $c_n^{+'}$ and $c_{(n-1)}^{+\bullet}$ fragment ions for ECD of [LHRH_HH]$^{2+}$ relative to sum of c and z fragment ions combined. No error bars are present, since peak intensities only taken from Figure 3.8.

pEHWSYWLHPG-OH
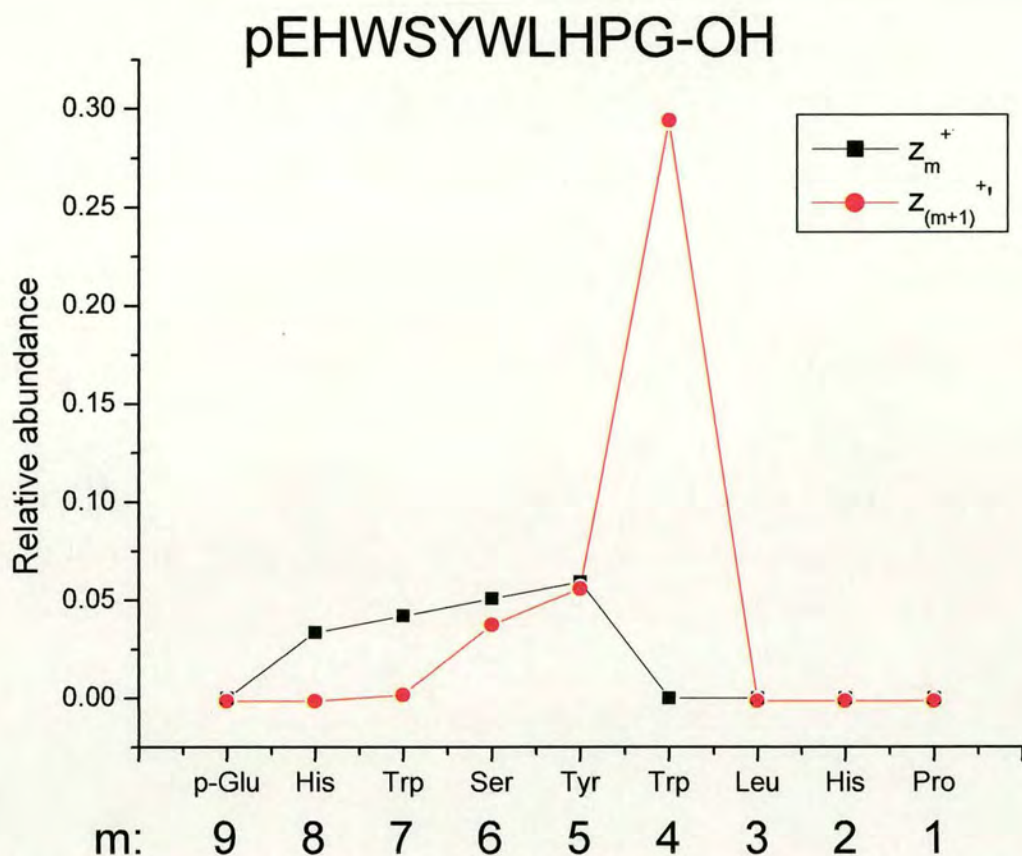
Figure 3.26 Relative abundances of $z_m^{+\bullet}$ and $z_{(m+1)}^{+'}$ fragment ions for ECD of $[LHRH\_HH]^{2+}$ relative to sum of c and z fragment ions combined. No error bars are present since peak intensities only taken from Figure 3.8.
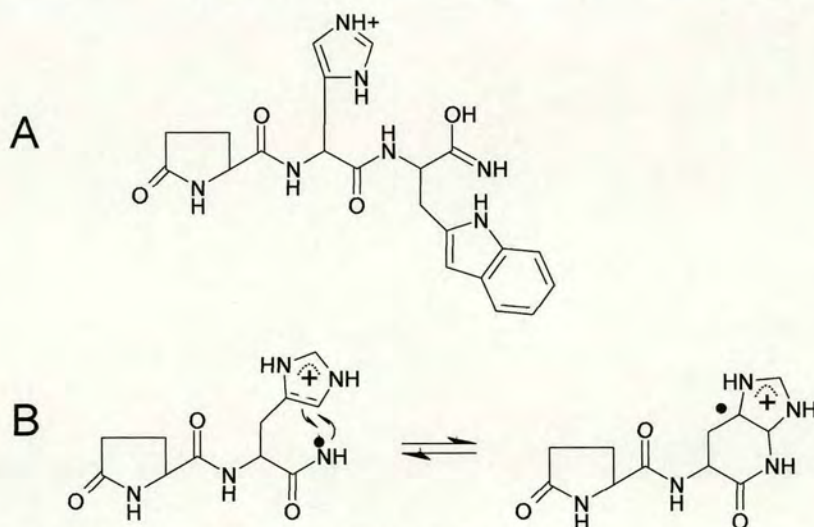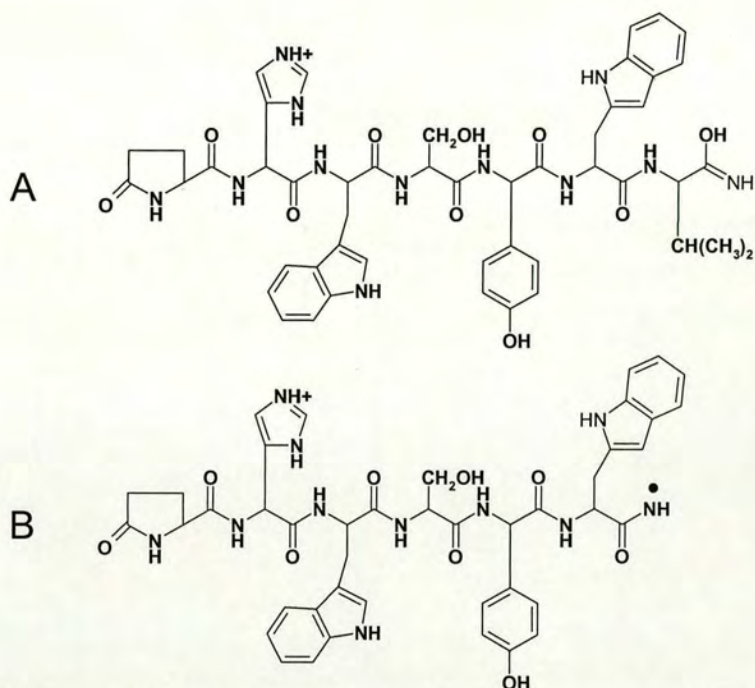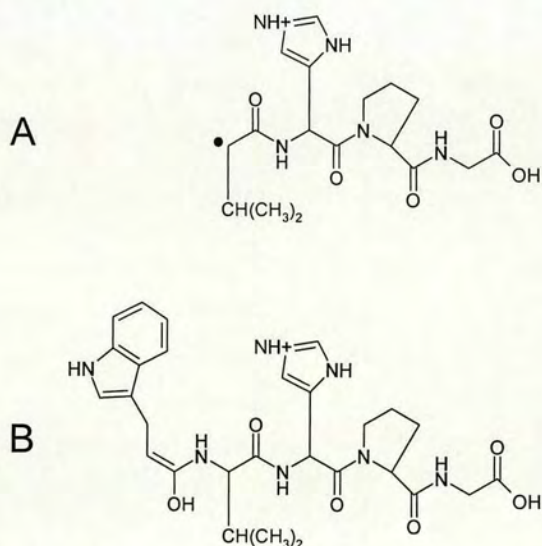

Figure 3.27 Proposed stick structures for hydrogen capture at tryptophan{3} backbone carbonyl, yielding a normal alpha-cleavage (A) $c_3^{+'}$ fragment ion, and reverse alpha-cleavage (B) $c_2^{+\bullet}$ fragment ion, which could re-arrange to a resonantly stabilised structure.

Figure 3.28 Proposed stick structures for hydrogen capture at leucine{7} backbone carbonyl of [LHRH_HH]$^{2+}$, yielding a 'normal' alpha-cleavage (A) $c_7^{+'}$ fragment ion, and a 'reverse' alpha-cleavage (B) $c_6^{+\bullet}$ fragment ion, which could again form cyclic structure with tryptophan side-chain (as shown in Figure 3.23).

Figure 3.29 Proposed stick structures for hydrogen capture at tryptophan{6} backbone carbonyl of [LHRH_HH]$^{2+}$, yielding a 'normal' alpha-cleavage (A) $z_4^{+\bullet}$ fragment ion, and a 'reverse' alpha-cleavage (B) $z_5^{+'}$ fragment ion.

3.4.6 ECD of $[LHRH\_RR]^{2+}$

The relative c and z fragment abundances for ECD of $[LHRH\_RR]^{2+}$ are shown in Figures 3.30 and 3.31. The relative c fragmentation pattern in Figure 3.30 is similar to the relative c fragmentation pattern for ECD of $[LHRH\_HH]^{2+}$ (Figure 3.25), with c_9^{+}' being the most abundant fragment and $c_6^{+\bullet}$ more abundant than c_7^{+}'. This re-confirms the stability of $c_6^{+\bullet}$, which can be rationalised by the tryptophan side-chain.

The relative z fragmentation pattern for $[LHRH\_RR]^{2+}$ (Figure 3.31) shows a striking abundance of z_m^{+}' fragments, as opposed to $z_m^{+\bullet}$ fragments. Even-electron z fragments as large as z_8^{+}' and z_9^{+}' are observed for ECD of $[LHRH\_RR]^{2+}$, which is not the case for any of the other LHRH variants. The most pronounced discrimination, in terms of a competition between 'normal' and 'reverse' alpha-cleavage, occurs for $z_5^{+\bullet}$ and z_6^{+}', the structures of which are shown in Figure 3.32. The low abundance of $z_5^{+\bullet}$ is surprising, given that the fragment structure is the same as for ECD of $[LHRH\_(l)W]^{2+}$, for which the abundance of $z_5^{+\bullet}$ is higher (see Figure 3.19). Nevertheless, the gas-phase structures of $[LHRH\_RR]^{2+}$ and $[LHRH\_(l)W]^{2+}$ may be different, which could explain different fragmentation behaviour.
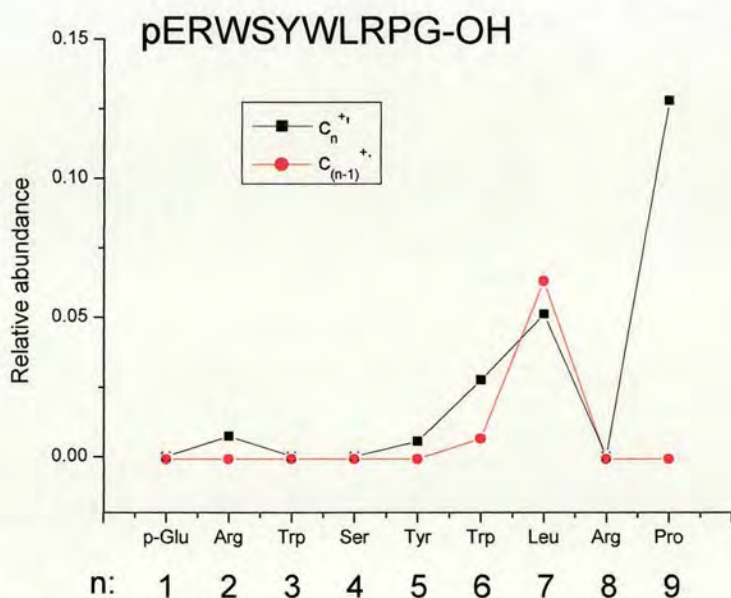
Figure 3.30 Relative abundances of c_n^{+}' and $c_{(n-1)}^{+\bullet}$ fragment ions for ECD of $[LHRH\_RR]^{2+}$ relative to sum of c and z fragment ions combined. No error bars are present, since peak intensities only taken from Figure 3.9.
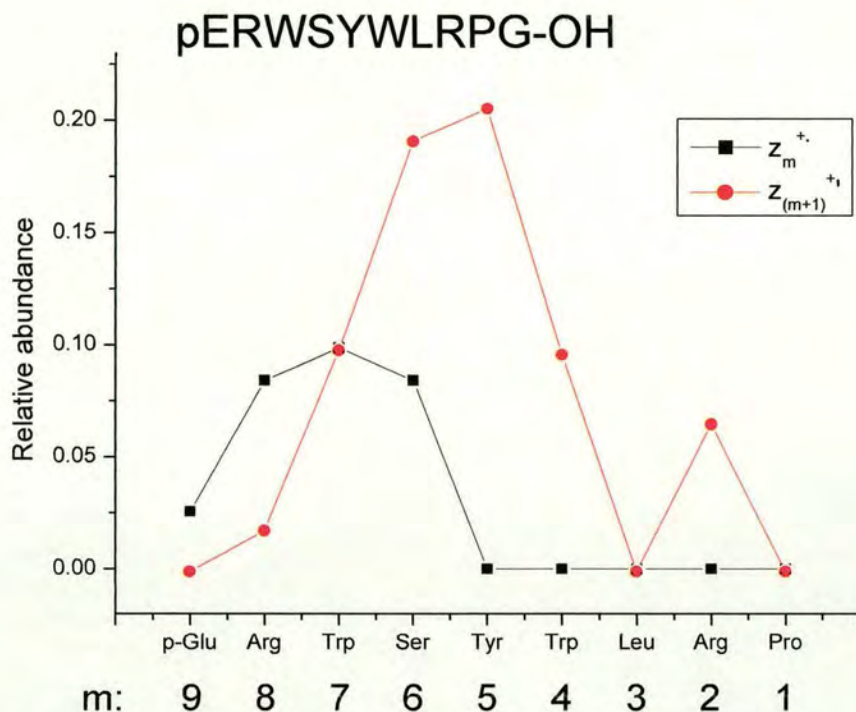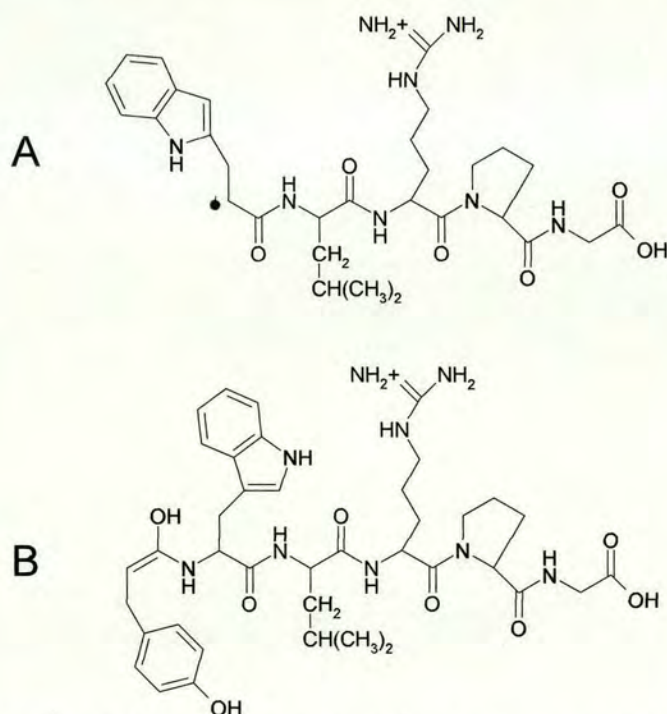
pERWSYWLRPG-OH

Figure 3.31 Relative abundances of $z_m^{+\bullet}$ and $z_{(m+1)}^{+'}$ fragment ions for ECD of $[LHRH\_RR]^{2+}$ relative to sum of c and z fragment ions combined. No error bars are present since peak intensities only taken from Figure 3.9.

Figure 3.32 Proposed stick structures for hydrogen capture at tyrosine{5} backbone carbonyl of $[LHRH\_RR]^{2+}$, yielding a 'normal' alpha-cleavage (A) $z_5^{+\bullet}$ fragment ion, and a 'reverse' alpha-cleavage (B) $z_6^{+'}$ fragment ion.

128

3.4.7 ECD of $[BK]^{2+}$ and $[Ac\_BK]^{2+}$

Finally, the relative c fragment abundances for $[BK]^{2+}$ and $[Ac\_BK]^{2+}$ are shown in Figures 3.33 and 3.34. The patterns look similar, although $c_8^{+\prime}$ is much more abundant for $[Ac\_BK]^{2+}$ (see Figure 3.34) than for $[BK]^{2+}$ (see Figure 3.33). The higher abundance of $c_5^{+\bullet}$ relative to $c_6^{+\prime}$ (which is not observed) can be rationalised due to the presence of a proline residue (in position 7) on the C-terminal side of serine{6}. These bradykinin peptides have 3 proline residues in their sequences, which limits the backbone cleavages that can be observed and therefore makes these peptides less suitable to study the relative occurrence of radical and even-electron c and z fragments. Further, it is not clear how to explain the appearance of an $a_8^{+\bullet}$ fragment, which could be due to further fragmentation from $c_8^{+\bullet}$, as suggested by Tureček [38], but could also be due to other mechanisms. The relative abundances of $z_m^{+\bullet}$ and $z_m^{+\prime}$ are not shown since they are of little informative value.
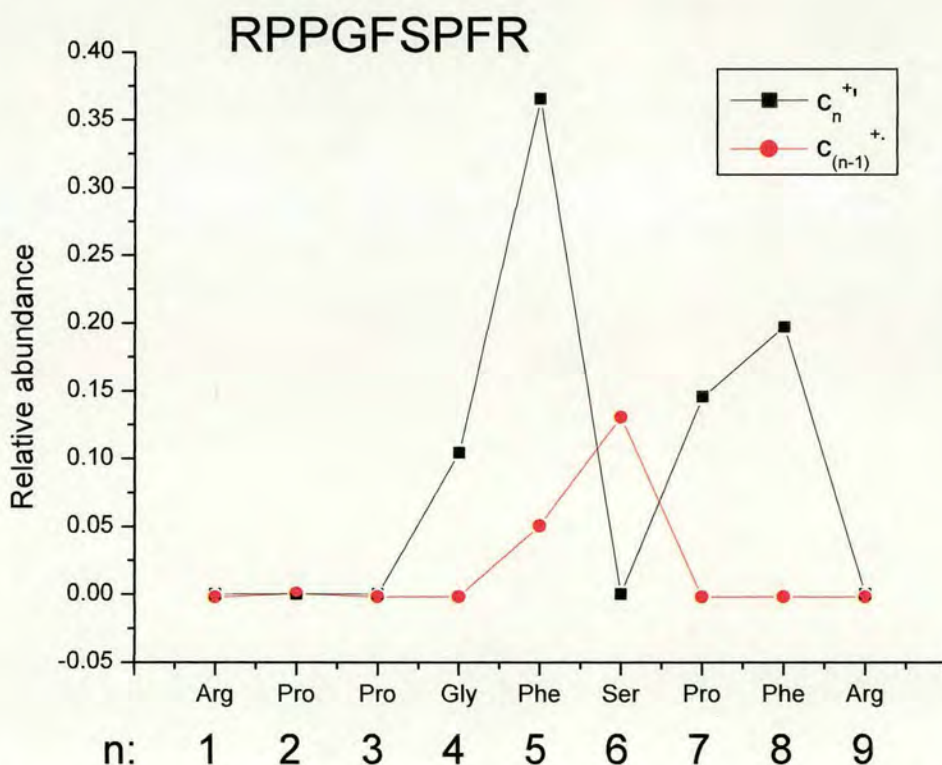
Figure 3.33 Relative abundances of $c_n^{+\prime}$ and $c_{(n-1)}^{+\bullet}$ fragment ions for ECD of $[BK]^{2+}$. No error bars are present, since peak intensities only taken from Figure 3.10.
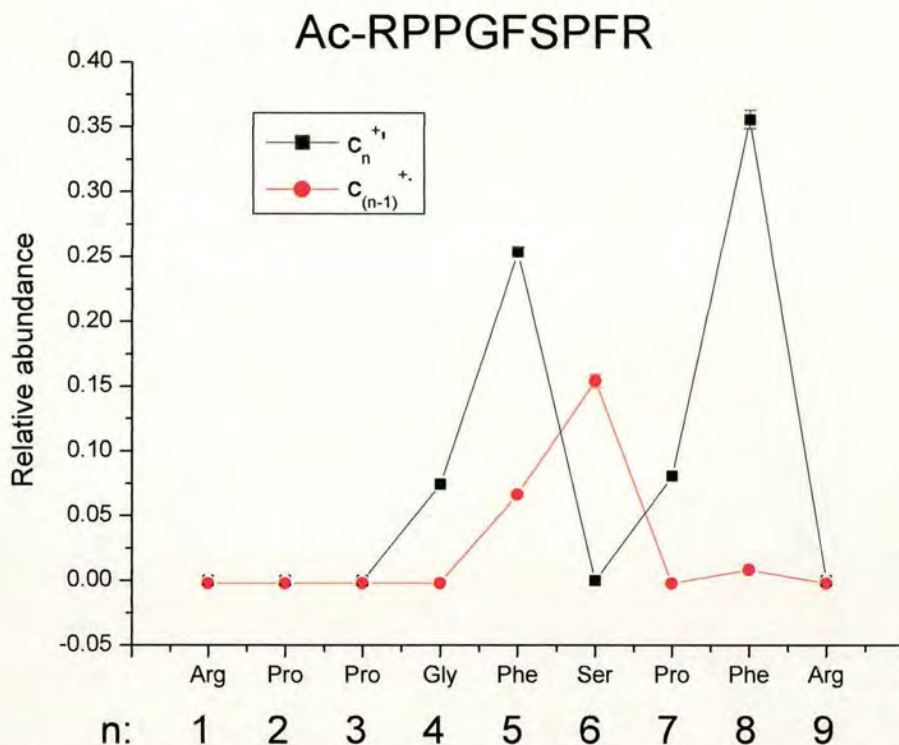
Ac-RPPGFSPFR

Figure 3.34 Relative abundances of $c_n^{+\prime}$ and $c_{(n-1)}^{+\bullet}$ fragment ions for ECD of $[Ac\_BK]^{2+}$. Error bars represent standard deviation obtained for 3 spectra (of 100 time domain transients each) taken at hexapole accumulation times of 0.9 and 1.0 s. (30 ms electron irradiation).

3.4.8 Conclusions

This study has shown that the relative occurrence of radical and even-electron c and z fragments for these peptides could be explained in terms of a competition between a 'normal' and 'reverse' alpha-cleavage. The occurrence of even-electron $z_m^{+\prime}$ fragments appears to be favoured for smaller size fragments, whereas larger size fragments are more likely to be odd-electron $z_m^{+\bullet}$ fragments. This observation may be linked to the vicinity of the radical C_α in $z_m^{+\bullet}$ to the protonated site; i.e. $z_m^{+\bullet}$ is favoured further away from the protonated group. These results also suggest that the amino acid side-chain adjacent to the radical C_α in $z_m^{+\bullet}$ can play a role in stabilising the fragment (compare abundance of $z_5^{+\bullet}$ for ECD of $[LHRH\_HR\_NH_2]^{2+}$ and ECD of $[LHRH\_(l)W]^{2+}$) if the gas-phase structure is right (compare abundance of $z_5^{+\bullet}$ for ECD of $[LHRH\_(l)W]^{2+}$ and ECD of $[LHRH\_(d)W]^{2+}$).

In the case of c fragments, even-electron $c_n^{+\prime}$ are much more common than the corresponding odd-electron $c_n^{+\bullet}$ ions for ECD of these peptides. Odd-electron $c_n^{+\bullet}$ fragments are observed for larger fragments and there appears to be a correlation between the occurrence of $c_n^{+\bullet}$ and the presence of an aromatic hetero-cyclic side-chain (i.e. tryptophan and histidine side-chains). It is conceivable that $c_n^{+\bullet}$ forms a cyclic structure with the side-chain as shown for $[BOMB]^{2+}$ (Figure 3.23 B), $[LHRH\_HH]^{2+}$ (Figure 3.27 B and 3.28 B) and $[LHRH\_RR]^{2+}$.

Thus, in terms of the discussion above, the relative occurrence of the 'normal' and 'reverse' alpha-cleavages would be due to a combination of the immediate chemical environment of the amino acid side-chains where the hydrogen atom is transferred to and the gas-phase conformation of the peptide. In terms of the results for $[LHRH\_(l)W]^{2+}$ and $[LHRH\_(d)W]^{2+}$ the relative fragment abundances in ECD may be able to distinguish between amino acid stereoisomers.

3.5 Relative fragmentation

In analogy with the discussion in sections 3.3 and 3.4, in this section I assume that the ECD process is fast and that the efficiency of hydrogen atom transfer is linked to the ease of abstraction from the protonated site and ease of acceptance from the hydrogen atom acceptance site. For the molecule $[LHRH\_HR]^{2+}$ this implies that only electron capture at the histidine side-chain gives rise to efficient backbone fragmentation, whereas electron capture at the arginine side-chain would merely result in neutral loss fragments. For $[LHRH\_HR]^{2+}$ the hydrogen affinities of the accepting backbone carbonyls are likely to be highly equivalent. Hence, the rate of hydrogen atom transfer is probably linked to the relative vicinity of the electron-capturing protonated site (i.e. H^+) to the hydrogen-accepting backbone carbonyl. Table 3.5 shows the relative occurrences of the different processes in ECD of $[LHRH\_HR]^{2+}$. Only 1 in 4 electron captures results in hydrogen atom desorption for this species (see Table 3.5), which suggests that the intra-molecular hydrogen atom transfer is surprisingly efficient. It is therefore conceivable that hydrogen bonding plays a role in hydrogen atom transfer. Here, it is put forward that preferential hydrogen bonding of the electron-capturing sites to particular backbone carbonyls enhances backbone cleavage at those sites. Since the ECD mechanism is assumed to be *non-ergodic*, the structure of the peptide ion <u>before</u> electron capture would determine the fragmentation pattern. This will be discussed in terms of modelling peptide structures in Chapter 4.

Process	Relative occurrence / %
Reduced species $[M+2H]^{+\bullet}$	8.5
Hydrogen desorption $[M+H]^+$	25
Sum of neutral losses	13.5
Sum of ECD backbone fragment ions	53

Table 3.5 Relative occurrence of different processes in ECD of $[LHRH\_HR]^{2+}$.

3.5.1 ECD of LHRH variants

Figure 3.35 shows the relative z fragment intensities for $[LHRH\_HR]^{2+}$ associated with respective backbone carbonyls. This Figure was generated by summing the relative intensities for $z_m^{+\bullet}$ and $z_{(m+1)}^{+'}$ in Figure 3.15. Note that $z_m^{+\bullet}$ and $z_{(m+1)}^{+'}$ are due to hydrogen atom transfer to the same backbone carbonyl, according to the 'normal' and 'reverse' alpha-cleavage mechanisms (see Figure 3.13). Note also that the z_3^+ fragment intensity is not included in this study due to the fact that the $z_2^{+\bullet}$ does not appear due to the cyclic structure of proline, and that the c_9^+ fragment is not included due to the ambiguity with the 59 Da loss (see section 3.2).
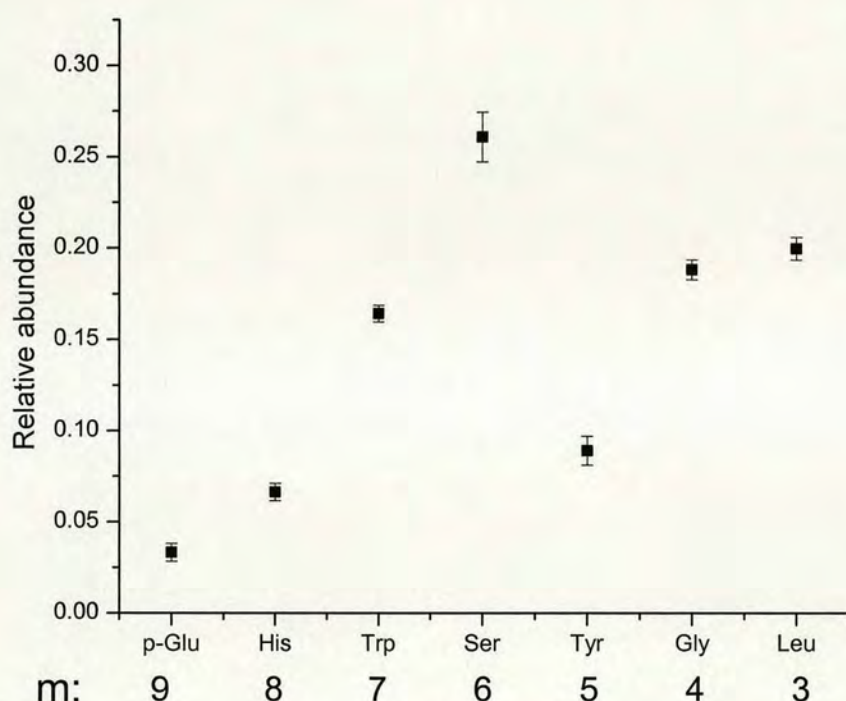
Figure 3.35 Relative ECD z_m^+ (= $z_m^{+\bullet}$ + $z_{(m+1)}^{+'}$) fragment abundances for $[LHRH\_HR]^{2+}$ associated with respective backbone carbonyls. (data taken from Figure 3.15) (9.4 T APEX III, dispenser cathode)

The relative ECD z fragment abundances for $[LHRH\_HR]^{2+}$ as determined with the electron filament are shown in Figure 3.36. Figures 3.35 and 3.36 show very similar patterns, which supports the assumption that the relative fragmentation pattern in ECD is reproducible and does not depend on the experimental parameters as much as in CAD.
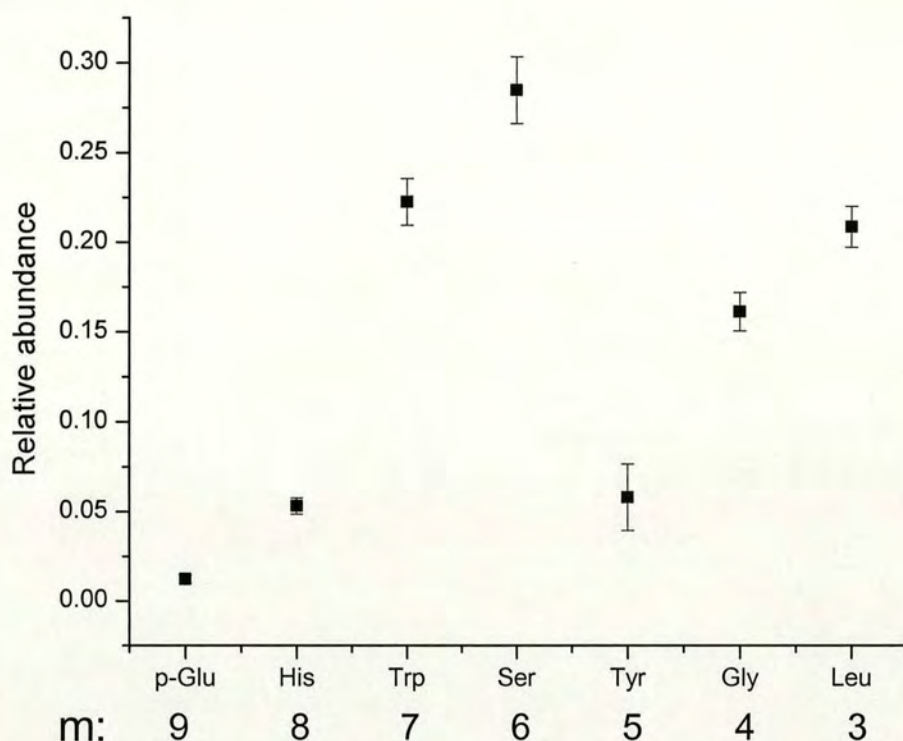
Figure 3.36 Relative ECD z_m^+ (= $z_m^{+\bullet}$ + $z_{(m+1)}^{+\prime}$) fragment abundances for [LHRH_HR]$^{2+}$ associated with respective backbone carbonyls. (data taken from Figure 3.16) (3 T APEX II, rhenium filament)

The relative ECD z fragment intensities for [LHRH_HR]$^{2+}$, [LHRH_HR_NH$_2$]$^{2+}$ (data from Figure 3.18), [LHRH_(l)W]$^{2+}$ (data from Figure 3.19) and [LHRH_(d)W]$^{2+}$ (data from Figure 3.21) are overlaid in Figure 3.37. The relative fragmentation pattern for these peptides is very similar. In terms of the efficiency of hydrogen atom transfer this could imply that these peptides have similar gas-phase conformations (see Chapter 4).
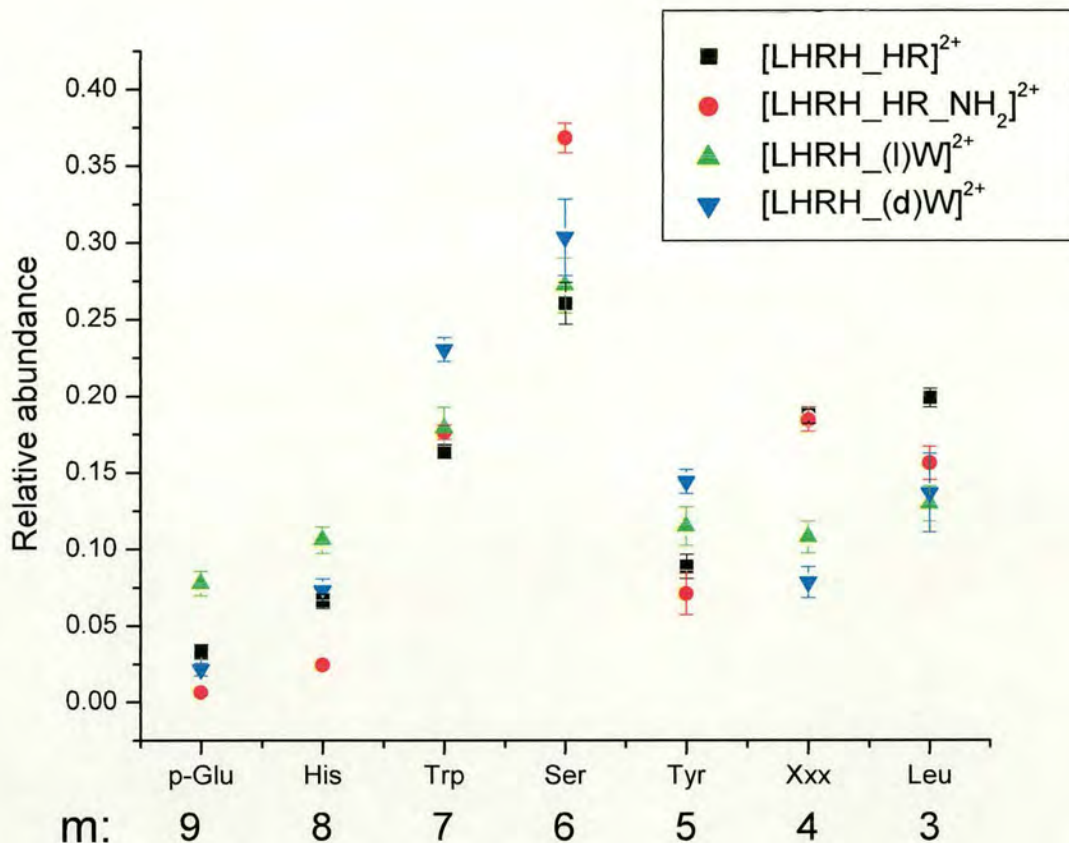
Figure 3.37 Relative ECD z_m^+ (= $z_m^{+\bullet}$ + $z_{(m+1)}^{+'}$) fragment abundances for [LHRH_HR]$^{2+}$, [LHRH_HR_NH$_2$]$^{2+}$, [LHRH_(l)W]$^{2+}$ and [LHRH_(d)W]$^{2+}$, where Xxx denotes the variable amino acid in position 6.

3.5.2 ECD of [BOMB]$^{2+}$

The relative ECD c and z fragment intensities for [BOMB]$^{2+}$ are shown in Figure 3.38. The ECD mass spectrum of [BOMB]$^{2+}$ (Figure 3.7) also includes two z fragments ($z_{12}^{+\bullet}$ and $z_{13}^{+\bullet}$) which are included in this analysis.

The relative cleavage efficiency for [BOMB]$^{2+}$ in Figure 3.38 is quite different to the one for the LHRH variants in Figure 3.37. One of the higher cleavage sites in both systems is on the C-terminal side of tryptophan, which has been reported before [12]. Otherwise, the cleavage efficiency over the length of the peptide does vary significantly and is seemingly not connected to the hydrogen atom affinity of

the backbone carbonyl; whereas backbone cleavage associated with the backbone carbonyl of glycine{11} is enhanced, the cleavage associated with glycine{5} is of low abundance.
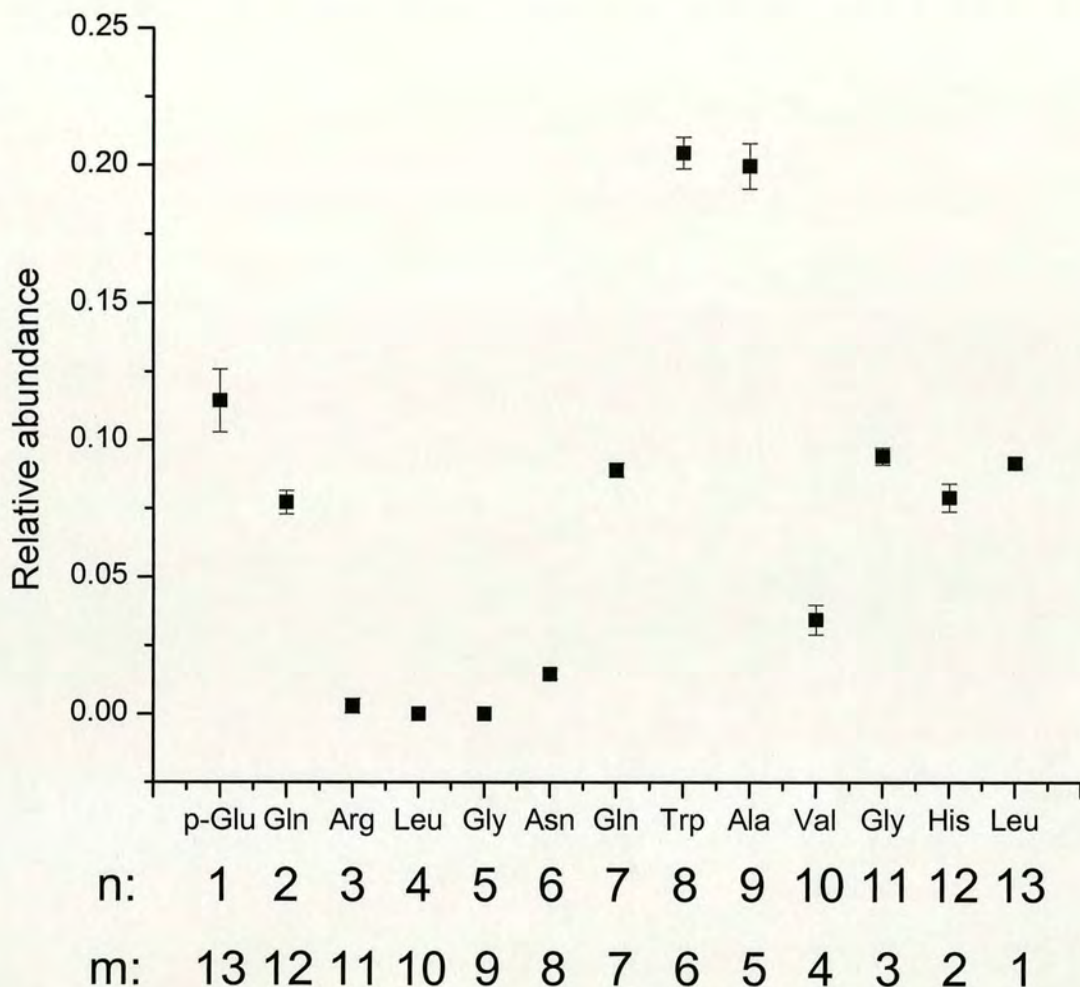
Figure 3.38 Relative ECD c_n^+ ($= c_n^+ + (c-1)_{(n-1)}^{+\bullet}$) and z_m^+ ($= z_m^{+\bullet} + z_{(m+1)}^{+\prime}$) fragment abundances for $[BOMB]^{2+}$.

The highest fragments are in general not due to hydrogen atom transfer to a backbone carbonyl in the immediate vicinity of the histidine side-chain from a stick structure point of view. Since the stick structure of the peptide does not give a representative idea of the actual vicinity of the histidine side-chain to various backbone carbonyls, the gas-phase structures of the above peptides will be modelled using mechanical force field modelling (Amber force field [47,48]) in Chapter 4.

3.5 Conclusions

This mechanistic study of ECD of quite well-defined peptide systems has shown that the spectra contain much information which can be useful to obtain further primary sequence information. In this light, the neutral losses in ECD have shown to be of diagnostic value for the presence of some amino acids. The 82 Da loss in particular is highly specific and sensitive to the presence of the histidine residue. Some other losses (17 Da, 59 Da) are more ambiguous to interpret in these peptides and more work needs to be done in that area. The 131 Da loss (isobaric with the loss of the tryptophan side-chain plus a hydrogen atom) may be a probe of tryptophan protonation.

The correlation study of linking the gas-phase basicity of the protonated site to the occurrence of charged fragments has shown that the most basic site is preferentially incorporated in the charged fragment. This is particularly useful in the *de novo* sequencing of tryptic digest peptides, since these often give doubly charged peptides in the gas-phase [10] and they generally contain a more basic amino acid on the C-terminus (arginine or lysine), therefore resulting in primarily C-terminal fragment ions.

The occurrence of radical and even-electron c and z ions was investigated in terms of the 'normal' and 'reverse' alpha-cleavage mechanisms. This study suggests that the immediate chemical environment can play a role in stabilising particular odd-electron fragments. Further, the differences found for the doubly charged peptides of $[LHRH\_(l)W]^{2+}$ and $[LHRH\_(d)W]^{2+}$ suggest that the peptide conformation may play a role in determining which mechanism is favoured. Modelling of peptide ions will be shown in Chapter 4.

The relative abundances of fragments linked to a particular backbone carbonyl show that these do vary significantly over the length of the peptide. For the following discussion in Chapter 4, the assumption is made that ECD is based on a *non-ergodic* mechanism and that the transfer efficiency of the radical hydrogen from

the protonated site to the backbone carbonyl is related to the vicinity of both groups prior to electron capture.

References:

(1)   Loo, J. A.; Udseth, H. R.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 207.

(2)   Eng, J. K.; McCormack, A. L.; Yates, J. R., 3rd *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976.

(3)   Yates, J. R., 3rd; Eng, J. K.; McCormack, A. L. *Anal. Chem.* **1995**, *67*, 3202.

(4)   Yates, J. R., 3rd; Eng, J. K.; McCormack, A. L.; Schieltz, D. *Anal. Chem.* **1995**, *67*, 1426.

(5)   Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.

(6)   Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *J. Mass Spectrom.* **1996**, *31*, 339-350.

(7)   Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 5142-5154.

(8)   Tsaprailis, G.; Somogyi, A.; Nikolaev, E. N.; Wysocki, V. H. *Int. J. Mass Spectrom.* **2000**, *195*, 467-479.

(9)   Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399-1406.

(10)   Tabb, D. L.; Smith, L. L.; Breci, L. A.; Wysocki, V. H.; Lin, D.; Yates, J. R., III *Anal. Chem.* **2003**, *75*, 1155-1163.

(11)   Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

(12)   Kruger, N. A.; Zubarev, R. A.; Carpenter Barry, K.; Kelleher, N. L.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom. Ion Proc.* **1999**, *182,183*, 1-5.

(13)   Budnik, B. A.; Nielsen, M. L.; Olsen, J. V.; Haselmann, K. F.; Hörth, P.; Haehnel, W.; Zubarev, R. A. *Int. J. Mass Spectrom.* **2002**, *219*, 283-294.

(14)   Cooper, H. J.; Hudgins, R. R.; Hakansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 241-249.

(15)   Falick, A. M.; Hines, W. M.; Medzihradszky, K. F.; Baldwin, M. A.; Gibson, B. W. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 882.

(16)   Haselmann, K. F.; Budnik, B. A.; Kjeldsen, F.; Polfer, N. C.; Zubarev, R. A. *Eur. J. Mass Spectrom.* **2002**, *8*, 461-469.

(17)   Breuker, K.; Oh, H.; Cerda, B. A.; Horn, D. M.; McLafferty, F. W. *Eur. J. Mass Spectrom.* **2002**, *8*, 177-180.

(18)   Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn David, M.; Carpenter Barry, K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.

(19)   Kruger, N. A.; Zubarev, R. A.; Horn, D. M.; McLafferty, F. W. *Int. J. Mass Spectrom.* **1999**, *185,186,187*, 787-793.

(20)   Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.

(21)   Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563-573.

(22)   Håkansson, K.; Emmett, M. R.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2001**, *73*, 3605-3610.

(23)   Kjeldsen, F.; Budnik, B. A.; Haselmann Kim, F.; Jensen, F.; Zubarev, R. A. *Chem. Phys. Lett.* **2002**, *356*, 201-206.

(24)   Nold, M. J.; Cerda, B.; Wesdemiotis, C. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 1-8.

(25)   Chan, T. W. D.; Ip, W. H. H. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1396-1406.

(26)   Tsybin, Y. O.; Witt, M.; Baykut, G.; Kjeldsen, F.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1759-1768.

(27)   Freitas, M. A.; Marshall, A. G. *Int. J. Mass Spectrom.* **1999**, *182*, 221-231.

(28)   Mao, D.; Douglas, M. J. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 85-94.

(29)   Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1997**, *119*, 11988-11989.

(30)   Julian, R. R.; Hodyss, R.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2001**, *123*, 3577-3583.

(31)   Julian, R. R.; Beauchamp, J. L.; Goddard III, W. A. *J. Phys. Chem. A* **2002**, *106*, 32-34.

(32)   Strittmatter, E. F.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 6069-6076.

(33)   Scherer, J. J.; Voelkel, D.; Rakestraw, D. J.; Paul, J. B.; Collier, C. P.; Saykally, R. J.; O'Keefe, A. *Chem. Phys. Lett.* **1995**, *245*, 273-280.

(34)   Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 12956-12957.

(35)   Nielsen, M. L.; Budnik, B. A.; Haselmann, K. F.; Olsen, J. V.; Zubarev, R. A. *Chem. Phys.Lett.* **2000**, *330*, 558-562.

(36)   Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Anal. Chem.* **2001**, *73*, 2998-3005.

(37)   Bakken, V.; Helgaker, T.; Uggerud, E. In *16th IMSC*: Edinburgh, UK, 2003.

(38)   Turecek, F. *J. Am. Chem. Soc.* **2003**, *125*, 5954-5963.

(39)   Turecek, F.; Syrstad, E. A.; Seymour, J. L.; Chen, X. In *16th IMSC*: Edinburgh, UK, 2003.

(40)   Siegbahn, P. E. M.; Blomberg, M. R. A.; Crabtree, R. H. *Theor. Chem. Acc.* **1997**, *97*, 289-300.

(41)   Datz, S. *J. Phys. Chem. A* **2001**, *105*, 2369-2373.

(42)   McLafferty, F. W., Private communication.

(43)   Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. *Eur. J. Mass Spectrom.* **2002**, *8*, 337-349.

(44)   Haselmann Kim, F. In *Chemistry Department*; University of Southern Denmark: Odense, 2002; p 111.

(45)   Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8949-8958.

(46)   McDonnell, L. A.; Giannakopulos, A. E.; Derrick, P. J.; Tsybin, Y. O.; Hakansson, P. *Eur. J. Mass Spectrom.* **2002**, *8*, 181-189.

(47)   Weiner, S. J.; Kollmann, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta Jr., S.; Weiner, P. J. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.

(48)   Weiner, S. J.; Kollmann, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230-252.

Chapter 4
Molecular mechanics force field modelling of peptide gas-phase ions

4.1 Introduction

The LHRH (luteinizing hormone releasing hormone) peptides are a family of decapeptides that are also known as GnRH (gonadropin-releasing hormone) peptides which play a key physiological role as a mediator of neuroendocrine regulation in the mammalian reproductive system [1,2]. Fifteen structural variants of GnRH have been identified [3]. All GnRH peptides are characterised by the conservation of the N-terminal residues (pGlu-His-Trp-Ser) and the C-terminal residues (Pro-Gly-NH$_2$) with a few exceptions. In mammalian GnRH (mGnRH) (sequence pEHWSYGLRPG-NH$_2$, i.e. [LHRH_HR_NH$_2$]$^{2+}$ is doubly protonated variant of mGnRH) residues 5-8 are believed to adopt a β-II' turn conformation which allows interaction with mGnRH type I receptors [4]. The identity of residue 6 plays a particular role in determining whether the β-II' turn conformation is adopted or not. Nuclear magnetic resonance (NMR) studies for mGnRH have shown that a β-type turn is present about Gly{6} (i.e. glycine residue in position 6 from N-terminus) [5]. Substitution with a D-amino acid in position 6 increases the receptor binding affinity, which is thought to be due to constraining the conformation to a β-type turn [6]. Determining the structure of these peptides is difficult due to their high flexibility [7]. Guarnieri *et al.* [8] have performed solution phase modelling of mGnRH using the Amber force field [9,10] and the generalised Born/surface area (GB/SA) solvation model of water [11]. They probed the conformational space of the peptides by Monte Carlo simulated random walks[1]. They predicted 'families' of GnRH structures of which 70% have a β-type turn involving residues 5-8. A review of the

---

[1] Monte Carlo-type simulations involve a stochastic approach based on the use of random numbers and probability statistics to investigate a problem. The peptide mGnRH has 35 rotatable torsional angles, which makes it impossible to probe the totality of the conformational space that mGnRH has available. Hence, Guarnieri *et al* have used the Monte Carlo approach where the value of a torsional angles are set to particular values by the 'roll of a die'. This approach allows for a statistical evaluation of the different conformations that mGnRH can adopt.

structural dependence of GnRH peptides and their interaction with the GnRH receptor has been published by Millar and co-workers [12].

So far, no gas-phase structure modelling has been published on GnRH peptides. Since the dielectric constant of the vacuum (ε_{vacuum}= 1) is more similar to the immediate environment of the GnRH type I receptor ($\varepsilon_{peptide/protein}$ = 4) than for water (ε_{water} = 80), the gas-phase is an appropriate 'medium' to study the structures of these systems. Further, hydrogen bonding between e.g. pGlu(1) and Gly-NH$_2$(10) is believed to play a role in stabilising the β-turn conformation [6] and can be studied more readily in the absence of water.

In this Chapter modelling of the GnRH/LHRH peptides investigated in Chapter 3 [LHRH_HR]$^{2+}$, [LHRH_HR_NH$_2$]$^{2+}$, [LHRH_(l)W]$^{2+}$ and [LHRH_(d)W]$^{2+}$ (see Figure 3.35) will be attempted *in vacuo* (i.e. the gas-phase) using the AMBER^2 suite of programs [13]. A correlation is drawn between the relative ECD fragmentation pattern in Chapter 3 (section 3.5) and preferential hydrogen bonding (Hbond) of the protonated sites to particular backbone carbonyls as predicted by modelling. For the sake of simplicity this is referred to as the *ECD/Hbond correlation* study. In the light of the ECD mechanism, the assumption is made that the ECD mechanism is *non-ergodic* and hence the gas-phase structure prior to electron capture determines the dissociation pattern. In particular, preferential hydrogen bonding of the protonated site to particular backbone carbonyls may enhance hydrogen atom transfer to those carbonyls resulting in higher backbone cleavage at those sites. Thus, the applicability of ECD as a tool to study gas-phase structure will be investigated.

The approach to perform force field modelling in this study is outlined in section 4.2. The modelling results for the peptides in this study are presented thereafter [LHRH_HR]$^{2+}$ (section 4.3), [LHRH_HR_NH$_2$]$^{2+}$ (section 4.4), and [LHRH_(l)W]$^{2+}$ and [LHRH_(d)W]$^{2+}$ (section 4.5). Gas-phase structure modelling is also carried out for the larger peptide [BOMB]$^{2+}$ (section 4.6) (for amino acid sequences, see Table 3.1 Chapter 3).

---

^{2} Note that AMBER refers to the computer modelling package, whereas Amber refers to the mechanical force field that is used in this modelling package.

4.2 Approach to force field modelling

4.2.1 Making new residues

Molecular modelling of peptides in this study has been carried out using Amber7 [13] based on the 1999 force field (ff99) [14]. The atom types are mostly those developed by Cornell *et al.* in the 1994 force field [15], however, changes have been made in many torsional parameters.[3] Standard amino acid residues are available in the AMBER suite of programs to build polypeptide chains [15]. These residues are suited for solution phase modelling and hence the standard residues have protonation sites reflected by the basicities of amino acids in solution (i.e. pKa values, see Chapter 1 section 1.4.2). As discussed in Chapter 1, the order of basicities of the amino acids in the gas-phase is different to the one in solution. When carrying out modelling in the gas-phase, it is often necessary to make new residues that reflect the protonation sites in the gas-phase.

Figure 4.1 shows the flow chart for the procedure to make new residues in AMBER that was used in this work. The procedure involves using the program *LeAP* (imbedded in the AMBER suite of programs) which allows visualisation of the residue and saving of the residue as a coordinate file; e.g. pdb file (Protein Data Bank file). Initial electrostatic potential (ESP) charge fitting on the atoms for new amino acid residues is derived from Gaussian 98 [17] single point calculations at HF/6-31G* level of theory. The basic idea underlying the ESP charge fitting procedure is to produce a set of atom-centred point-charges meant to accurately describe the electrostatic potential of the molecule at given multiples of the Van der Waals radii (1.4, 1.6, 1.8 and 2.0). The restrained electrostatic potential (RESP) [18] fitting procedure is then carried out in *antechamber* (part of AMBER suite of programs) in order to minimise high charges on embedded atoms and differences in charges on equivalent atoms. The charges obtained from the ESP charge fitting procedure in

---

[3] Note that the 2002 force field (ff02) is a polarisable variant of ff99, where the charges were determined at the B3LYP/cc-pVTZ//HF/6-31G* level [16]. In the ff02 force field the atomic charges are more similar to 'gas-phase' charges, whereas the charges derived by Cornell *et al.* exaggerate the dipole moment by 10-20 % thereby accounting for the polarisation in an aqueous environment. At the time of modelling these peptides, I was not aware of this new force field (ff02) and hence the 'newest' known variant was used (i.e. ff99).

Gaussian are not equivalent for a static molecule, however, in a dynamic molecule where free bond rotation takes place these charges should be equivalent. Apart from new atomic charges, sometimes also new bond and torsion angles are required when making new residues. The structures of the modified residues as well as the atomic charges are shown in Appendix 4: pyro-glutamic acid (Figure A4.1), neutral C-terminal glycine (Figure A4.2), neutral amidated C-terminal glycine (Figure A4.3), protonated amidated C-terminal glycine (Figure A4.4), neutral amidated C-terminal methionine (Figure A4.5) and protonated amidated C-terminal methionine (Figure A4.6). The atoms in these residues are represented using CPK[4] with colour-coding of the atoms: C = green, H = white, O = red, N = blue and S = yellow. This convention of colour-coding of the atoms will also be employed for other structures shown in this Chapter. The (d) isomer of tryptophan was made by changing the order of atoms in the (l) tryptophan residue and hence did not require new charges to be made.
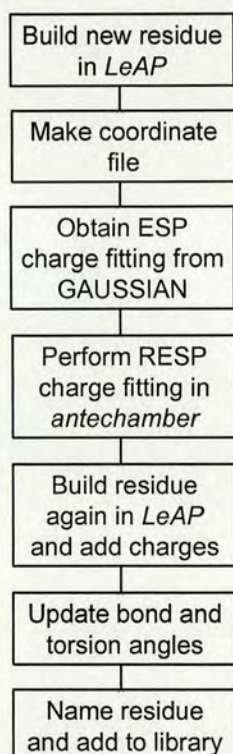
Figure 4.1 Schematic flow chart for procedure to make new amino acid residue in the AMBER suite of programs.

---

[4] All graphical representations of structures were effected using the VMD software (Visual Molecular Dynamics) available on website: http://www.ks.uiuc.edu/Research/vmd/

4.2.2 Simulated annealing and molecular dynamics

Figure 4.2 shows the flow chart outlining the procedure used in investigating the gas-phase structures. The polypeptides in this study are subjected to a simulated annealing procedure involving heating to 800 K[5] followed by controlled cooling to 0 K to give a minimised structure (see Figure 4.3). This minimised structure then serves as a seed structure for the next annealing cycle. In total 100 candidate structures are generated with this procedure, which is thought to give a reasonable probing of the potential energy surface of the possible structures.[6] Since the temperature reflects the velocities of atoms in the molecule, the cooling cycle causes the molecule to minimise its potential energy (i.e. fall into a minimum on the potential energy surface).

A 0 K structure represents a structure where no atomic movement takes place and therefore it needs to be interrogated by molecular dynamics at a higher temperature (e.g. 300 K) to determine its gas-phase stability at higher internal energies. The temperature 300 K represents a reasonable internal energy to probe the secondary structure of peptides [19] using the Amber force field. The hydrogen bonding interactions of the protonated site of lower gas-phase basicity than arginine (i.e. histidine side-chain or C-terminal protonated amide) to backbone carbonyls are determined for the lowest energy structures and these are investigated at higher internal energies with dynamics. Hydrogen bonding in this analysis refers to an atomic distance of < 3 Å for the heavy atom distance N...O and an angle N-H...O > 150°, where 180° refers to linear alignment of N-H...O. The cut-offs employed for hydrogen bonding differ among research groups. Whereas Kurz et $al.$ [20] used distance N...O < 3.5 Å and angle N-H...O > 120°, McDonald et $al.$ [21] used distance N-H...O < 2.5 Å and angle N-H...O > 120°. The reasons for using the stringent definition for hydrogen bonding, as employed in this study, is explained in section 4.3.1. The fundamental time-step in these dynamics calculations is 1 fs

---

[5] As explained in Chapter 1 (section 1.5), the temperature expresses the kinetic energy (K.E.) of the atoms. $K.E. = \frac{1}{2} m v^2 = N_f k T/2$, where m = mass of atom, v = velocity of atom, N_f = degrees of freedom ($N_f = 3$, for perfect gas), k = Boltzmann constant and T = temperature.
[6] This procedure has been developed by the group of Prof. Michael Bowers in Santa Barbara, California, USA.

(femto second $= 10^{-15}$ s), however, the structures of the molecular dynamics simulations will be extracted only every 1000 fs (i.e. 1 ps).
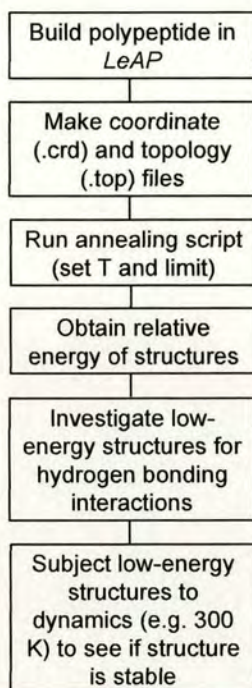

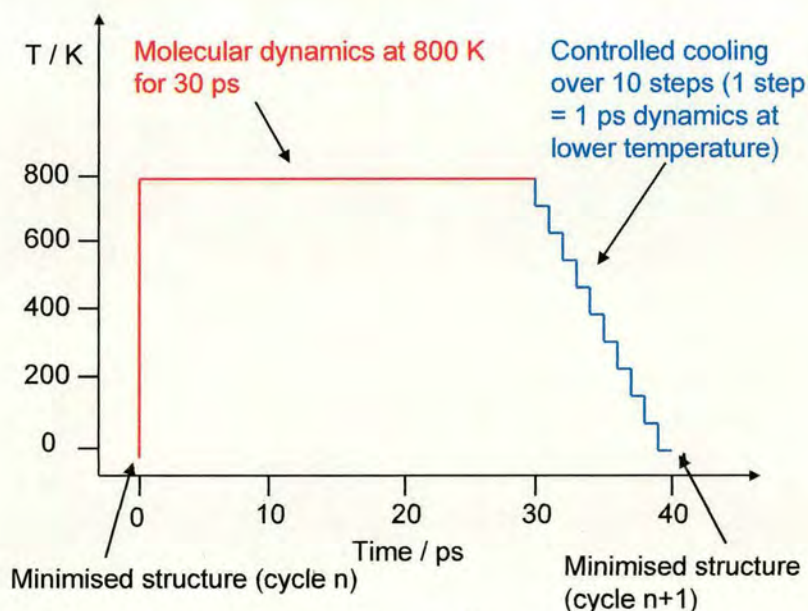**Figure 4.2** Schematic flow chart showing procedure for analysis of gas-phase structures.

Figure 4.3 Schematic for simulated annealing procedure (cycle n+1) showing heating of the minimised structure from the previous cycle (cycle n) to 800 K and dynamics at 800 K for 30 ps followed by cooling in 10 steps to 0 K to give new minimised structure.

4.3 Modelling of [LHRH_HR]$^{2+}$

4.3.1 Minimised structures

The peptide [LHRH_HR]$^{2+}$ (sequence: pEHWSYGLRPG-OH) was built by protonating the side-chains of histidine and arginine. This peptide was chosen due to the relative unambiguous assignment of the protonation sites. Figure 4.4 shows the relative energies of the 100 minimised candidate structures from the annealing procedure which have been normalised with respect to the strain energy of the lowest energy structure E_S[LHRH_HR]$^{2+}$\{min73\} = -128.06 kcal mol^{-1}.
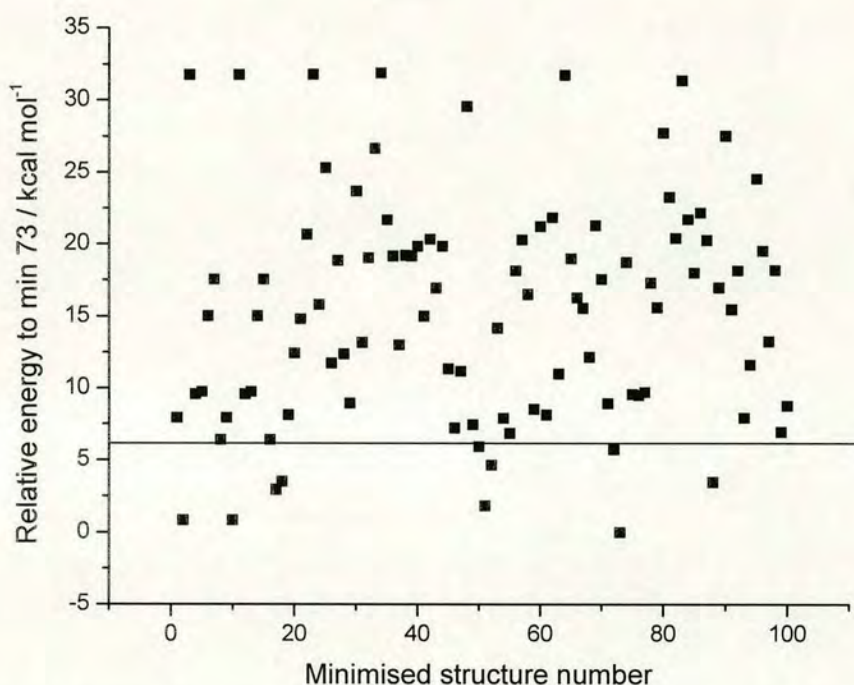
Figure 4.4 Relative energies of minimised structures for [LHRH_HR]$^{2+}$ obtained from the simulated annealing procedure, normalised with respect to E_S[LHRH_HR]$^{2+}$\{min73\} = -128.06 kcal mol^{-1}. The horizontal line indicates the cut-off for the 10 lowest energy structures.

An analysis of the hydrogen bonding interactions using the hydrogen bonding definition as used by Kurz et al. [20] (i.e. bond distance N...O < 3.5 Å and bond angle N-H...O > 120°) showed hydrogen bonding for [\{2\}(His)N-H...O=C(His)\{2\}]

for 50 of the 100 minimised structures. Hydrogen bonding of the histidine side-chain to other backbone carbonyls was found to be much reduced (i.e. ~ 20 candidate minimised structures for each backbone carbonyl). Since the hydrogen atom is believed to be desorbed from the protonated site as a hot hydrogen H$^\bullet$ (~ 6 eV) [22] it is reasonable to argue that the angle for N-H...O should be more linear and hence the cut-off angle for a hydrogen bonding interaction was chosen to be a more conservative N-H...O > 150°. Moreover, the bond distance N...O was chosen as N...O < 3 Å. Calculated equilibrium hydrogen bonding distances for bond distance N...O by Kim *et al.* using quantum chemical calculations were ~ 3 Å [23]. Quantum chemical calculations by Huang *et al.* gave bond distances for N...O < 3 Å and angle N-H...O ~ 180° for the minimum interaction geometries of hydrogen bonded biological complexes [24]. As noted by McDonald *et al.* the hydrogen bond populations predicted by modelling vary with definition stringency for a hydrogen bonding interaction [21]. Nevertheless, constraints for hydrogen bonding interactions must be established in order to analyse the modelled structures even if this approach is somewhat arbitrary.

Table 4.1 shows which of the 100 minimised structures for [LHRH_HR]$^{2+}$ have hydrogen bonding interactions (length < 3 Å, angle > 150°) with particular backbone carbonyls. The number of minimised structures showing hydrogen bonding of the histidine side-chain N-H's to a particular backbone carbonyl is summarised in Figure 4.5. No structures are found showing hydrogen bonding for [{2}(His)N-H...O=C(pGlu){1}] and only 3 structures are found to show hydrogen bonding for [{2}(His)N-H...O=C(His){2}]. This is in agreement with the relative ECD fragmentation pattern for [LHRH_HR]$^{2+}$ (Chapter 3, Figure 3.35), which shows low abundances for z fragments associated with the backbone carbonyls of pGlu{1} and His{2}. However, the relative distribution for the number of minimised structures showing hydrogen bonding to any of the other backbone carbonyls does not match the relative abundances of ECD fragments linked with particular backbone carbonyls (see Chapter 3, Figure 3.35). This shows that the statistical distribution of minimised structures predicting hydrogen bonding to particular backbone carbonyls for this annealing procedure does not correlate with the statistical distribution of ECD fragments linked to particular backbone carbonyls in Figure 3.35. Conversely, the

modelling studies by Guarnieri *et al.* involved a Monte Carlo-type approach [8], which allows for a statistical assessment of gas-phase structures, which would be more appropriate for this study.

Backbone carbonyl	Structure number	Relative energy to min73 / kcal mol⁻¹
pGlu{1}	none	NA
His{2}	min21	14.79
His{2}	min72	5.71
His{2}	min91	15.43
His{2}	**Average:**	**11.98 ± 5.44**
Trp{3}	min15	17.53
Trp{3}	min16	6.39
Trp{3}	min27	18.85
Trp{3}	min28	12.37
Trp{3}	min29	8.92
Trp{3}	min35	21.66
Trp{3}	min45	11.32
Trp{3}	min47	11.16
Trp{3}	min48	29.45
Trp{3}	min56	18.13
Trp{3}	min66	16.25
Trp{3}	min69	21.27
Trp{3}	min73	0
Trp{3}	min85	17.97
Trp{3}	min94	11.60
Trp{3}	min100	8.74
Trp{3}	**Average:**	**14.88 ± 7.05**
Ser{4}	min10	0.84
Ser{4}	min32	19.03
Ser{4}	min49	7.43
Ser{4}	min60	21.20
Ser{4}	min74	18.71
Ser{4}	min93	7.90
Ser{4}	min95	24.52
Ser{4}	min96	19.50
Ser{4}	min98	18.16
Ser{4}	min99	6.93
Ser{4}	**Average:**	**14.42 ± 7.88**
Tyr{5}	min4	9.58
Tyr{5}	min6	14.98
Tyr{5}	min7	17.53
Tyr{5}	min30	23.67
Tyr{5}	min31	13.16
Tyr{5}	min36	19.15
Tyr{5}	min43	16.91

Tyr{5}	min78	17.27
Tyr{5}	min79	15.56
Tyr{5}	min89	16.95
Tyr{5}	**Average:**	**16.48 ± 3.69**
Gly{6}	min12	9.58
Gly{6}	min14	14.98
Gly{6}	min29	8.92
Gly{6}	min40	19.82
Gly{6}	min57	20.26
Gly{6}	min64	31.69
Gly{6}	min75	9.55
Gly{6}	min81	23.24
Gly{6}	min84	21.69
Gly{6}	**Average:**	**17.75 ± 7.67**
Leu{7}	min1	7.92
Leu{7}	min2	0.84
Leu{7}	min9	7.92
Leu{7}	min15	17.53
Leu{7}	min16	6.39
Leu{7}	min20	12.42
Leu{7}	min37	12.99
Leu{7}	min44	19.82
Leu{7}	min45	11.32
Leu{7}	min46	7.20
Leu{7}	min47	11.16
Leu{7}	min55	6.84
Leu{7}	min65	18.96
Leu{7}	min67	15.52
Leu{7}	min69	21.27
Leu{7}	min70	17.53
Leu{7}	min80	27.20
Leu{7}	min85	17.97
Leu{7}	min88	3.44
Leu{7}	**Average:**	**12.85 ± 6.82**

Table 4.1 Table summarising hydrogen bonding interactions of histidine side-chain with backbone carbonyls for 100 minimised structures of $[LHRH\_HR]^{2+}$. The average energy of the structures that show hydrogen bonding to a particular backbone carbonyl is shown, as well as the standard deviation.
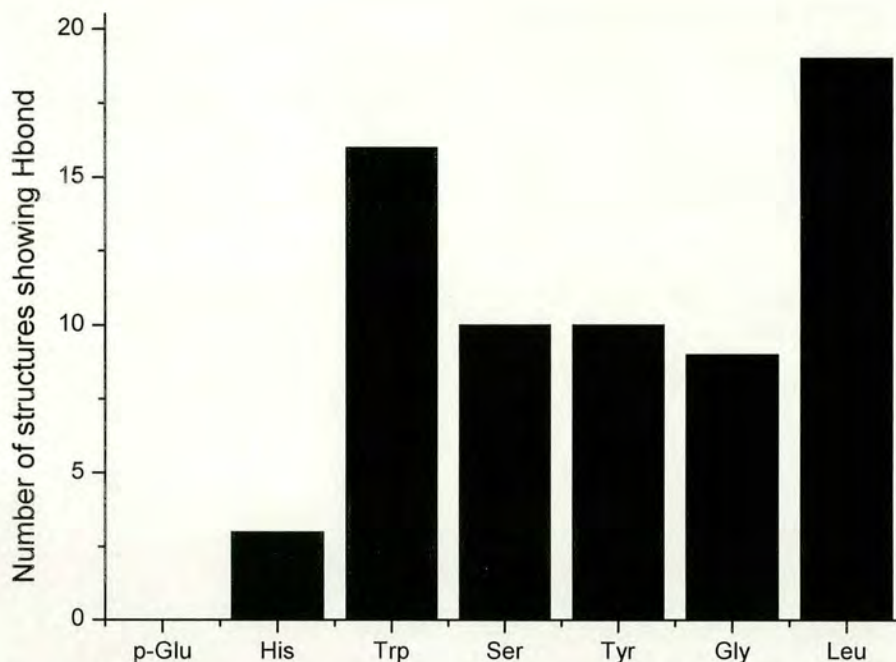
Figure 4.5 Number of minimised structures for [LHRH_HR]$^{2+}$ showing hydrogen bonding (Hbond) to particular backbone carbonyls.

An alternative way to look at the modelled results is to consider the lowest energy structures only, since the ions may exist in low-energy conformations. Figure 4.6 shows the 10 lowest energy structures. For the sake of clarity only the side-chains of histidine and one other amino acid are shown. The backbone of the peptide structures is represented as a ribbon, with the N-terminus on the left-hand-side and the C-terminus on the right-hand-side. 5 of the minimised structures (min2, min10, min17, min52 and min73, Figure 4.5 A, B, C, G and I) can be classified as 'loop' structures involving a β-turn in the middle of the molecule and both termini close in space. Other minimised structures have backbones that vary more significantly, despite an energy difference of merely < 6 kcal mol^{-1}. This compares to calculated energy stabilisation of biologically relevant hydrogen bonds by *ab initio* calculations in the range of 5-30 kcal mol^{-1} [24].
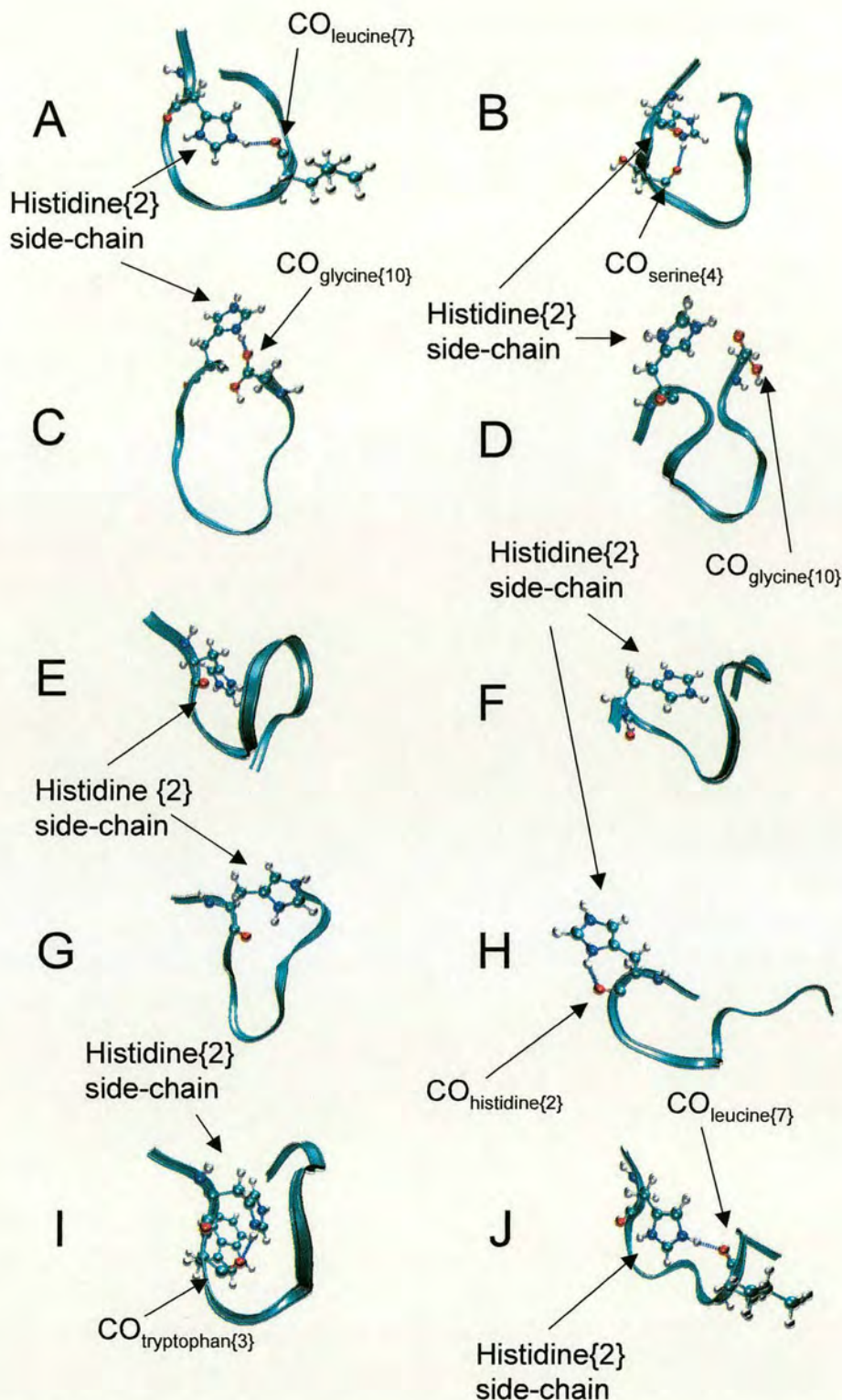
Figure 4.6 10 lowest energy structures for simulated annealing of [LHRH_HR]$^{2+}$, showing backbone ribbon, the side-chain of histidine and the backbone carbonyl that histidine nitrogen-bound hydrogens are in hydrogen bonding contact with (indicated by blue lines). (A) min2, (B) min10, (C) min17, (D) min18, (E) min50, (F) min51, (G) min52, (H) min72, (I) min73 and (J) min88. CO$_{Xxx\{X\}}$ refers to backbone carbonyl of amino acid residue Xxx in position {X} from N-terminus.

An analysis of the 10 lowest energy structures reveals that six of them show hydrogen bonding of one of the nitrogen-bound hydrogens (N-H's) on the histidine side-chain to backbone carbonyls (see Figure 4.6). Structure min17 (Figure 4.6 C) shows a hydrogen bonding interaction of the histidine side-chain to the C-terminal backbone CO (carbonyl). The other five show hydrogen bonding of the histidine side-chain to the backbone carbonyls of histidine{2} (Figure 4.6 H), tryptophan{3} (Figure 4.6 I), serine{4} (Figure 4.6 B) and leucine{7} (Figure 4.6 A and J).

The three lowest energy structures predicted from this annealing procedure are min2 (Figure 4.6 A), min10 (Figure 4.6 B) and min73 (Figure 4.6 I), which show hydrogen bonding to the backbone carbonyls of leucine{7}, serine{4} and tryptophan{3}. In the study of the relative z fragment abundances for $[LHRH\_HR]^{2+}$ (see Figure 3.35, Chapter 3), the fragmentation associated with the backbone carbonyls of tryptophan{3}, serine{4}, glycine{6} and leucine{7} were found to be enhanced. Hence, the hydrogen bonding pattern of the histidine side-chain to backbone carbonyls as predicted by the <u>lowest</u> energy minimised structures has confirmed 3 of the 4 enhanced ECD cleavage sites in Figure 3.35 (Chapter 3).

Figure 4.7 shows the detail of the lowest energy structure min73, where the side-chain of histidine is in hydrogen bonding contact with the backbone carbonyl of tryptophan $[\{2\}(His)N-H...O=C(Trp)\{3\}]^7$, resulting in enhanced hydrogen atom transfer to that carbonyl followed by higher backbone cleavage. *Ab initio* calculations by Tureček and co-workers also showed that the activation barrier for such a one-step transfer of the hydrogen atom to a backbone carbonyl is very low and would be preferred over the two-step process (hydrogen atom desorption followed by re-capture) [25]. The predicted rate constants for this one-step process were as high as 10^{12} s^{-1}, which approaches *non-ergodic* behaviour. It should be noted that no hydrogen bonding was found in these modelling studies to amide nitrogens and it is well known that hydrogen bonding to oxygen is stronger than to nitrogen due to the four lone pair electrons on oxygen.

---

[7] Note that $[\{X\}(Xxx)N-H...O=C(Yyy)\{Y\}]$ refers to the atomic distance between the nitrogen-bound hydrogen on the side-chain of amino acid residue Xxx in position {X} (from N-terminus) to the oxygen atom on the backbone carbonyl belonging to the amino acid residue Yyy in position {Y} (from N-terminus).
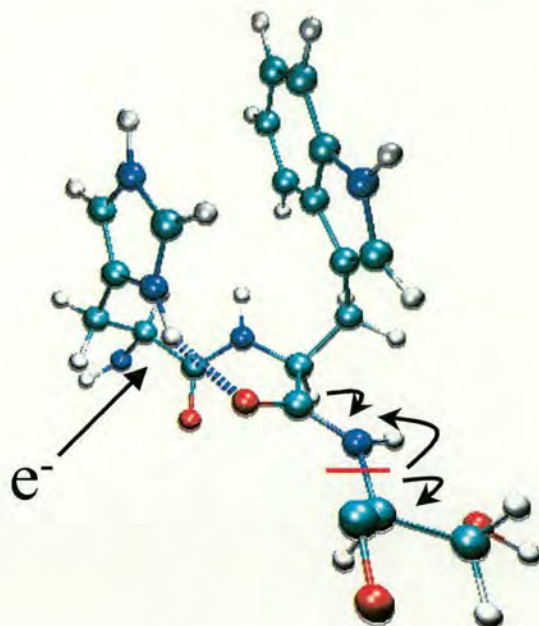
Figure 4.7 Detail of minimised structure min73, showing hydrogen bond between histidine N-H to backbone carbonyl of tryptophan [{2}(His)N-H...O=C(Trp){3}]. Schematic exemplifies that electron capture at histidine side-chain results in enhanced hydrogen atom transfer to the backbone carbonyl of tryptophan and therefore higher cleavage at that site (indicated by red line). (The 'normal' α-cleavage is shown, see Chapter 3 Figure 3.13. Note that arrows refer to single electron movement.)

4.3.2 Molecular dynamics of min73

Molecular dynamics of the lowest energy structure (min73) at 300 K shows that the hydrogen bonding interaction for [{2}(His)N-H...O=C(Trp){3}] is conserved at higher internal energies (see Figure 4.8).[8] Despite an initial increase in the interatomic distance, the structure quickly (< 30 ps) stabilises to within a hydrogen bonding interaction distance for [{2}(His)N-H...O=C(Trp){3}]. Molecular dynamics of the structure 300K.1000ps[9] at 500 K shows that both hydrogens of the histidine side-chain are alternately in hydrogen bonding contact, suggesting a flipping of the histidine side-chain (see Figure 4.9). The higher temperature of the

---

[8] Note that the SHAKE algorithm [26] is used in these simulations. This restricts the movement of the hydrogen atoms.

[9] For the sake of simplicity the gas-phase structures of dynamic simulations will be referred to as XK.Yps, where X is the temperature of the simulation in Kelvin and Y is the number of pico seconds of simulation. Hence 300K.1000ps refers to the structure obtained after 1000 ps molecular dynamics at 300 K.

molecule represents a higher velocity of the atoms, manifesting itself in side-chain flipping. The higher temperature can also be understood as a speeding up of the simulation and hence 1000 ps at 300 K is equivalent to a time frame of < 1000 ps simulation at 500 K.
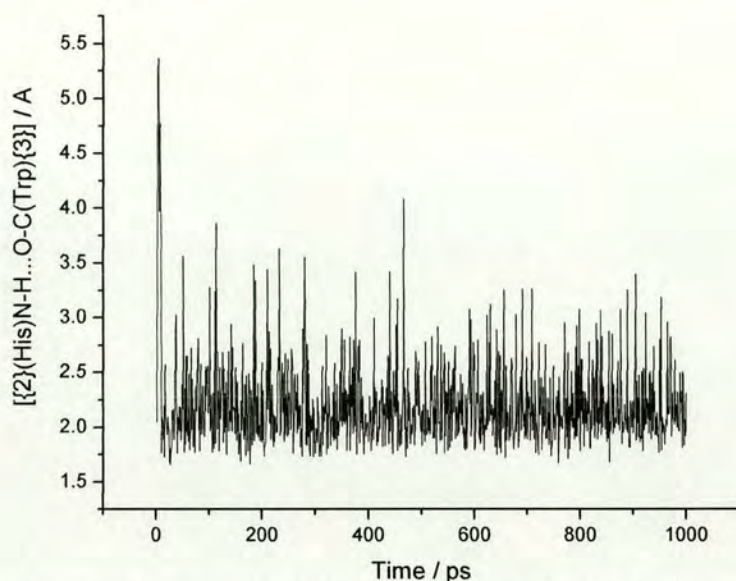
Figure 4.8 Atomic distance [{2}(His)N-H…O=C(Trp){3}] (in Ångstrom) for molecular dynamics of [LHRH_HR]$^{2+}$ min73 at 300 K.
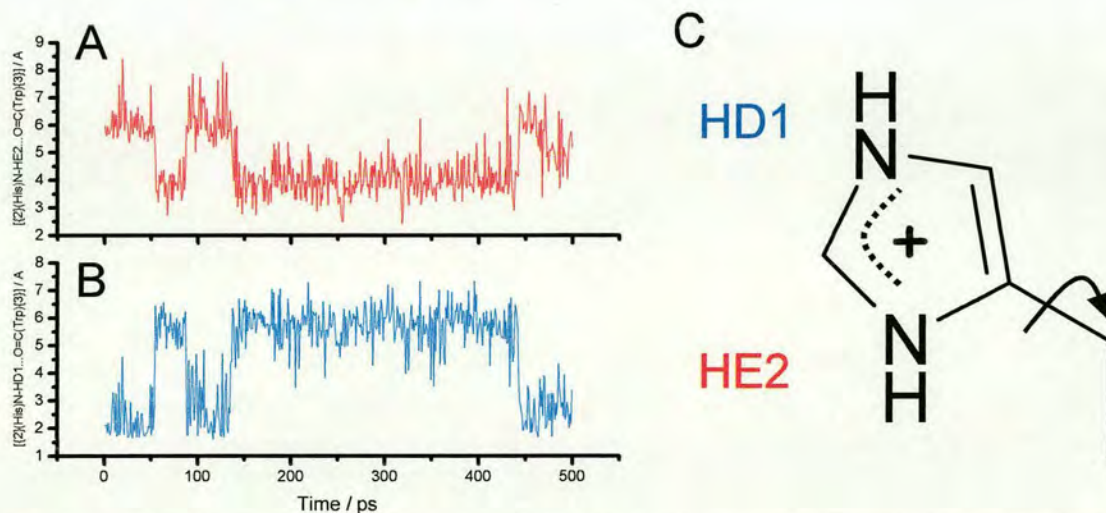
Figure 4.9 (A) Atomic distances [{2}(His)N-HD1…O=C(Trp){3}] and (B) [{2}(His)N-HE2…O=C(Trp){3}] for molecular dynamics of [LHRH_HR]$^{2+}$ min73 at 500 K. (C) Structure (stick structure) of protonated histidine side-chain showing delocalisation of positive charge and two different nitrogen-bound hydrogens (HE2 and HD1) on histidine side-chain. Curvy arrow indicates rotation of side-chain.

154

Figure 4.10 shows the root-mean-square (RMS) deviation of the 500 K molecular dynamics simulation. In Figure 4.10 RMS refers to the average spatial deviation of each heavy backbone atom (i.e. N, C_α, C and O) for 500K.Yps (where Y is number of ps of simulation) compared to the reference structure 500K.1ps. This structure appears to be stable and hence constitutes a likely gas-phase structure of this molecule.
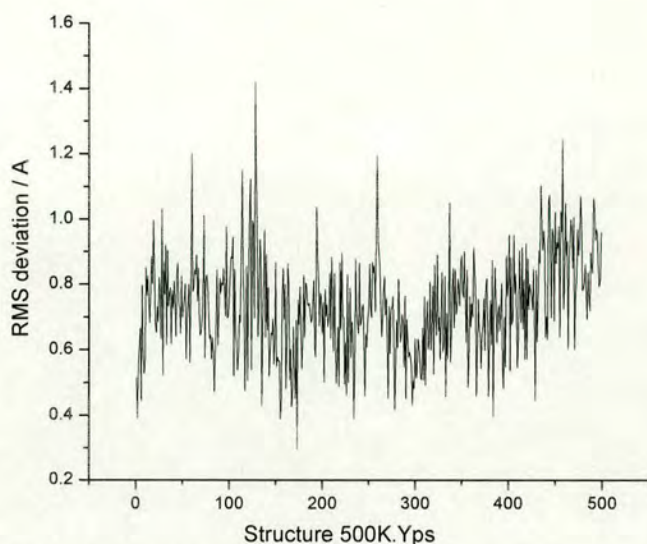
Figure 4.10 RMS deviation of peptide backbone (atoms N, C_α, C and O) for min 73 of [LHRH_HR]$^{2+}$ at dynamics of 500 K for reference = 500K.1ps.

Figure 4.11 shows the gas-phase structure of min73 at 500 K after 50 ps of simulation (500K.50ps). The β-turn centred at residues 4-6 can clearly be seen. Given the high internal energy of the molecule, the similarity of this structure to the minimised structure of min73 (Figure 4.6 I) implies a significant thermodynamic stability of the β-turn.
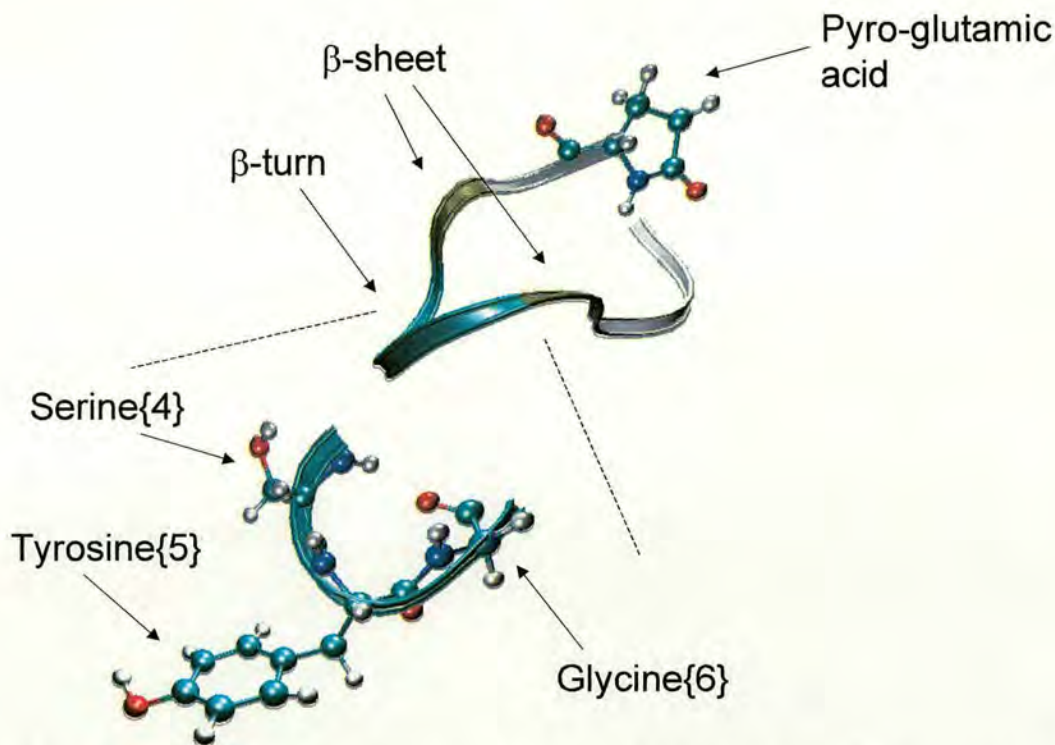
Figure 4.11 Ribbon gas-phase structure of min73 of [LHRH_HR]$^{2+}$ at dynamics of 500 K after 50 ps simulation (500K.50ps) showing pyro-glutamic acid N-terminus and secondary structure (β-sheet – brown, β-turn – cyan). Insert shows detail of β-turn for residues 4-6 (serine, tyrosine and glycine).

The interaction of the arginine side-chain N-H's with the backbone carbonyls of [LHRH_HR]$^{2+}$ for the 300 K dynamics run is shown in Figure 4.12. Three of the 5 nitrogen-bound hydrogens (N-H's) are in hydrogen bonding contact to the backbone carbonyls of leucine and tyrosine. Figure 4.13 shows a representative ribbon backbone structure for this simulation and the hydrogen bonding interactions for the arginine side-chain. If preferred hydrogen bonding of the charged site hydrogens to particular backbone carbonyls results in enhanced backbone cleavage next to those carbonyls, one would expect c_5^+ and c_7^+ fragments. As discussed in Chapter 3, in ECD of [LHRH_HR]$^{2+}$ no such fragments are observed. In terms of the *hot hydrogen re-arrangement mechanism* these modelling results suggest that the poor transfer of hydrogen atoms from the arginine side-chain to the backbone (resulting in backbone cleavage) is not due to an inability of the arginine side-chain to solvate backbone carbonyls (see discussion on 'ease of hydrogen atom transfer from protonated site', section 3.3, Chapter 3).
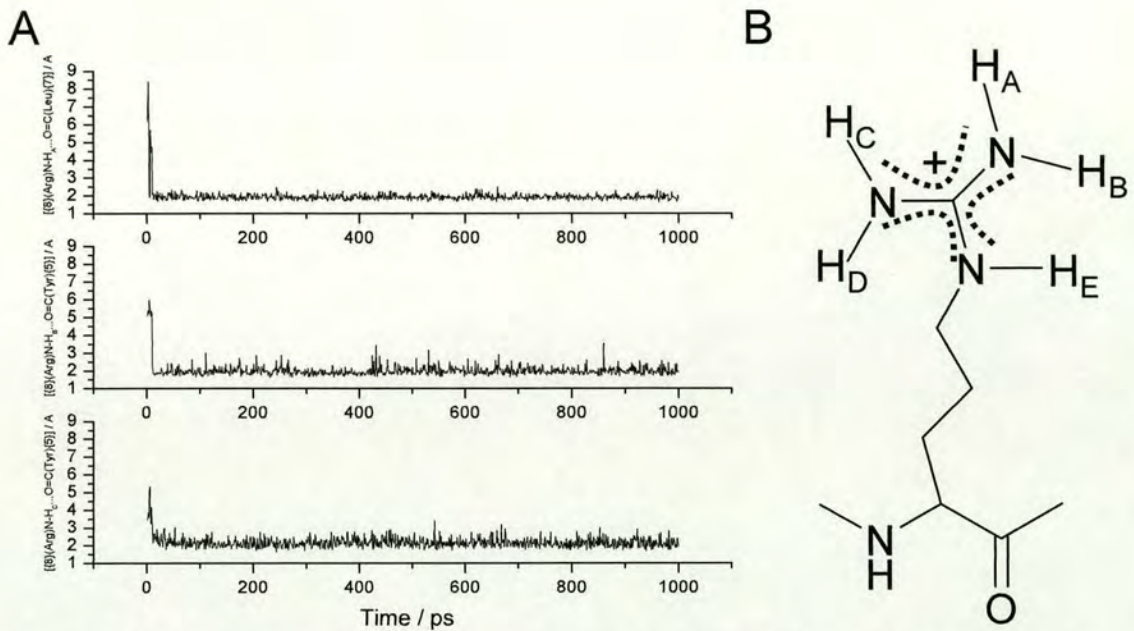
Figure 4.12 (A) Atomic distances of arginine side-chain N-H$_A$, N-H$_B$ and N-H$_C$ to backbone carbonyls of tyrosine{5} and leucine{7} for dynamics of min73 at 300 K. (B) Schematic structure (stick structure) of arginine residue showing 5 N-H on side-chain: H$_A$, H$_B$, H$_C$, H$_D$ and H$_E$. Positive charge (+) nominally located on C, but in fact delocalised over guanidine group (-NH-C-(NH$_2$)$_2$).
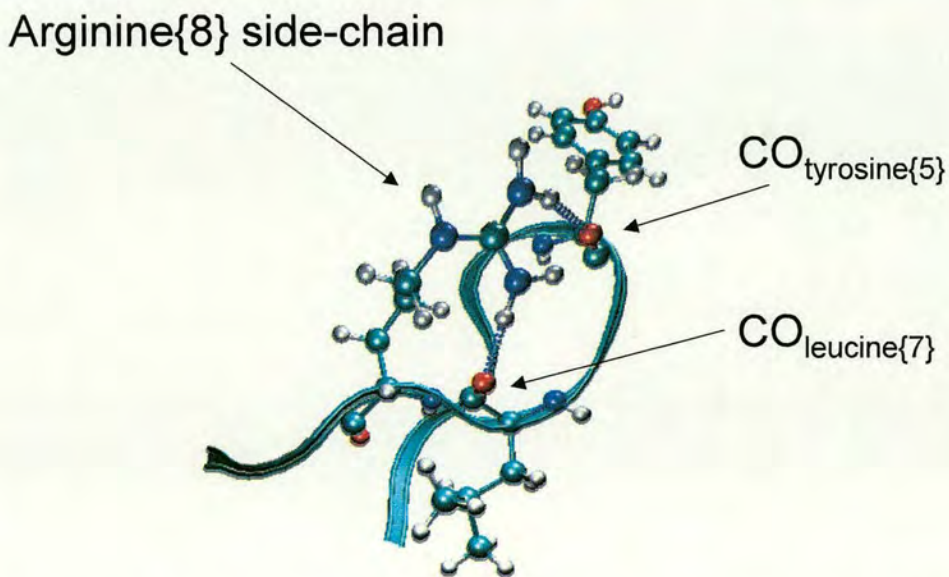
Figure 4.13 Ribbon structure of min73 at molecular dynamics of 300 K showing hydrogen bonding of N-H's to backbone carbonyls of tyrosine{5} and leucine{7}. Structure 300K.100ps.

157

4.3.4 Molecular dynamics of min10

Molecular dynamics at 100 K of the minimised structure $[LHRH\_HR]^{2+}$ min10 maintains $[\{2\}(His)N-HD1...O=C(Ser)\{4\}]$ in a hydrogen bonding interaction (Figure 4.14 A). However, at higher internal energies of 300 K (Figure 4.14 B), this hydrogen bond is broken. Figure 4.15 shows the RMS deviations for the molecular dynamics simulation at 300 K, comparing the following structures in the dynamics simulation to the initial structure or the structure obtained after 30 ps. The peptide backbone is shown to readily adopt a β-turn conformation around residues 4-6, similarly to min73 (see Figure 4.11). This structure can be described as a 'loop' structure with a β-turn involving residues 4-6 and both N and C-termini close together. Nevertheless, neither of the two histidine side-chain N-H's are found to be in hydrogen bonding to backbone carbonyls.
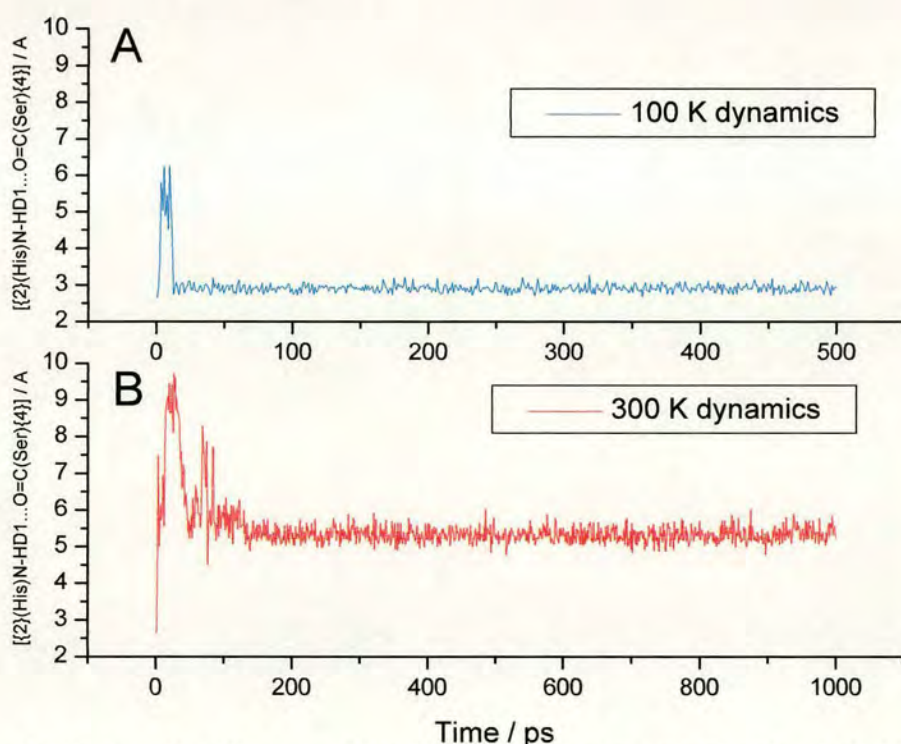
Figure 4.14 Atomic distance $[\{2\}(His)N-H...O=C(Ser)\{4\}]$ for $[LHRH\_HR]^{2+}$ min10 at molecular dynamics of (A) 100 K, and (B) 300 K.
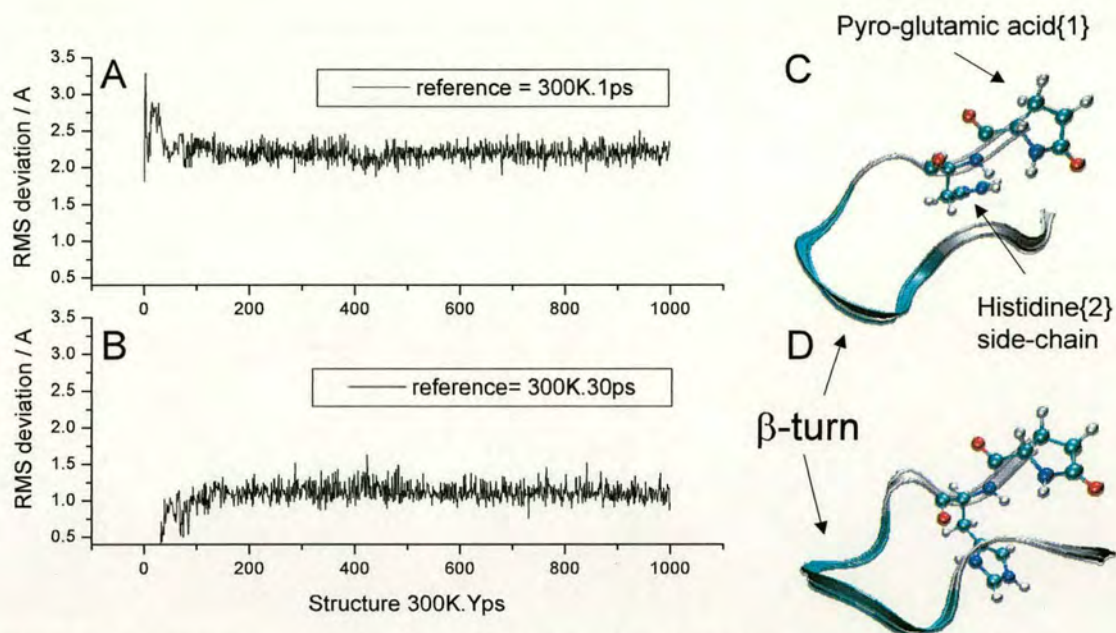
Figure 4.15 RMS deviation for [LHRH_HR]$^{2+}$ min10 (A) structures 300K.2ps – 300K.1000ps in 300 K dynamics simulation compared to reference 300K.1ps (B) structures 300K.31ps – 300K.1000ps in 300 K dynamics simulation compared to reference 300K.30ps (C) ribbon structure for 300K.30ps (D) ribbon structure for 300K.500ps. β-turn is shown for both structures.

4.3.5 Molecular dynamics of min2

Molecular dynamics at 300 K of [LHRH_HR]$^{2+}$ min2 shows that this structure falls apart to an elongated structure (Figure 4.16) and that the hydrogen bond for [{2}(His)N-HE2...O=C(Leu){7}] is broken (Figure 4.17 A). Instead, N-HE2 of the histidine side-chain is found to be in hydrogen bonding contact with the backbone carbonyl of tyrosine. This structure is fairly dynamic and hence the hydrogen bonding interaction of the atomic distance [{2}(His)N-HE2...O=C(Tyr){5}] is broken at times (Figure 4.17 B). In the relative ECD fragmentation study of [LHRH_HR]$^{2+}$ (see Chapter 3, Figure 3.35) the z fragmentation associated with Tyr{5} is a lower fragmentation channel.
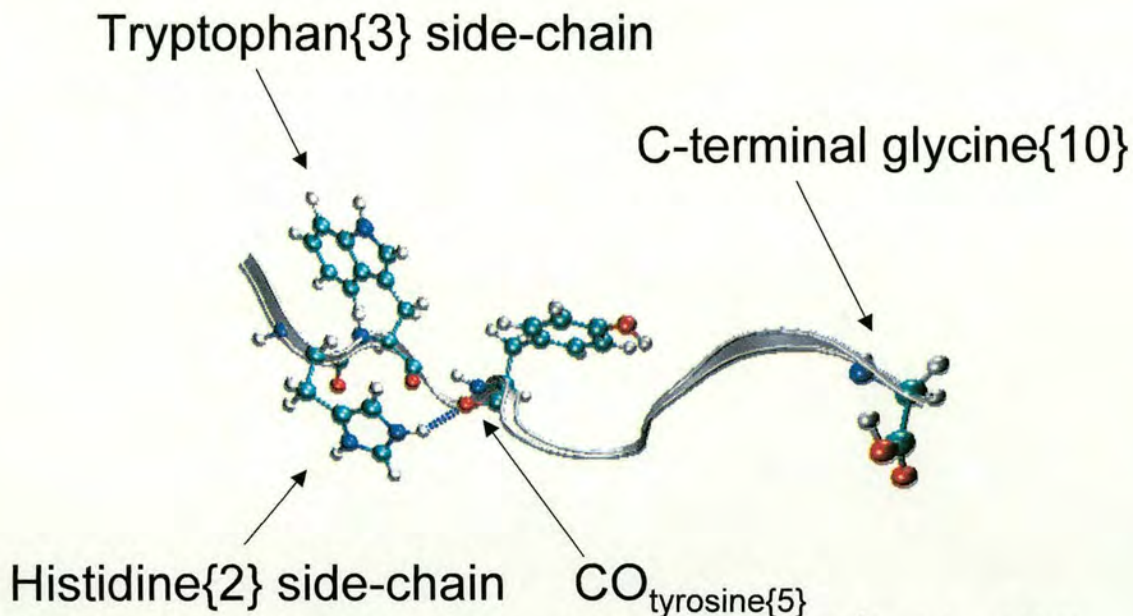
Figure 4.16 Molecular dynamics ribbon structure of [LHRH_HR]$^{2+}$ min2 at 300 K. Structure 300K.200ps.
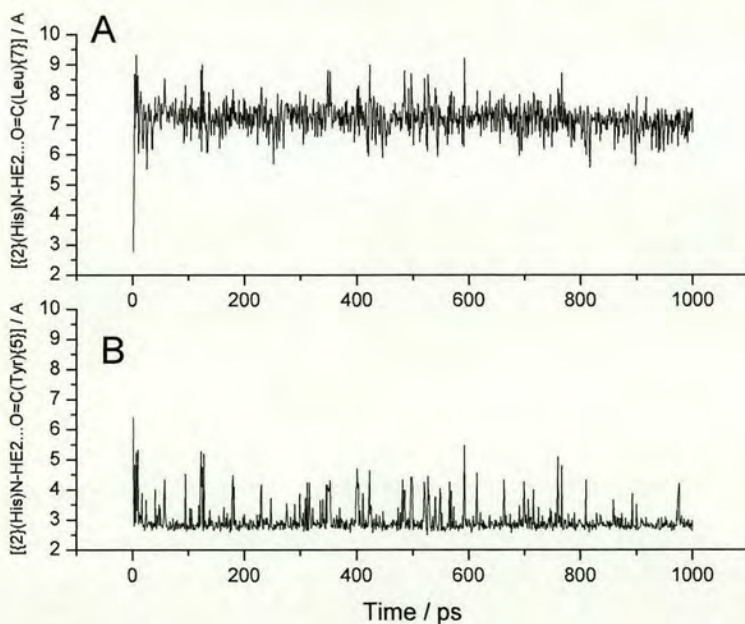
Figure 4.17 Atomic distances for dynamics of [LHRH_HR]$^{2+}$ min2 at 300 K (A) [{2}(His)N-HE2...O=C(Leu){7}] and (B) [{2}(His)N-HE2...O=C(Tyr){5}].

4.3.6 Conclusions for [LHRH_HR]$^{2+}$

In summary, an analysis of the three lowest energy structures of [LHRH_HR]$^{2+}$ obtained by simulated annealing has shown that hydrogen bonding of the histidine side-chain N-H's to backbone carbonyls correlates with high ECD backbone fragmentation associated with those carbonyls. Hydrogen bonding to the backbone carbonyl of glycine{6} was not predicted from the lowest energy structures. Molecular dynamics at low energy (100 K) appears to conserve the hydrogen bonding interaction (Figure 4.14 A), however higher energy (300 K) dynamics changes the gas-phase conformation for min10 and min2 (Figure 4.14 B and 4.17 A). This could be explained by the fact that the ions in the ECD experiments of [LHRH_HR]$^{2+}$ have low internal energies. The 'loop' structure involving a β-turn conformation at residues 4-6 was independently found for different minimised structures and found to be stable at higher internal energies (with the exception of min 2, see Figure 4.16).

4.4 Modelling of C-terminal amidated LHRH

4.4.1 Modelling of $[LHRH\_HR\_NH_2]^{2+}$

4.4.1.1 Minimised structures

In order to assess the comparability of the simulated annealing procedure, the peptide $[LHRH\_HR\_NH_2]^{2+}$ (mGnRH, sequence: pEHWSYGLRPG-NH$_2$, protonation sites: side-chains of arginine and histidine) was modelled using the same approach as for $[LHRH\_HR]^{2+}$. Both peptides differ merely by their C-termini and hence the annealing procedure is expected to yield low-energy structures that are similar.
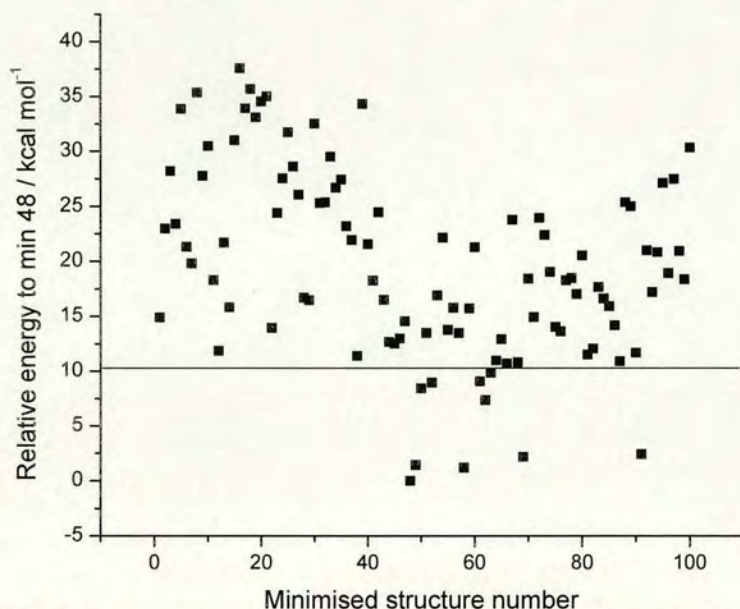
Figure 4.18 Relative energies of 100 minimised structures obtained by simulated annealing of $[LHRH\_HR\_NH_2]^{2+}$, normalised with respect to $E_S[LHRH\_HR\_NH_2]^{2+}\{min48\} = -184.17\ kcal\ mol^{-1}$.

Figure 4.18 shows the relative energies of the 100 minimised candidate structures. The energy spread from the lowest (min48) to the highest energy structure (min16) is 37.5 kcal mol^{-1}, which is comparable to the simulated annealing for

162

[LHRH_HR]$^{2+}$ (difference = 31.8 kcal mol^{-1}, see Figure 4.4). Table 4.2 shows which of the 100 minimised structures for [LHRH_HR_NH$_2$]$^{2+}$ have hydrogen bonding interactions (length $\leq$ 3 Å, angle $\leq$ 30°) with particular backbone carbonyls. The number of minimised structures showing hydrogen bonding of the histidine side-chain N-H's to a particular backbone carbonyl is summarised in Figure 4.19. The relative distribution of structures showing hydrogen bonding in Figure 4.19 does not correlate with the relative ECD fragmentation pattern for [LHRH_HR_NH$_2$]$^{2+}$ (Chapter 3, Figure 3.37).

Backbone carbonyl	Structure number	Relative energy to min48 / kcal mol^{-1}
pGlu{1}	min42	24.44
pGlu{1}	min71	14.94
pGlu{1}	min100	30.30
pGlu{1}	**Average:**	**23.23 ± 7.75**
His{2}	min1	14.90
His{2}	min9	27.76
His{2}	min11	18.28
His{2}	min12	11.86
His{2}	min15	31.01
His{2}	min69	2.17
His{2}	min70	18.41
His{2}	min88	25.33
His{2}	**Average:**	**18.72 ± 9.35**
Trp{3}	min13	21.68
Trp{3}	min16	37.55
Trp{3}	min19	33.08
Trp{3}	min21	35.00
Trp{3}	min32	25.34
Trp{3}	min33	29.49
Trp{3}	min34	26.67
Trp{3}	min40	21.54
Trp{3}	min73	22.35
Trp{3}	min74	19.01
Trp{3}	min92	20.97
Trp{3}	min97	27.45
	Average:	**26.678 ± 6.03**
Ser{4}	min2	22.97
Ser{4}	min5	33.88
Ser{4}	min30	32.52
Ser{4}	min39	34.31
	Average:	**30.92 ± 5.35**
Tyr{5}	min14	15.83
Tyr{5}	min20	34.54

Tyr{5}	min25	31.73
Tyr{5}	min27	26.05
Tyr{5}	min28	16.68
Tyr{5}	min29	16.47
Tyr{5}	min58	1.20
Tyr{5}	min59	15.71
Tyr{5}	min93	17.18
Tyr{5}	min95	27.08
Tyr{5}	**Average:**	**20.25 ± 9.75**
Gly{6}	min36	**26.67**
Leu{7}	min10	30.48
Leu{7}	min54	22.12
Leu{7}	min55	13.76
Leu{7}	min61	9.06
Leu{7}	min62	7.34
Leu{7}	min79	17.02
Leu{7}	**Average:**	**16.63 ± 8.65**

Table 4.2 Table summarising hydrogen bonding interactions of histidine side-chain with particular backbone carbonyls for 100 minimised structures of $[LHRH\_HR\_NH_2]^{2+}$.
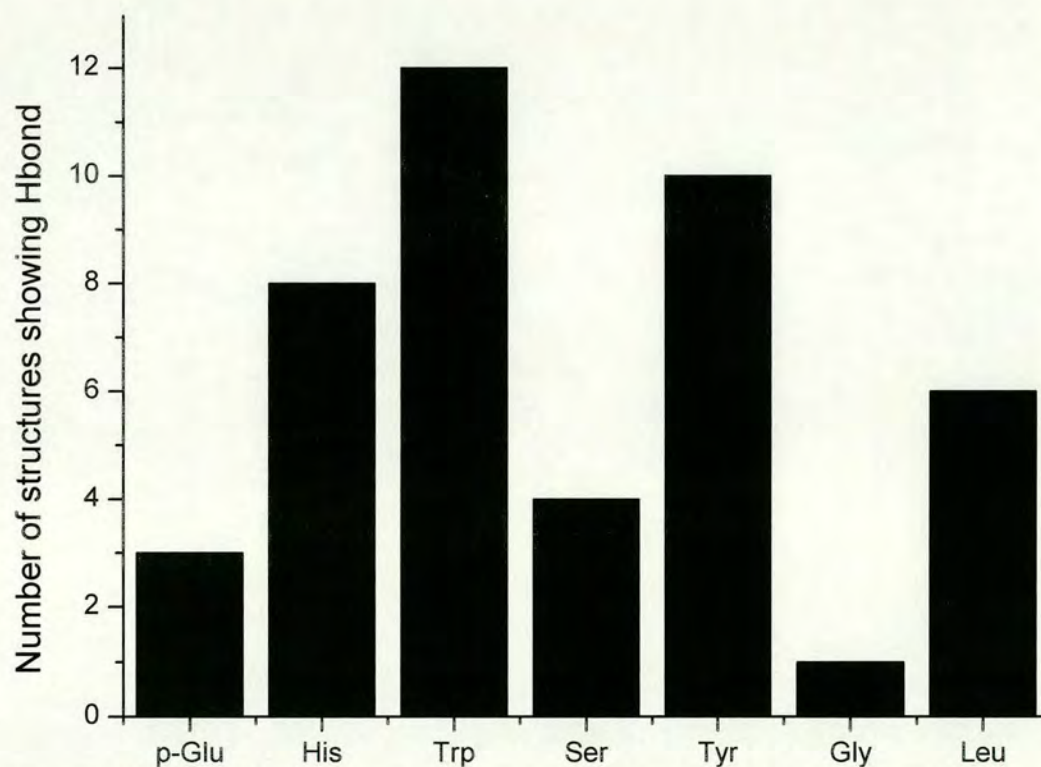
Figure 4.19 Number of minimised structures for $[LHRH\_HR\_NH_2]^{2+}$ confirming hydrogen bonding (Hbond) to particular backbone carbonyls.

164

Figure 4.20 shows the ribbon structures of the 10 lowest energy minimised structures for the simulated annealing of [LHRH _HR_NH$_2$]$^{2+}$. One 'family' of structures includes min48, min49, min50 and min52 (Figure 4.20 A, B, C and D), which all show hydrogen bonding of (His)N-HD1 to the backbone carbonyl of arginine{8}. Two other structures predict hydrogen bonding of (His)N-HD1 to the backbone carbonyl of leucine{7} (min61 and min62, Figure 4.20 F and G), whereas min63 and min69 (Figure 4.20 H and I) predict hydrogen bonding of (His)N-HD1 to the C-terminal glycine{10} carbonyl. This means that none of the low-energy structures for [LHRH_HR_NH$_2$]$^{2+}$ predict hydrogen bonding of the histidine side-chain N-H's to backbone carbonyls with enhanced ECD fragmentation (see Chapter 3, Figure 3.37).
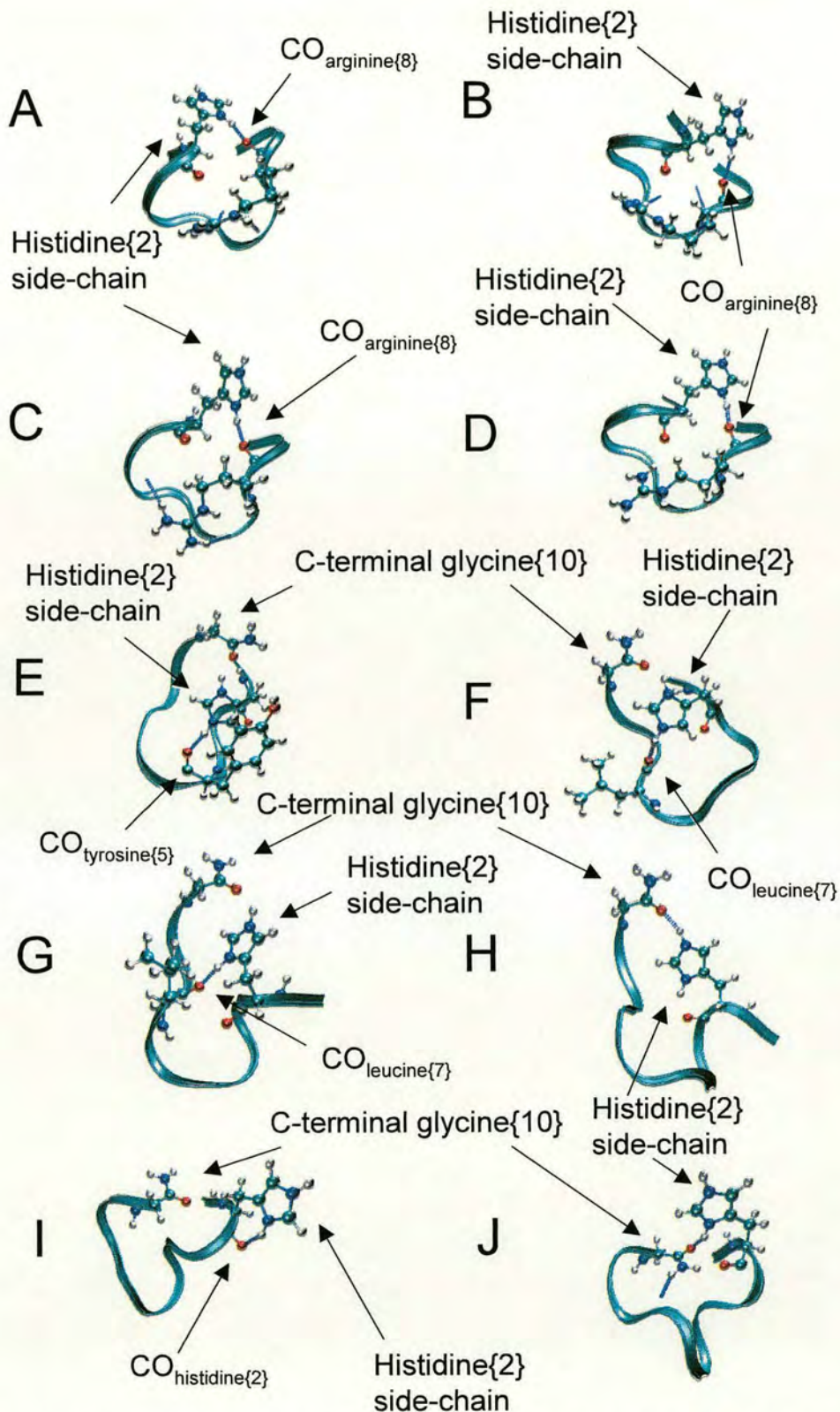
Figure 4.20 Low-energy ribbon structures for annealing of [LHRH _HR_NH$_2$]$^{2+}$ showing histidine side-chain and hydrogen bonding interaction with backbone carbonyl if present (A) min48, (B) min49, (C) min50, (D) min52, (E) min58, (F) min61, (G) min62, (H) min63, (I) min69 and (J) min91.

4.4.1.2 Molecular dynamics of min48

The lowest energy structure min48 was subjected to molecular dynamics at 300 K. Figure 4.21 shows the atomic distances of N-HD1 on the histidine side-chain to the backbone carbonyls of arginine{8}, proline{9} and leucine{10}, all of which are in hydrogen bonding distance. Figure 4.22 shows the ribbon backbone structure of the peptide after 500 ps dynamics at 300 K (300K.500ps). Once again, the N-H's on the side-chain of arginine can be seen to be in hydrogen bonding with the backbone structure.
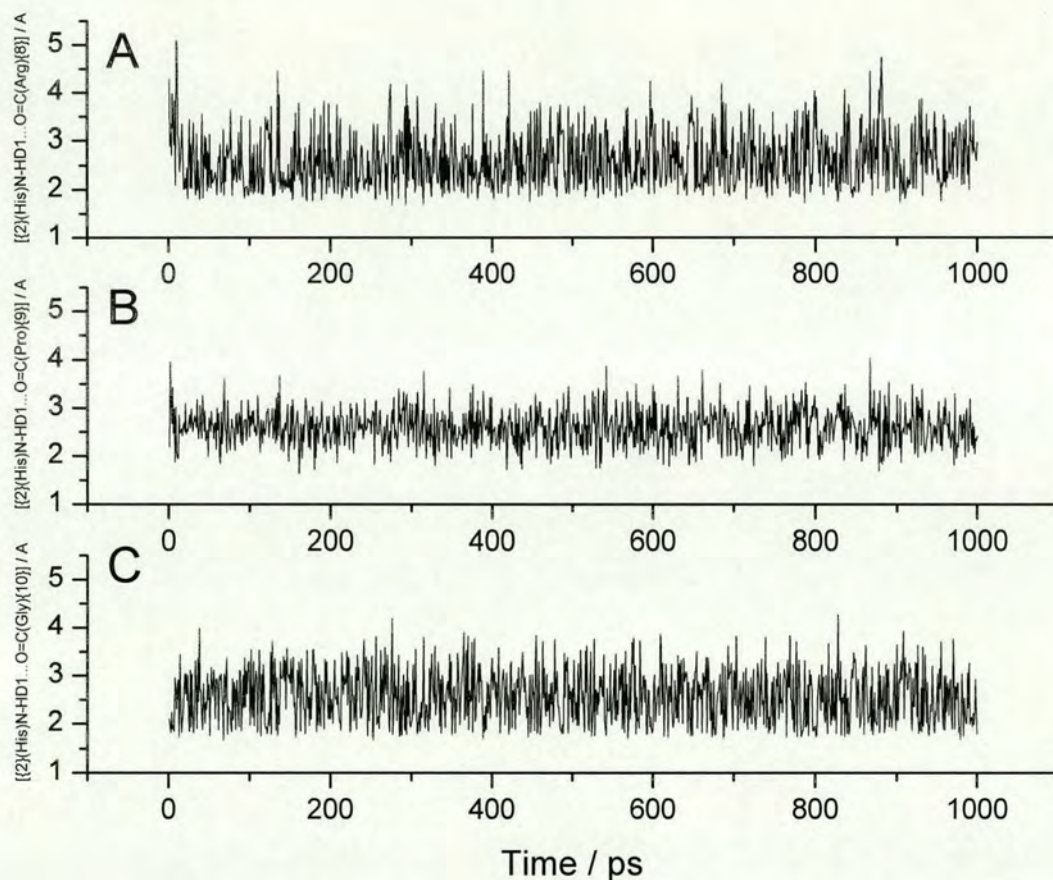
Figure 4.21 Atomic distances (in Ångstrom) for molecular dynamics at 300 K of [LHRH_NH$_2$_HR]$^{2+}$ min48 (A) [{2}(His)N-HD1...O=C(Arg){8}], (B) [{2}(His)N-HD1...O=C(Pro){9}] and (C) [{2}(His)N-HD1...O=C(Gly){10}].
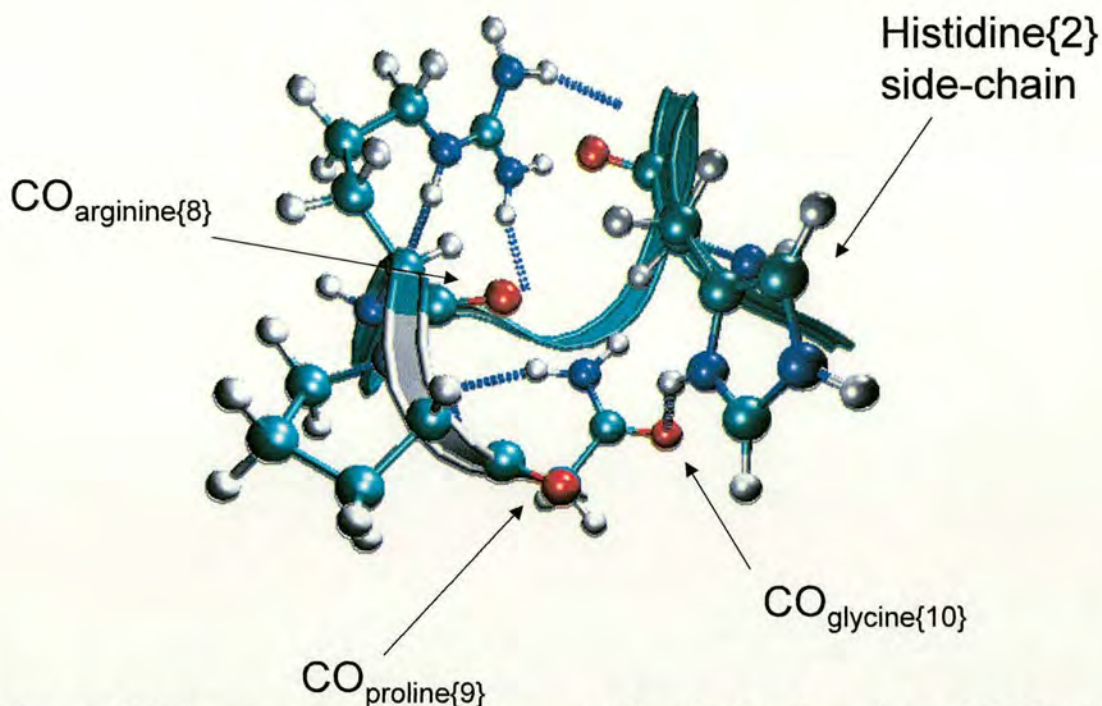
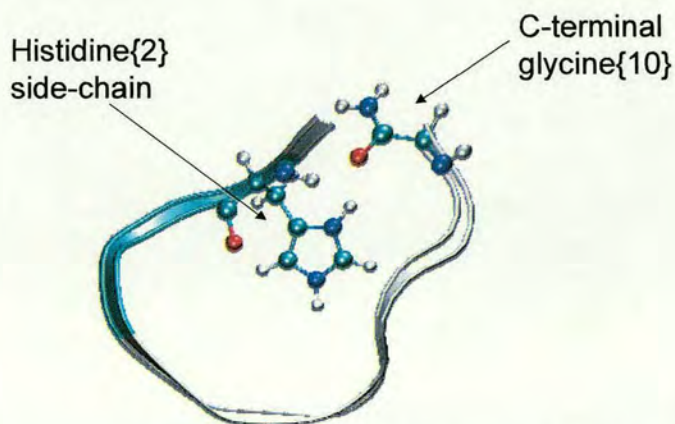**Figure 4.22** Representative ribbon structure of molecular dynamics at 300 K of [LHRH _HR_NH$_2$]$^{2+}$ min48. Structure 300K.500ps.

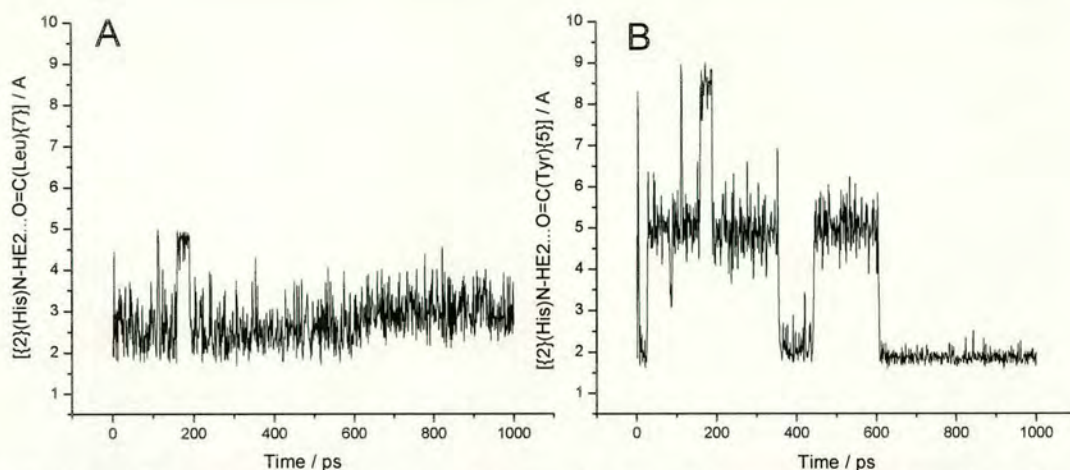### 4.4.1.3 Molecular dynamics of min61

The minimised structure min61 (Figure 4.20 F) was also subjected to dynamics at 300 K. The atomic distances for (His)N-HD1 to various backbone carbonyls are shown in Figure 4.23. A representative structure of the dynamics simulation is shown in Figure 4.24. Enhanced hydrogen atom transfer from the histidine side-chain to the backbone carbonyls of arginine could explain the high abundance of the $z_3^{+\prime}$ fragment in the ECD mass spectra of all LHRH compounds. The presence of the proline residue in position 9 prevents the corresponding normal alpha-cleavage $z_2^{+\bullet}$ respectively c_8^+ fragment to be seen. Hydrogen atom transfer to the backbone carbonyl of proline could explain the high abundance of the c_9^+ fragment (once again the $c_8^{+\bullet}$ fragment could not be observed in this case, due to the presence of proline).

The atomic distances of (His)N-HE2 with backbone carbonyls for this dynamics simulation are shown in Figure 4.25. Whereas, the hydrogen bond to the

backbone carbonyl of leucine can be rationalised (see Chapter 3, Figure 3.37), the hydrogen bond to the backbone carbonyl of tyrosine cannot be rationalised with the relative ECD z fragment intensities for $[LHRH\_HR\_NH_2]^{2+}$.
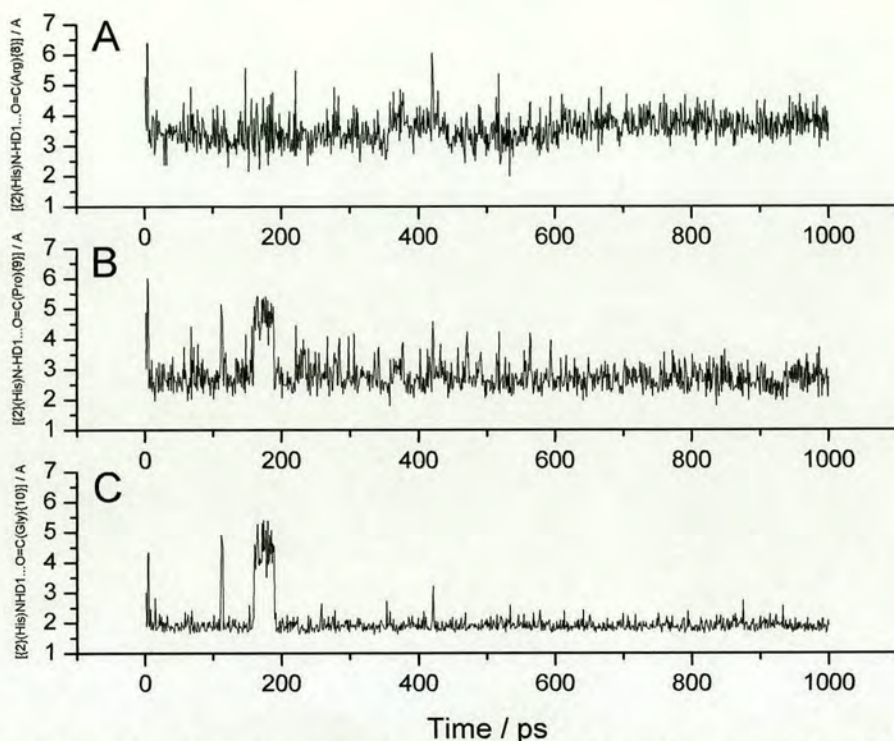
Figure 4.23 Atomic distances for molecular dynamics at 300 of [LHRH _HR_NH₂]²⁺ min61 (A) [{2}(His)N-HD1…O=C(Arg){8}], (B) [{2}(His)N-HD1…O=C(Pro){9}] and (C) [{2}(His)N-HD1…O=C(Gly){10}].

Figure 4.24 Representative ribbon structure of molecular dynamics at 300 K of [LHRH _HR_NH₂]²⁺ min61. Structure 300K.300ps.

169

Figure 4.25 Atomic distances for molecular dynamics at 300 K of [LHRH_NH$_2$_HR]$^{2+}$ min61 (A) [{2}(His)N-HE2…O=C(Leu){7}] and (B) [{2}(His)N-HE2…O=C(Tyr){5}].

4.4.1.4 Conclusions for [LHRH_HR_NH$_2$]$^{2+}$

In summary, the 10 lowest energy structures of [LHRH _HR_NH$_2$]$^{2+}$ have not confirmed hydrogen bonding of the histidine N-H's to backbone carbonyls that show enhanced ECD backbone fragmentation and in particular not to CO$_{serine\{4\}}$, which shows the highest ECD fragmentation (Chapter 3, Figure 3.37). Nevertheless, the 'loop' structure was again found to be adopted for the molecular dynamics of both min48 and min61.

4.4.2 Modelling of [LHRH_HR_NH$_3$]$^{2+}$

4.4.2.1 Minimised structures

A possible competitive protonation site in amidated LHRH peptides is the C-terminal amide NH$_2$ group instead of the histidine side-chain. The gas-phase basicity of this amide group is likely to be similar to the gas-phase basicity of asparagine or glutamine (see Appendix 2, Table A2.2), which is lower than the gas-phase basicity of histidine and even lower than other residues in the molecule (i.e. proline and tryptophan). Since all naturally occurring LHRH peptides are amidated at the C-terminus, is was deemed to be interesting to determine how the modelled structures

170

change when this site is protonated instead of the histidine side-chain. The [LHRH_HR_NH$_3$]$^{2+}$ (sequence pEHWSYGLRPG-NH$_2$; protonation sites: side-chain of arginine and amidated C-terminus) peptide was built in LeAP and the conformational space was probed using the annealing procedure.

The relative energies of the 100 lowest energy structures are shown in Figure 4.26. Of these, structures min47, min76 and min89 are lowest in energy. The strain energy for the lowest energy conformer, as derived by AMBER, is E_S[LHRH_HR_NH$_3$]$^{2+}${min89} = -214.64 kcal mol^{-1}, which is energetically more favourable to the strain energy determined for the corresponding lowest energy structure for [LHRH_HR_NH$_2$]$^{2+}$ (E_S[LHRH_HR_NH$_2$]$^{2+}${min48} = -184.17 kcal mol^{-1}). Since [LHRH_HR_NH$_2$]$^{2+}$ and [LHRH_HR_NH$_3$]$^{2+}$ are structural isomers, their relative energies can in principle be compared. The difference in gas-phase basicity of histidine (GB$_{histidine}$ = 223.7 kcal mol^{-1}) and the C-terminal amide group (GB$_{asparagine}$ = 214.2 kcal mol^{-1}) is approximately 10 kcal mol^{-1}, whereas the strain energy difference for the structural conformers [LHRH_HR_NH$_2$]$^{2+}$ min48 and [LHRH_HR_NH$_3$]$^{2+}$ min89 is approximately 30 kcal mol^{-1}. Hence, these results suggest that protonation of the amidated C-terminus is in fact energetically favourable to protonation at the histidine side-chain. The ribbon structures, as well as the hydrogen bonding interactions of the C-terminal NH$_3$$^+$ group (if present) are shown in Figure 4.27. Whereas, min47 and min76 (Figure 4.26 A and B) have a β-turn for residues 5-8, the lowest energy structure min89 shows a β-turn for residues 1-5.
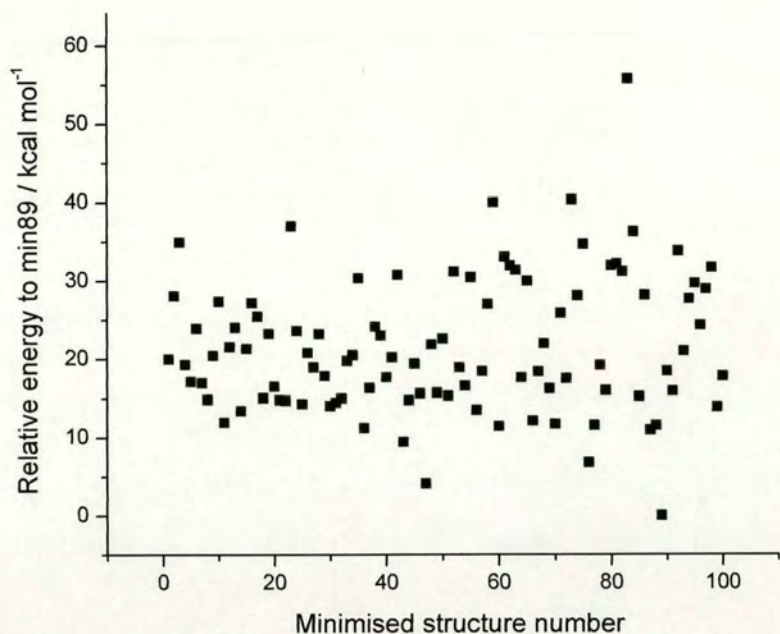
Figure 4.26 Relative energies of minimised structures obtained from simulated annealing of $[LHRH\_HR\_NH_3]^{2+}$, normalised with respect to $E_S[LHRH\_HR\_NH_3]^{2+}\{min89\} = -214.64$ kcal mol^{-1}.
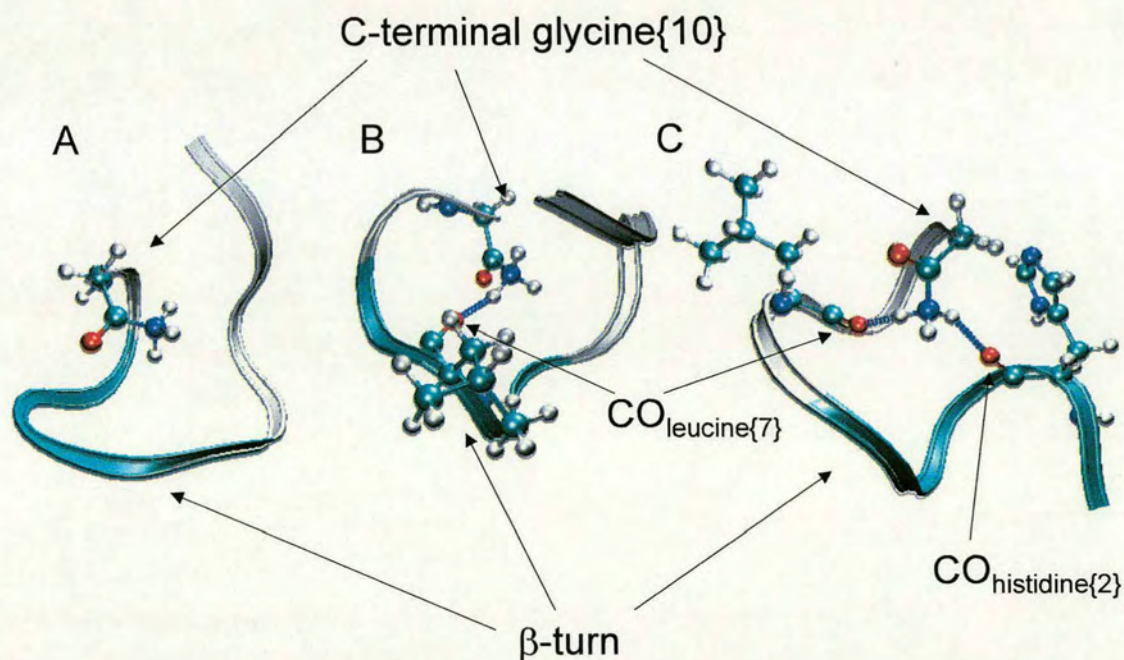

Figure 4.27 Ribbon structures of minimised structures for $[LHRH\_HR\_NH_3]^{2+}$ (A) min47, (B) min76 and (C) min89. β-turn represented in cyan.

4.4.2.2 Molecular dynamics of min89

Min89 was subjected to molecular dynamics at 300 K (see Figure 4.28). The initial structure changes due to the higher internal energy, but stabilises after about 300 ps dynamics (Figure 4.28 B). The β-turn around residues 4-6 for the 300 K dynamics is reproduced, although the N-terminus of the molecule still points away thus causing a 'non-loop' type structure.
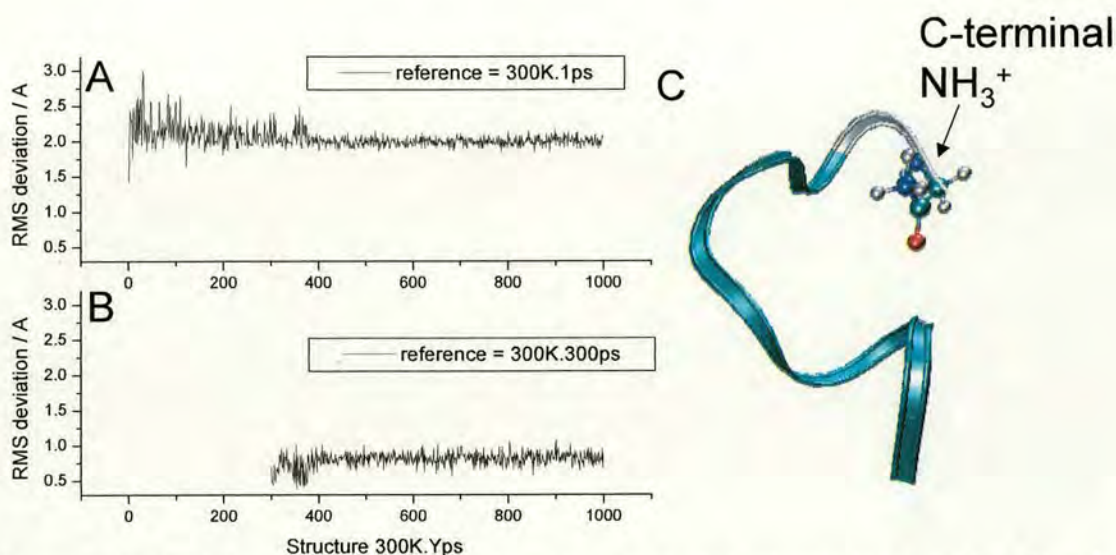
Figure 4.28 RMS deviation of molecular dynamics simulation at 300 K of $[LHRH\_HR\_NH_3]^{2+}$ min89 for backbone N, C_α, C and O for (A) reference = 300K.1ps and (B) reference = 300K.300ps. (C) Representative structure for simulation (300K.300ps).

4.4.2.3 Conclusions for $[LHRH\_HR\_NH_3]^{2+}$

Protonation of the C-terminal amide group renders the analysis of hydrogen bonding interactions of this group with backbone carbonyls more complicated, since no preferential hydrogen bonding exists. While a β-turn involving residues 4-6 is found for min89 at higher internal energies, the termini of the molecule are not close in space (i.e. non-loop type structure), showing that different protonation of the peptide yields different minimised gas-phase structures. Nevertheless, protonation of the amidated C-terminus may only represent a small proportion of ions in the gas-phase due to the higher gas-phase basicity of histidine. Moreover, basicity of amides in solution is very low [27] and since ions are made by electrospray ionisation (ESI)

the solution phase structure may need to be taken into account. On the other hand, ions are held in an accumulation hexapole for 600-1000 ms, which will cause heating to the ions and hence gas-phase conformations of the ions should be favoured.

4.5 Modelling of [LHRH_(l)W]$^{2+}$ and [LHRH_(d)W]$^{2+}$

4.5.1 Minimised structures

LHRH peptides with a tryptophan residue in position 6 were modelled. The lowest energy structures for [LHRH_(l)W]$^{2+}$ (sequence pEHWSY(l)WLRPG-NH$_2$; protonation sites at the side-chains of arginine and histidine) and [LHRH_(d)W]$^{2+}$ (sequence pEHWSY(d)WLRPG-NH$_2$; protonation sites at the side-chains of arginine and histidine) are shown in Figure 4.29. While these structures are quite dissimilar, the calculated strain energies E$_S$ for both are very similar.
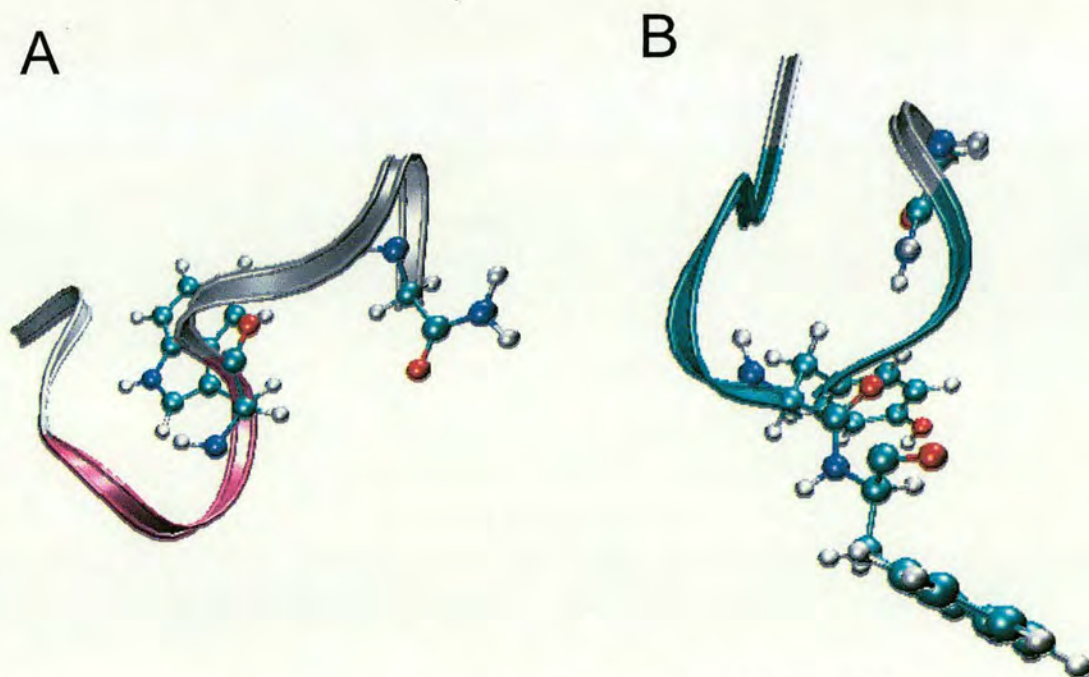
Figure 4.29 Lowest energy structures of (A) [LHRH_(l)W]$^{2+}$ min23 E$_S$ = -172.78 kcal mol^{-1} and (B) [LHRH_(d)W]$^{2+}$ min47 E$_S$ = -169.79 kcal mol^{-1} from simulated annealing. Purple colour indicates α-helix and cyan colour indicates β-turn. Note that VMD does not recognise (d)Trp as a residue to show ribbon structure.

4.5.2 Molecular dynamics

Molecular dynamics was carried out on both these structures at 300 K. Figures 4.30 and 4.31 show representative gas-phase conformations of both these peptide ions for the dynamics simulation.

The lowest energy structure for $[LHRH\_(l)Trp]^{2+}$ (Figure 4.29 A) shows an α-helix around residues 4-6 (Ser-Tyr-Trp), which is conserved at higher internal energies (Figure 4.30). This structure resembles a 'twist' and is very different to the 'loop' type structure found for the lowest energy structure of $[LHRH\_(d)Trp]^{2+}$ (Figure 4.29 B). The 'loop' structure is conserved for $[LHRH\_(d)Trp]^{2+}$ at higher internal energies (Figure 4.31). For $[LHRH\_HR]^{2+}$ (Figure 4.11) and $[LHRH\_HR\_NH_2]^{2+}$ (Figure 4.22) the 'loop' structure was also found to be stable at higher internal energies. Whereas, ECD of $[LHRH\_HR]^{2+}$, $[LHRH\_HR\_NH_2]^{2+}$ and $[LHRH\_(d)Trp]^{2+}$ had shown that the abundance of $z_6^{+\prime} > z_5^{+\bullet}$ (Chapter 3, Figures 3.15, 3.18 and 3.21), the opposite is the case for $[LHRH\_(l)Trp]^{2+}$ (Chapter 3, Figure 3.19). These modelling results seem to indicate that the competition between 'normal' and 'reverse' alpha-cleavage could be affected by the gas-phase conformation of the peptide ion.
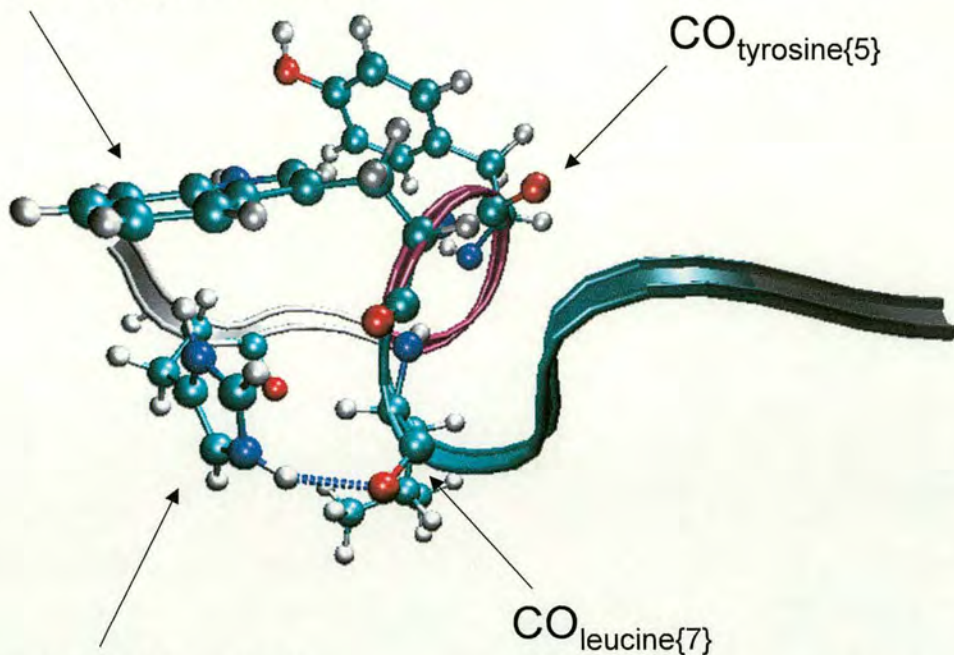
Figure 4.30 Molecular dynamics structure for $[LHRH\_(l)W]^{2+}$ min23 300K.300ps. Structure resembles a 'twist'.
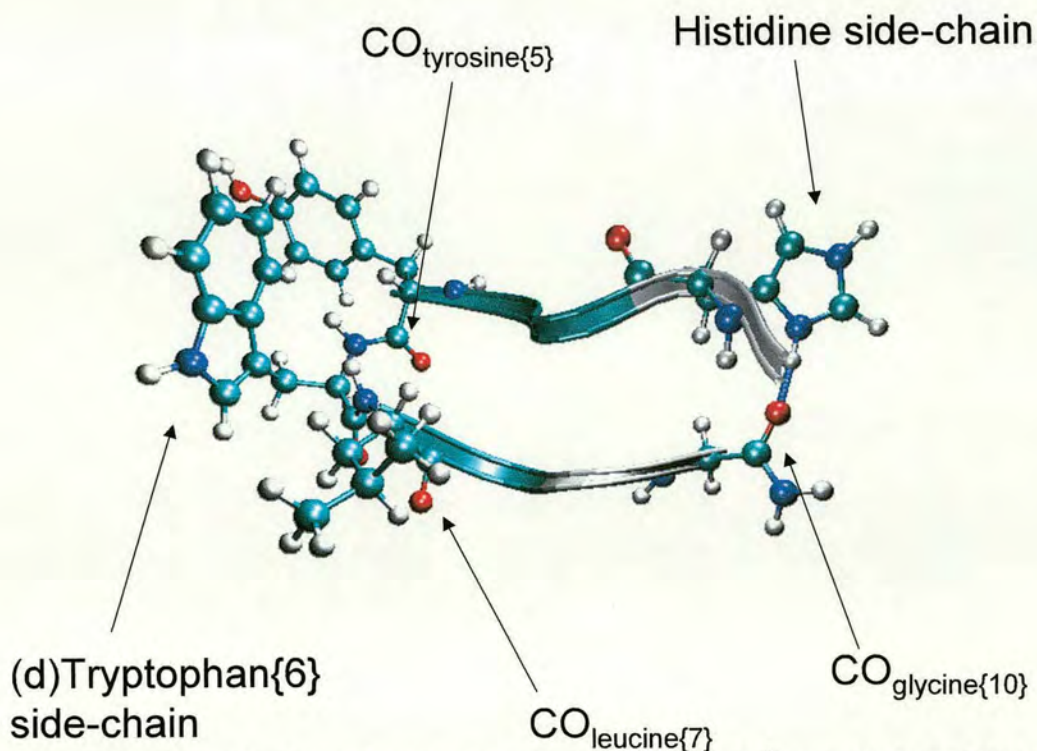

Figure 4.31 Molecular dynamics structure for $[LHRH\_(d)W]^{2+}$ min47 300K.300ps.

4.5.3 Conclusions for LHRH variants

In summary, modelling of these LHRH variants *in vacuo* has shown that a 'loop' structure involving a β-turn is energetically favourable for the variants with protonation sites on the arginine and histidine side-chain, which therefore confirms solution phase modelling results by Guanieri [8]. The exception to this rule seems to be [LHRH_(l)Trp]$^{2+}$, which has also been confirmed by NMR structure results [5]. In terms of the correlation between hydrogen bonding of the protonated site to backbone carbonyls and enhanced cleavage at those carbonyls, results for [LHRH_HR]$^{2+}$ suggest good correlation for the lowest energy minimised structures, whereas results for [LHRH_HR_NH$_2$]$^{2+}$ show no such agreement.

The relative ECD fragmentation patterns for these LHRH variants (Chapter 3, Figure 3.37) are similar, which could be explained by similar gas-phase structures. If [LHRH_(l)Trp]$^{2+}$ has a different gas-phase conformation to the other LHRH variants, but gives a similar ECD fragmentation pattern, this would show that the secondary structure of the peptide does not influence the fragmentation pattern significantly. This could, however, rationalise whether the relative ECD fragment abundances for z_6^{+} and $z_5^{+\bullet}$ are related to the gas-phase conformation.

4.6 Modelling of bombesin

4.6.1 Modelling of [BOMB_NH$_2$]$^{2+}$

4.6.1.1 Minimised structures

Since hydrogen bonding in 'small' peptide such as the LHRH variants may not be as robust, the larger peptide bombesin was also modelled using the AMBER force field. The doubly protonated bombesin peptide (sequence: pEQRLGNQWAVGHLM-NH$_2$) was first attempted with protonation of the side-chains of arginine and histidine – [BOMB_NH$_2$]$^{2+}$. The relative energies of the 100 candidate structures generated by the simulated annealing procedure are shown in Figure 4.32.
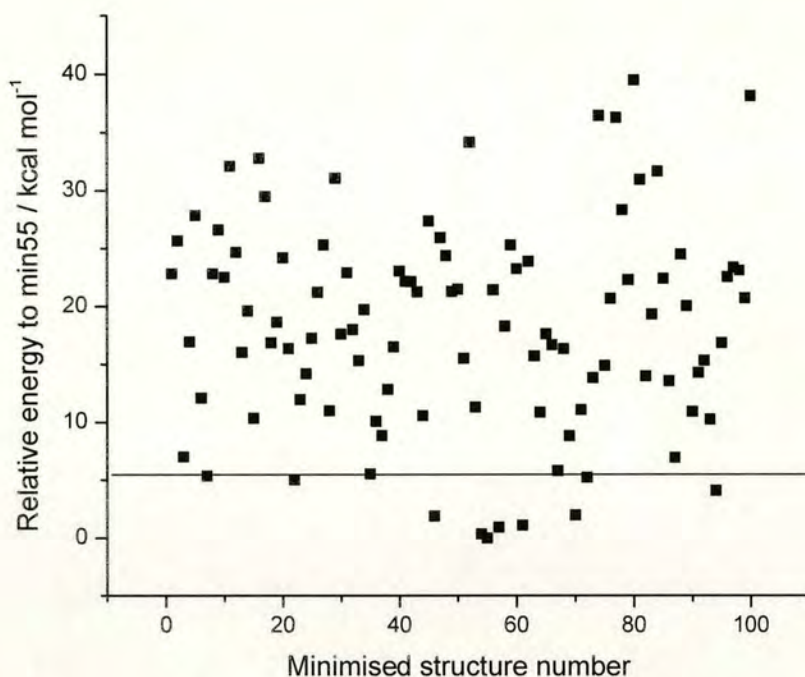
Figure 4.32 Relative energies of minimised structures obtained from simulated annealing of $[BOMB\_NH_2]^{2+}$, normalised with respect to $E_S[BOMB\_NH_2]^{2+}\{min55\}$ = –409.74 kcal mol^{-1}.

The ribbon structures of the 10 lowest energy minimised structures are shown in Figure 4.33. An analysis of the hydrogen bonding interactions of the histidine side-chain N-H's with backbone carbonyls for the 10 lowest energy structures is summarised in Table 4.3.
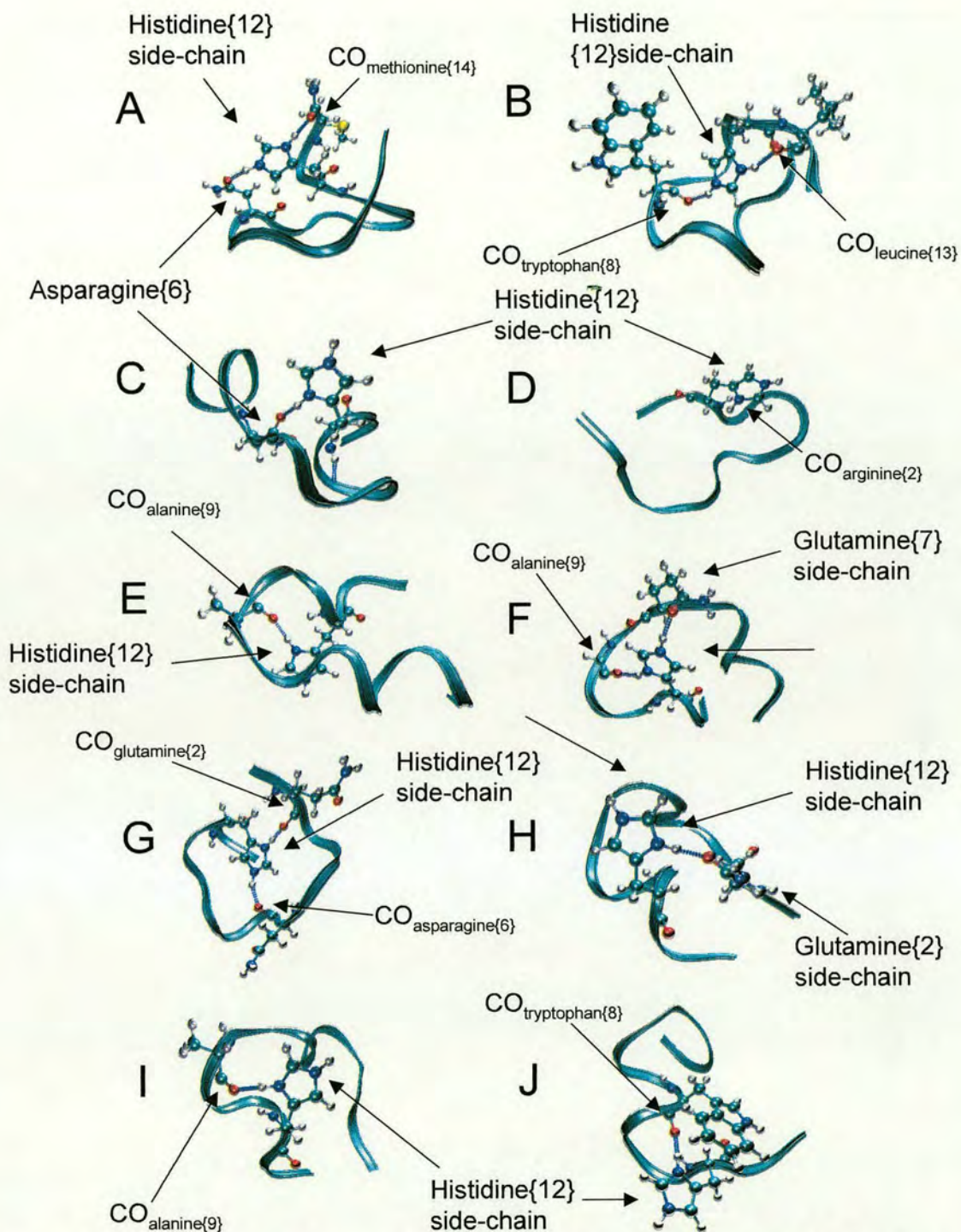
Hydrogen bonding of the histidine side-chain N-H's to the backbone carbonyl of tryptophan{8} is predicted for 2 minimised structures, whereas hydrogen bonding to the backbone carbonyl of alanine{9} is predicted for 3 minimised structures. Experimentally, these backbone carbonyls show higher ECD cleavage (see Chapter 3, Figure 3.38). On the other hand, 3 of the lowest energy structures predict [{12}(His)N-H...O=C(Asn){6}] to be hydrogen bonded, which however does not exhibit a high ECD cleavage rate.

CO	min 7	min 22	min 46	min 54	min 55	min 57	min 61	min 70	min 72	min 94	Total
pGlu{1}											0
Gln{2}							Y	Y			2
Arg{3}											0
Leu{4}											0
Gly{5}											0
Asn{6}	Y		Y				Y				3
Gln{7}						Y					1
Trp{8}		Y								Y	2
Ala{9}				Y	Y				Y		3
Val{10}											0
Gly{11}											0
His{12}											0
Leu{13}		Y									1
Met{14}	Y										1

Table 4.3 Table summarising hydrogen bonding interactions of histidine side-chain N-H's with backbone carbonyls (CO) for 10 lowest energy structures in Figure 4.33.

Figure 4.33 Low-energy ribbon structures for annealing of [BOMB_NH$_2$]$^{2+}$ showing histidine side-chain and hydrogen bonding interaction with backbone carbonyl if present (A) min7, (B) min22, (C) min46, (D) min54, (E) min55, (F) min57, (G) min61, (H) min70, (I) min72 and (J) min94.

4.6.1.2 Molecular dynamics of min22

Molecular dynamics at 300 K are shown for three of these structures: min22, min55 and min61. Figure 4.34 shows the RMS deviation of the backbone of min22 during the 300 K simulation. The structure seems to stabilise after 50 ps dynamics. (Figure 4.34 B). A representative 300 K dynamics structure (shown in Figure 4.34 C) shows that hydrogen bonding of the histidine N-H's to the backbone carbonyls of tryptophan{8} and leucine{13} is conserved (compare to Figure 4.33 B). Figure 4.35 shows the atomic distances for both of these hydrogen bonds.
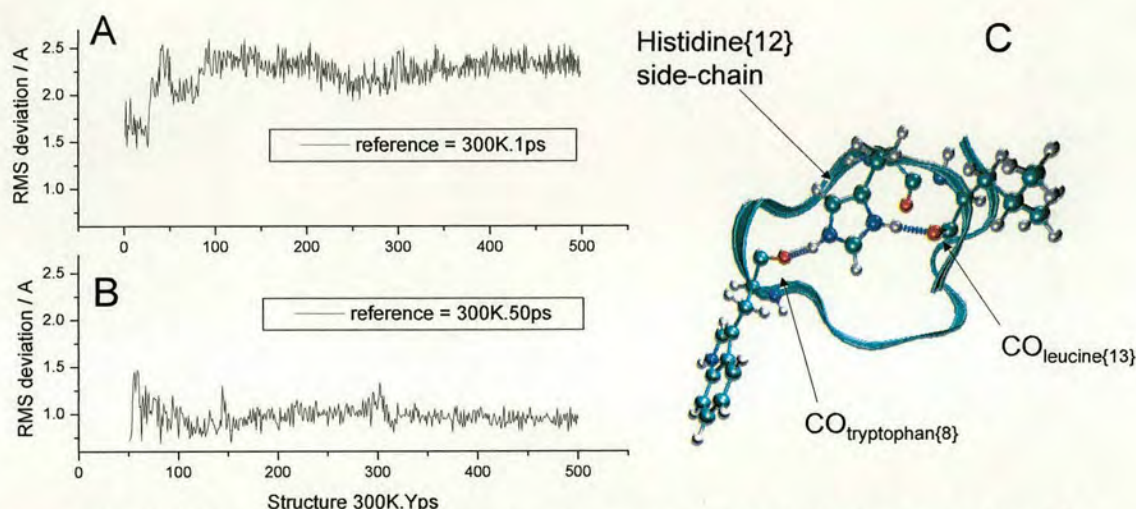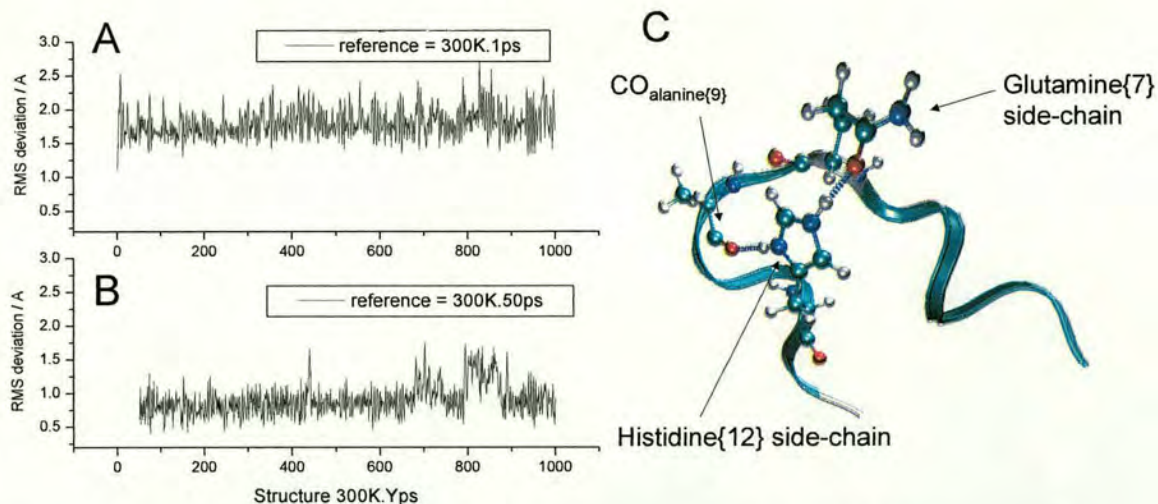

Figure 4.34 RMS deviation for molecular dynamics at 300 K of backbone atoms for [BOMB_NH2]$^{2+}$ min22 for (A) reference = 300K.1ps, (B) reference = 300K.50ps. (C) Ribbon structure 300K.59ps showing hydrogen bonding of histidine side-chain.
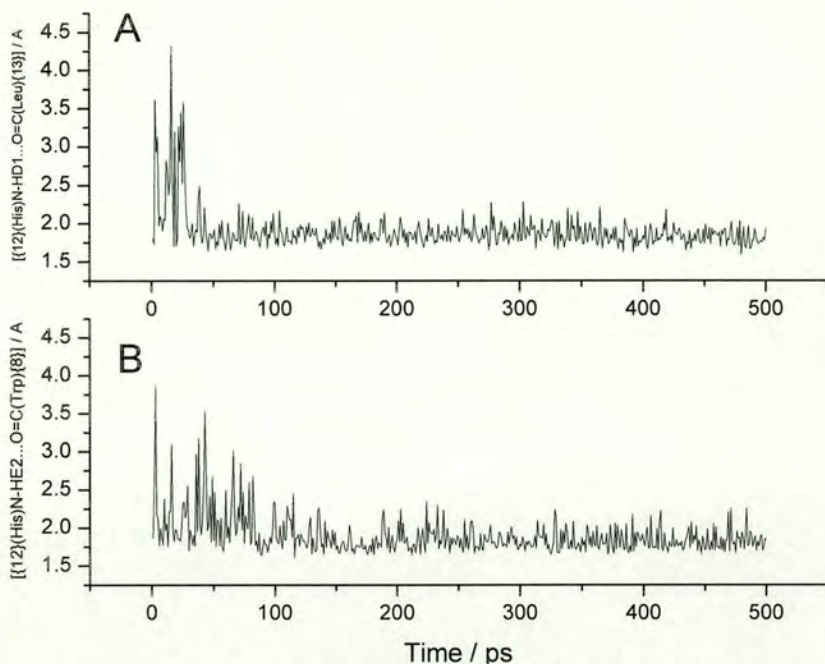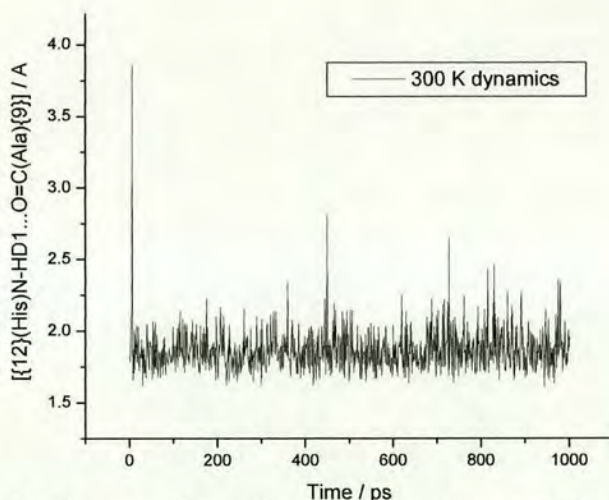
Figure 4.35 Atomic distances for molecular dynamics at 300 of [BOMB_NH2]$^{2+}$ min22 (A) [{11}(His)N-HD1...O=C(Leu){13}] and (B) [{11}(His)N-HE2...O=C(Trp){8}].

4.6.1.3 Molecular dynamics of min55

The backbone RMS deviation for the molecular dynamics simulation of min55 is shown in Figure 4.36. The structure stabilises after 50 ps dynamics and a representative structure is shown in Figure 4.36 C. This structure is very similar to the minimised structure min57 (see Figure 4.33 F), showing hydrogen bonding of the histidine N-H's to CO$_{alanine\{9\}}$ and the side-chain of glutamine{7}. Figure 4.37 supports the stability of the hydrogen bonding interactions for [{12}(His)N-HD1...O=C(Ala){9}] at 300 K. Dynamics at 500 K, however, lead to a break-up of this hydrogen bond (not shown).

Figure 4.36 RMS deviation for molecular dynamics at 300 K of backbone atoms for [BOMB_NH2]$^{2+}$ min55 for (A) reference = 300K.1ps, (B) reference = 300K.50ps. (C) Ribbon structure 300K.50ps showing hydrogen bonding of histidine.

Figure 4.37 Atomic distance for molecular dynamics at 300 of [BOMB_NH2]$^{2+}$ min55 [{12}(His)N-HD1…O=C(Ala){9}].

4.6.1.4 Molecular dynamics of min61

Figure 4.38 shows the backbone RMS deviation for min61 at 300 K dynamics, which deviates significantly from the initial structure (300K.1ps). A representative stabilised structure (300K.50ps) for the remainder of this dynamics simulation is shown in Figure 4.38 C. For this structure hydrogen bonding is observed for [{12}(His)N-HD1…O=C(p-Glu){1}] and [{12}(His)N-HE2…O=C(Gln){7}]. The hydrogen bonding of the histidine side-chain N-H's to the

backbone carbonyls of glutamine{2} and asparagine{6} in min61 are broken in favour of new hydrogen bonding to the backbone carbonyls of pyro-glutamic acid{1} and glutamine{7} (see Figure 4.39). Note that these hydrogen bonds correlate with higher ECD cleavage associated with the backbone carbonyls of pyro-glutamic acid{1} and glutamine{7} (see Chapter 3, Figure 3.38).
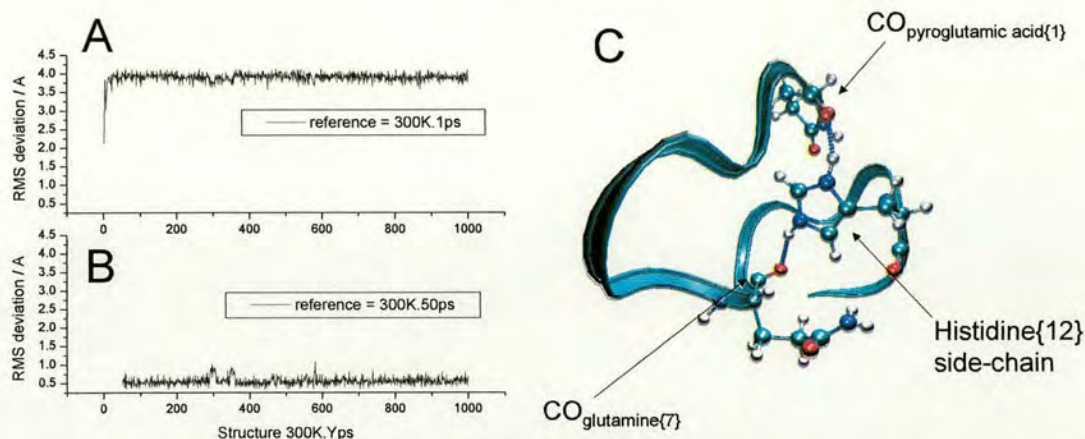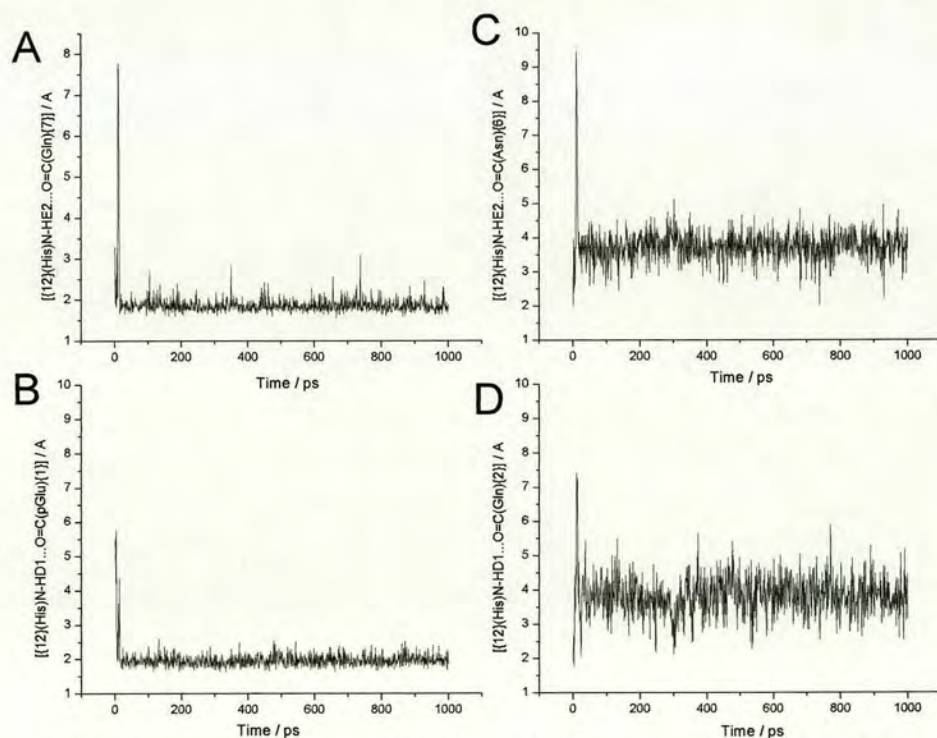

Figure 4.38 RMS deviation for molecular dynamics at 300 K of backbone atoms for [BOMB_NH2]$^{2+}$ min61 for (A) reference = 300K.1ps, (B) reference = 300K.50ps. (C) Ribbon structure 300K.50ps showing hydrogen bonding of histidine side-chain.

Figure 4.39 Atomic distances for [BOMB_NH$_2$]$^{2+}$ min61 dynamics at 300 K: (A) [{12}(His)N-HE2…O=C(Gln){7}], (B) [{12}(His)N-HD1…O=C(pGlu){1}], (C) [{12}(His)N-HE2…O=C(Asn){6}] and (D) [{12}(His)N-HD1…O=C(Gln){2}].

4.6.1.5 Conclusions for [BOMB_NH₂]²⁺

In summary, these results seem to indicate that hydrogen bonding of the histidine side-chain to the backbone carbonyls of alanine{9} and tryptophan{8} is predicted for low-energy structures and remains conserved at higher internal energies of 300 K. Hydrogen bonding of the histidine side-chain to the backbone carbonyl of asparagine{6} was also predicted for low-energy structures, however, at higher internal energies this hydrogen bond is not stable and is replaced by hydrogen bonding to glutamine{7}. These results correlate well with the relative ECD cleavage of [BOMB]²⁺ in Figure 3.38 (Chapter 3).

4.6.2 Modelling of [BOMB_NH₃]²⁺

4.6.2.1 Minimised structures

A competing site of protonation in the bombesin peptide is the amidated C-terminus. The peptide [BOMB_NH₃]²⁺ (sequence: pEQRLGNQWAVGLM-NH₂, protonation on the side-chain of arginine and the amidated C-terminus) was built in LeAP. The relative energies of the structures obtained from the simulated annealing are shown in Figure 4.40. Once again, in analogy to [LHRH_HR_NH₂]²⁺ and [LHRH_HR_NH₃]²⁺, the strain energy for the lowest energy structure for [BOMB_NH₃]²⁺ (E_S[BOMB_NH₃]²⁺{min55} = −491.59 kcal mol⁻1) is more favourable than in the case of [BOMB_NH₂]²⁺ (E_S[BOMB_NH₂]²⁺{min55} = −409.74 kcal mol⁻1). The ribbon backbone structures of the 10 lowest energy structures are shown in Figure 4.41.
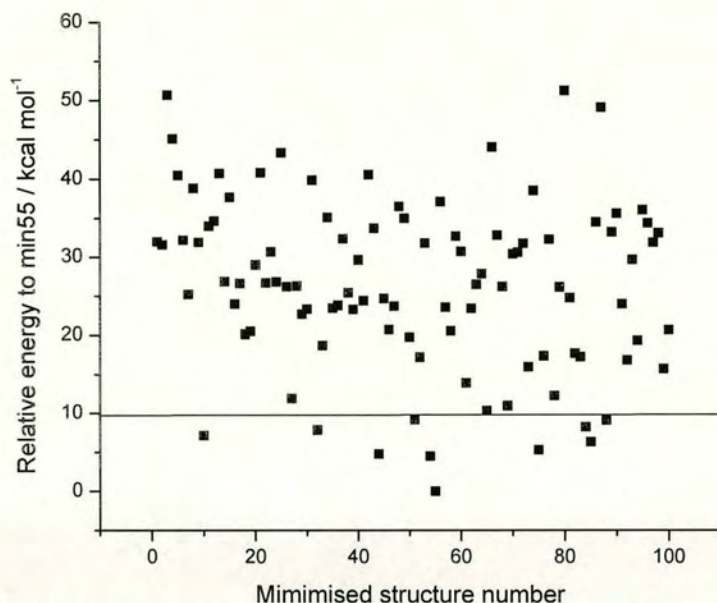
Figure 4.40 Relative energies of minimised structures of $[BOMB\_NH_3]^{2+}$ generated by simulated annealing, normalised with respect to $E_S[BOMB\_NH_3]^{2+}\{min55\} = -491.59$ kcal mol^{-1}.

The minimised structures obtained form the simulated annealing of $[BOMB\_NH_3]^{2+}$ are very different to the ones obtained for $[BOMB\_NH_2]^{2+}$. This shows that the protonation site plays a significant role in determining the gas-phase structure. The hydrogen bonding interactions of the C-terminal NH_3^+ group with backbone carbonyls are summarised in Table 4.4.
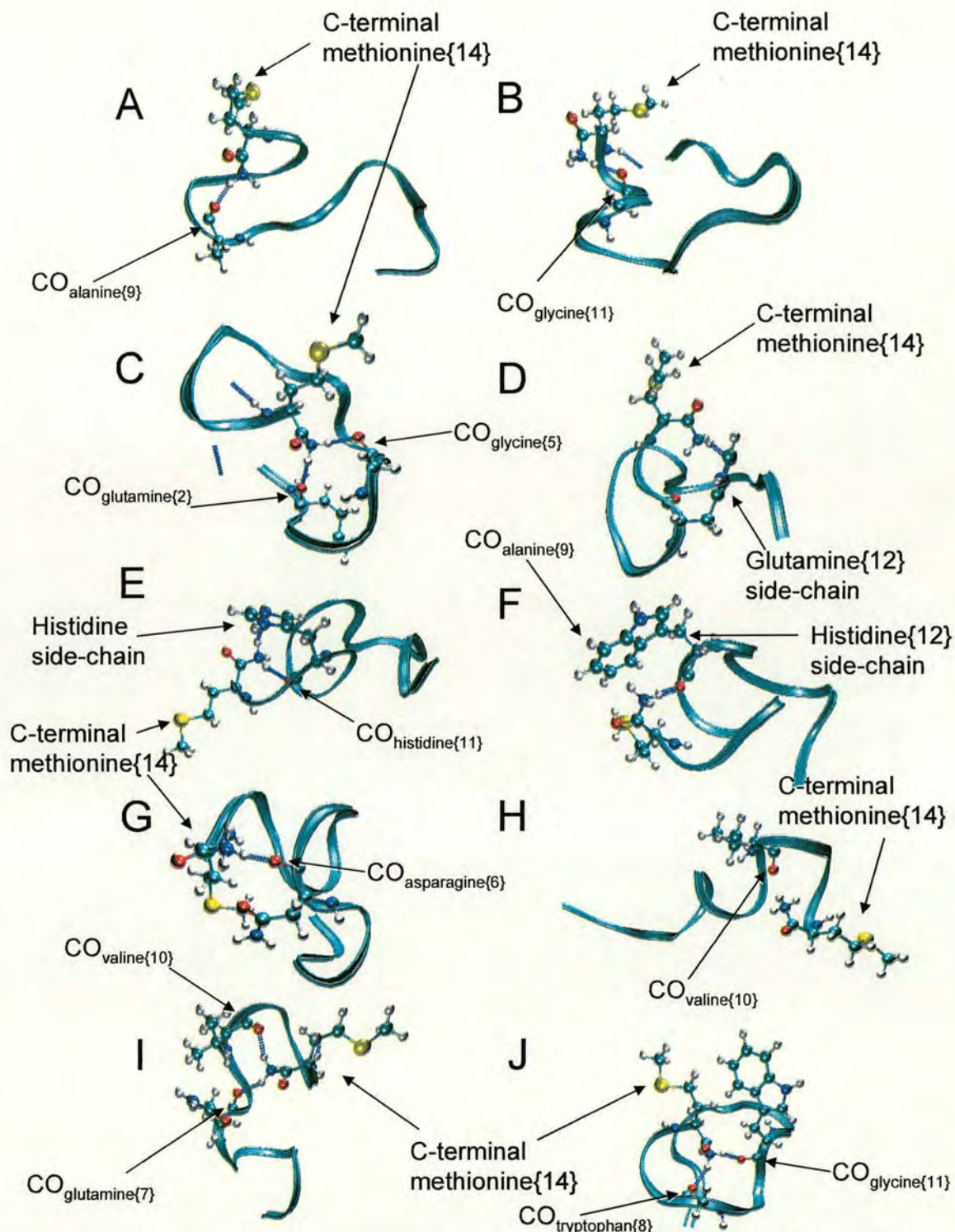
186

Figure 4.41 Low-energy ribbon structures for annealing of [BOMB_NH$_3$]$^{2+}$ showing histidine side-chain and hydrogen bonding interaction with backbone carbonyl if present (A) min10, (B) min32, (C) min44, (D) min51, (E) min54, (F) min55, (G) min75, (H) min84, (I) min85 and (J) min88.

CO	min 10	min 32	min 44	min 51	min 54	min 55	min 75	min 84	min 85	min 88	Total
pGlu{1}			Y								1
Gln{2}											0
Arg{3}											0
Leu{4}											0
Gly{5}			Y								1
Asn{6}							Y				1
Gln{7}									Y		1
Trp{8}						Y				Y	2
Ala{9}	Y										1
Val{10}									Y	Y	2
Gly{11}		Y									1
His{12}					Y						1
Leu{13}											0
Met{14}											0

Table 4.4 Summary of hydrogen bonding interactions of C-terminal NH_3^+ group with backbone carbonyls (CO) for 10 lowest energy structures in Figure 4.41.

4.6.2.2 Molecular dynamics and conclusions

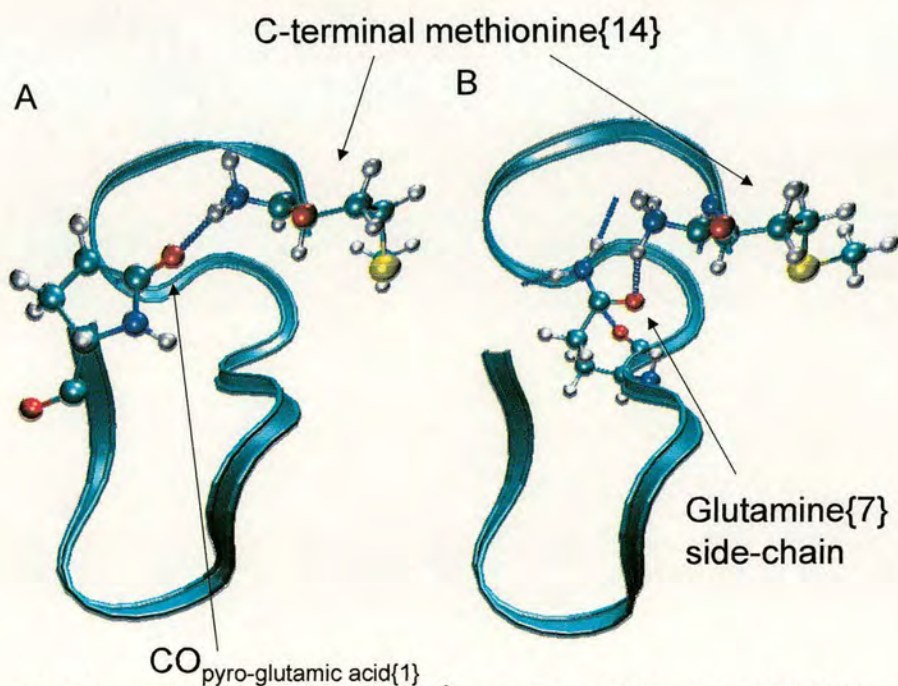**Figure 4.42** Dynamics of $[BOMB\_NH_3]^{2+}$ min55 at 300 K (A) 300K.100ps and (B) 300K.200ps.

Dynamics of min55 (lowest energy structure) shows that the C-terminal NH_3^+ group enters into hydrogen bonding contact with many backbone carbonyls (see Figure 4.42) and hence preferential hydrogen bonding is much more difficult to establish. In terms of an *ECD/Hbond correlation study* protonation of the C-terminal amide is less useful than protonation of the histidine side-chain. As argued before (section 4.4.2.3), the second protonation site in this molecule (first protonation site is arginine side-chain) is a matter of controversy.

4.7 Conclusions and further work

Molecular modelling of the doubly protonated LHRH variants $[LHRH\_HR]^{2+}$, $[LHRH\_HR\_NH_2]^{2+}$ and $[LHRH\_(d)W]^{2+}$ gives rise to low-energy 'loop' structures involving a β-turn around residues 4-6, which are stable at higher internal energies. For $[LHRH\_(l)W]^{2+}$ a different lowest-energy structure is predicted. This correlates with a difference in the relative abundances of ECD fragments $z_5^{+\bullet}$ and $z_6^{+'}$ for $[LHRH\_(l)W]^{2+}$ as opposed to the other three LHRH variants. More studies on structural LHRH variants should be conducted to see whether the preference for ECD fragments $z_5^{+\bullet}$ over $z_6^{+'}$ is related to the conformation of the peptide.

The results of the *ECD/Hbond correlation* study are ambiguous. Whereas, the lowest energy minimised structures for $[LHRH\_HR]^{2+}$ and $[BOMB\_NH_2]^{2+}$ and molecular dynamics show that such a correlation seems to exist, the results for $[LHRH\_HR\_NH_2]^{2+}$ do not support these findings. Further, the second protonation site in the C-terminal amidated peptides is more controversial. Whereas, the higher gas-phase basicity of histidine should favour protonation of the histidine side-chain, calculated strain energies for the low-energy structures of mGnRH and bombesin show that protonation of the C-terminal amide is energetically favourable.

Protonation of the C-terminal amide group in favour of the histidine side-chain results in modelling results with no preferential hydrogen bonding pattern. In terms of using ECD as a tool to study the gas-phase structure the protonation of this group is of less diagnostic value.

Modelling of these peptides has shown that the arginine side-chain N-H's do engage in hydrogen bonding interactions with backbone carbonyls even at higher internal energies. In terms of the *hot hydrogen re-arrangement* mechanism the low transfer of hydrogen atoms from the arginine side-chain to the backbone is therefore not due to lower hydrogen bonding interactions than is the case for the histidine side-chain N-H's.

The gas-phase structures of these peptides should be investigated by ion mobility measurements. The collision cross-sections of the 'loop' structure ($[LHRH\_RH]^{2+}$, $[LHRH\_HR\_NH_2]^{2+}$ and $[LHRH\_(d)W]^{2+}$) and the 'twist' structure ($[LHRH\_(l)W]^{2+}$) should be different enough to differentiate them. If the gas-phase structures of the studied LHRH variants, with the exception of $[LHRH\_(l)W]^{2+}$, confirm the 'loop' structure, this could explain the differences in relative abundances of ECD fragments $z_5^{+\bullet}$ and $z_6^{+\bullet}$. Conversely, this would also imply that the *ECD/Hbond correlation* does not always hold, since it would be difficult to explain why different gas-phase conformations give rise to very similar fragmentation patterns.

The accuracy of hydrogen bonding interactions as predicted by force field modelling (i.e. electrostatic models) has been seriously questioned by Marten *et al.* and Kim *et al.* due [23,24,28]. Kim *et al.* suggested the use of polarisable force fields to improve the accuracy of hydrogen bonding interactions [23].

The newer force field in AMBER, ff02, is a polarisable variant of ff99 which is more suited for gas-phase modelling. The LHRH peptides in this study should be modelled using ff02; it would be interesting to see if the 'loop' structure is again predicted as a stable low-energy structure and what preferential hydrogen bonding interactions are predicted. Further, a Monte-Carlo type approach to modelling these peptide structures (as employed by Guarnieri *et al.* [8]) could establish the relative proportion of families of gas-phase structures and hence add a quantitative dimension to the modelling studies.

References:

(1)   Chang, J.; Williams, R. H.; Humphries, A. J.; Johansson, N.; Folkers, K.; Bowers, C. Y. *Biochem. Biophys. Res. Commun.* **1972**, *47*.

(2)   Fujino, M.; Kobayashi, S.; Obayashi, M.; Fukuda, T.; Shinagawa, S.; Yamazaki, I.; Nakayama, R. *Biochem. Biophys. Res. Commun.* **1972**, *49*, 698-705.

(3)   Millar, R.; Lowe, S.; Conklin, D.; Pawson, A.; Maudsley, S.; Troskie, B.; Ott, T.; Millar, M.; Lincoln, G.; Sellar, R.; Faurholm, B.; Scobie, G.; Kuestner, R.; Teresawa, E.; Katz, A. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9636-9641.

(4)   Momany, F. A. *J. Am. Chem. Soc.* **1976**, *98*, 2990-2996.

(5)   Maliekal, J. C.; Jackson, G. E.; Flanagan, C. A.; Millar, R. P. *S. Afr. J. Chem.* **1997**, *50*, 217-219.

(6)   Pfleger, K. D. G.; Bogerd, J.; Millar, R. P. *Mol. Endocr.* **2002**, *16*, 2155-2162.

(7)   Chary, K. V. R.; Srivastava, S.; Hosur, R. V.; Roy, K. B.; Govil, G. *Eur. J. Biochem.* **1986**, *158*, 323-332.

(8)   Guarnieri, F.; Weinstein, H. *J. Am. Chem. Soc.* **1996**, *118*, 5580-5589.

(9)   Weiner, S. J.; Kollmann, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta Jr., S.; Weiner, P. J. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.

(10)   Weiner, S. J.; Kollmann, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230-252.

(11)   Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. J. *J. Am. Chem. Soc.* **1990**, *112*, 6127-6129.

(12)   Sealfon, S. C.; Weinstein, H.; Millar, R. P. *Endocrine Rev.* **1997**, *18*, 180-205.

(13)   Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K., M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R.; Duan, Y.; Pitera, J.; Massova, I. G.; Seibel, G. L.; Singh, U. C.; Weiner, P. J.; Kollmann, P. A. In;; University of California: San Francisco.

(14)   Wang, J.; Cieplak, P.; Kollmann, P. A. *J. Comput. Chem.* **2000**, *21*, 1049-1074.

(15)   Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K., M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollmann, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

(16)   Cieplak, P.; Caldwell, J. W.; Kollmann, P. A. *J. Comput. Chem.* **2001**, *22*, 1048-1057.

(17)   Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A. J.; Stratman, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, A. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavacchari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gombertz, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanyakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnsen, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. In; Gaussian 98.; Gaussian Inc.: Pittsburgh, PA, 1998.

(18)   Cieplak, P.; Cornell, W. D.; Bayly, C. I.; Kollmann, P. A. *Computat. Chem.* **1995**, *16*, 1357-1377.

(19)   Tobias, D. J.; Mertz, J. E.; Brooks 3rd, C. L. *Biochemistry* **1991**, *30*, 6054-6058.

(20)   Kurz, M.; Mierke, D. F.; Kessler, H. *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 210-212.

(21)   McDonald, D. Q.; Still, W. C. *J. Am. Chem. Soc.* **1994**, *116*, 11550-11553.

(22)   Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563-573.

(23)   Kim, K.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 12952-12961.

(24)   Huang, N.; MacKerell, A. D. J. *J. Phys. Chem.* **2002**, *106*, 7820-7827.

(25)   Turecek, F.; Syrstad, E. A. *J. Am. Chem. Soc.* **2003**, *125*, 3353-3369.

(26)   Ryckaert, J.-P.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(27)   McMurry, J. In *Organic Chemistry*; 3rd ed.; Pantzis, H. C., Ed.; Brooks/Cole Publishing Company: Pacific Grove, CA, 1992; pp 963-966.

(28)   Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 11775-11788.

Chapter 5
Zn isotope exchange in the metallothionein Zn$_4$SmtA

5.1 Background

Zinc is an important metal in biological systems due principally to its role in zinc fingers [1], which regulate DNA expression, and metallothioneins [2,3]. The interaction of the protein with the metal, and in particular the metal-ligand cluster, determines the thermodynamic and kinetic properties of these metals [4], including their redox potentials [5,6] and transfer rates [7,8]. Whereas mammalian metallothioneins contain two types of clusters: Zn$_3$Cys$_9$ and Zn$_4$Cys$_{11}$ [9], bacterial metallothioneins possess only one type of cluster [10], Zn$_4$Cys$_9$His$_2$ for the case of cyanobacterial Zn$_4$SmtA (Figure 5.1) [11].
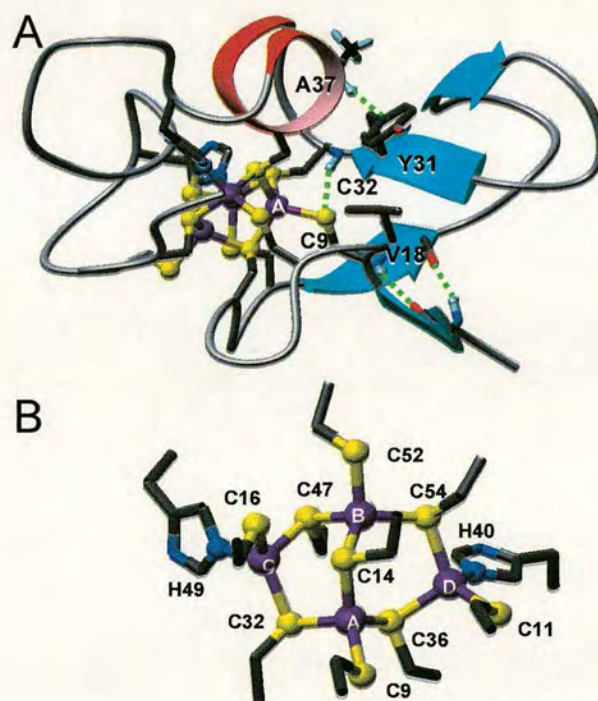
Figure 5.1 (A) 3D structure of Zn$_4$SmtA (PDB 1JJD). The Zn atoms are shown in purple, whereas the sulphur atoms are in yellow. The secondary structure shown includes an α-helix (red) and beta sheet (blue arrows). (B) Representation of Zn$_4$Cys$_9$His$_2$ cluster, showing Zn sites A-D and ligands binding to metals.

Mass spectrometry has been used to examine metalloproteins in conjunction with electrospray ionisation given the 'softness' of the ionisation [12-15]. Solvent conditions must be selected carefully to ionise the metal-protein complex without losing the metals in the process. In practical terms, this means that the solution pH must be buffered with, for example, ammonium acetate. FT-ICR MS is ideally suited to investigate these systems given its high mass resolution and accuracy [16]. When unit resolution is achieved the isotope distribution of the metal:protein complex can be fitted to the theoretical isotope distribution [17-20].

5.2 Materials and methods

The Zn_4SmtA protein is expressed in *Echerichia coli* BL21(DE3) cells and purified as described by Sadler and co-workers [11]. Samples for mass spectrometry were exchanged into 10 mM ammonium acetate (at pH 7.4) using pD10 gel filtration columns (Pharmacia), and were concentrated by ultrafiltration using Amicon YM3 filters (MW cutoff 3 kDa).

Apo-SmtA (i.e. not containing Zn) was prepared according to a literature procedure for mammalian metallothionein [21] by decreasing the pH to 1 and thus removing the zinc from the protein by gel filtration at pH 2 (0.01 M HCl) on a Pharmacia pD10 column under nitrogen. The protein was re-constituted as $^{67}Zn_4SmtA$ by adding a slight excess of a 0.1 M $^{67}ZnCl_2$ solution (prepared from ^{67}ZnO by dissolving in a minimum amount of 1 M HCl), and carefully raising the pH to 7.4 with 1 M NH_3. Samples were subsequently washed with 10 mM NH_4Ac (pH 7.4) using Amicon filters to remove excess zinc and chloride.

For the Zn exchange study, Zn_4SmtA with a natural Zn isotope distribution[1] (200 μM, in 10 mM NH_4Ac) was incubated at 310 K with a 10-fold molar excess (8 mM) of $^{67}ZnCl_2$ (prepared from ZnO, 93.11 % isotopic purity[2]). An aliquot (20 or 40 μL) was taken after various time intervals. The aliquot was diluted to 2.5 ml and

---

[1] Natural abundances of stable Zn isotopes: ^{64}Zn 48.6(3)%, ^{66}Zn 27.9(2)%, ^{67}Zn 4.1(1)%, ^{68}Zn 18.8(3)%, ^{70}Zn 0.6(1)%, taken from [22].

[2] ^{67}Zn isotopic purity = 93.11% ± 0.05, ^{64}Zn = 1.37 ± 0.02, ^{66}Zn = 2.58 ± 0.02, ^{68}Zn = 2.89 ± 0.02 and ^{70}Zn = 0.05 ± 0 (Oak Ridge National Laboratory, TN).

applied to a Pharmacia pD10 column, which had been equilibrated with 25 ml 10 mM NH_4Ac (for desalting). The total separation time was less than 3 min. The eluate (3.5 mL) was concentrated employing Amicon YM3 filters achieving a final concentration of about 20 μM protein in 10 mM NH_4Ac and 10% MeOH, and samples were analyzed by FT-ICR MS.

5.3 Preliminary results

In order to test the solution conditions necessary to study Zn_4SmtA by mass spectrometry, Zn_4SmtA was prepared with bovine ubiquitin (MW 8564.9, Sigma Aldrich) (4:1, Zn_4SmtA:ubiquitin) at a total concentration of 15 μM (1:1:0.12, $MeOH$:H_2O:CH_3COOH) and ionised by electrospray ionisation on the 3 Tesla FT-ICR MS instrument. Figure 5.2 shows the mass spectrum obtained by averaging 12 time-domain transients (512k). The mass spectrum was internally calibrated with the ubiquitin peaks (Ubi^{7+}, Ubi^{8+}, Ubi^{9+} and Ubi^{11+}.).

The ionisation efficiency of bovine ubiquitin is found to be higher than for Zn_4SmtA in these conditions. The 5+ charge state of Zn_4SmtA, $[M]^{5+}$, is observed to be the most abundant charge state for this species. Nevertheless, some losses do occur, such as the loss of threonine ($[M-Thr]^{5+}$) and, more importantly, the loss of Zn from the metalloprotein ($[M+7H-4Zn]^{7+}$, $[M+6H-4Zn]^{6+}$), giving rise to the apoprotein. The loss of threonine can be rationalised from the SmtA sequence[3], which would imply that the N-terminal threonine (T) is lost upon ionisation or in the transfer to the ICR cell. The loss of Zn takes place when the pH of the solution is too acidic or the solvent conditions too organic, both resulting in an unfolding of the protein. Thus, it appears that these solvent conditions were too harsh for this system.

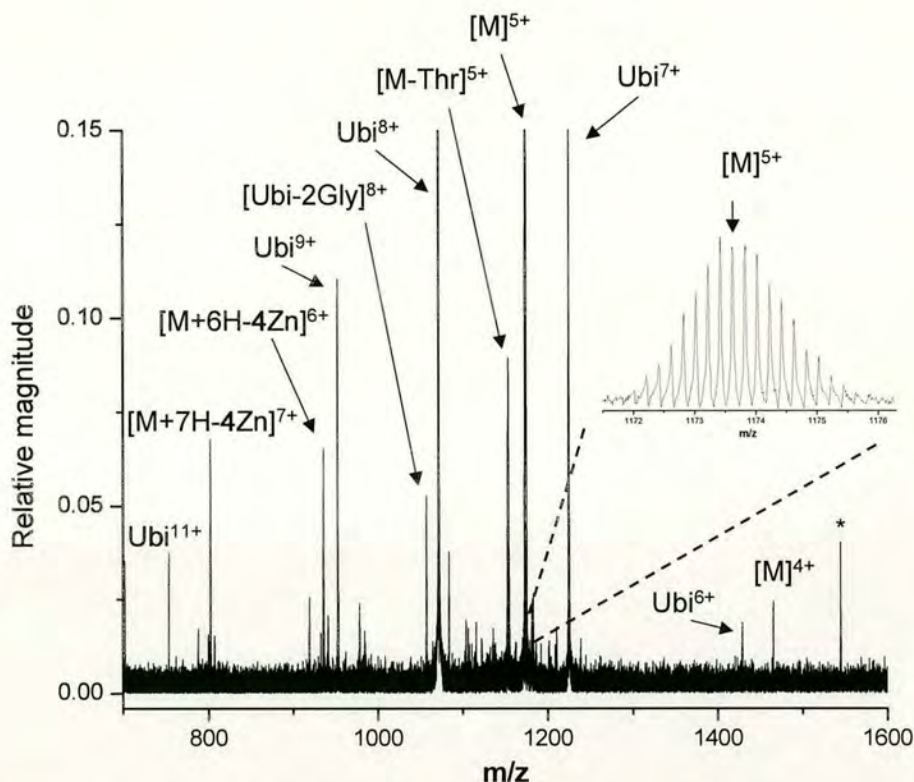[3] SmtA: TSTTLVKCACEPCLCNVDPSKAIDRNGLYYCSEACADGHTGGSKGCGHTGCNCHG

Figure 5.2 Raw ESI mass spectrum of Zn_4SmtA [M] and bovine ubiquitin (Ubi) (4:1, total concentration 15 μM) sprayed from $MeOH/H_2O/CH_3COOH$ (1:1:0.12). The base peak in the mass spectrum is Ubi^{8+}. The insert shows the isotopic distribution of $[M]^{5+}$; arrow indicates the centre peak. The pH was not measured.

The insert in Figure 5.2 shows the isotopic distribution for $[M]^{5+}$. The mass resolution on the centre peaks is measured to be >25000, amply sufficient for unit resolution, however, the intensity of the centre peak (marked with an arrow) clearly is suppressed. The theoretical value for this centre peak was also calculated using the 'isotope distribution' facility in the Xmass™ software, developed by Van Orden based upon work by Rockwood *et al.* [23]. The exact masses in this program are taken from Audi and Wapstra [24] and the relative abundances of the isotopes are taken from the CRC Handbook of Chemistry and Physics [22]. The elemental composition of Zn_4SmtA^4 is determined using GP-MAW™.[5] A comparison of the

---

[4] Elemental composition for neutral Zn_4SmtA: $C_{225}H_{348}N_{70}O_{80}S_9Zn_4$

experimental (m/z 1173.62) and the theoretical (m/z 1173.61) value of the centre peak of the isotope distribution of $[M]^{5+}$ yields a mass error of < 5 ppm. Internal calibration of the Zn_4SmtA peaks with the bovine ubiquitin peaks can give good mass accuracy. Nevertheless, a distortion in the isotope distribution seen in the example of a suppressed centre peak for $[M]^{5+}$ can result in a misallocation of the mass by 1 Da. The relatively low magnetic field of 3 Tesla has a low tolerance of space charge effects in the cell and hence more precautions must be taken to minimise these effects.

5.4 Re-constituted $^{67}Zn_4SmtA$

Solution conditions were optimised to render the electrospray ionisation process less destructive. Zn_4SmtA was examined at a concentration of 10 μM together with bovine ubiquitin at a concentration of 5 μM in a 5 mM NH_4Ac water/methanol/formic acid (70:30:0.05, v/v/v) throughout these experiments. The ESI generated ions were accumulated in the hexapole for 2000-3000 ms and then transferred to the ICR cell. Initial trapping of the ions was effected at 3 V (rear trapping plate, PV2) and 2 V (front trapping plate, PV1), followed by dynamic ramping to 0.1 V on both trapping plates. These very low trapping plate voltages enable the ion cloud to occupy a larger space in the direction of the magnetic field (z direction), thereby reducing space charge effects in the cell. Conversely, the previous experiment (5.3 Preliminary Results) was conducted using static trapping potentials of 0.8 V. The population of ions in the cell was limited, again for the purpose of minimising space charge effects. 30 time-domain transients were co-averaged in order to reduce distortions in the isotope distribution due to statistical variations in the isotope abundances.

Figure 5.3 shows a mass spectrum of Zn_4SmtA spiked with ubiquitin standard run under the above solution conditions. Ionisation conditions appear to be gentle enough, since no losses were observed. The charge state distribution here (Figure 5.3) is narrower than previously (Figure 5.2) and additionally the species appear at lower charge states. The 4+ charge state, $[M]^{4+}$ (at m/z 1467), is the most abundant,

---

[5] http://welcome.to/gpmaw

which perhaps reflects the natural charge state of Zn_4SmtA in solution. At pH 5.5 (measured), the conditions were only mildly acidic compared to natural pH conditions (pH 7.4).

The 6+ charge state of ubiquitin, Ubi^{6+}, appearing at m/z 1428 was chosen as a suitable standard for calibration of $[M]^{4+}$. Care was taken to have similar peak intensities for Ubi^{6+} and $[M]^{4+}$ in the mass spectra (i.e. intensity of Ubi^{6+} < 3 times the intensity of $[M]^{4+}$), thereby achieving the best mass accuracy. If required, the ratio of Zn_4SmtA to bovine ubiquitin concentration was modified to account for this discrepancy.
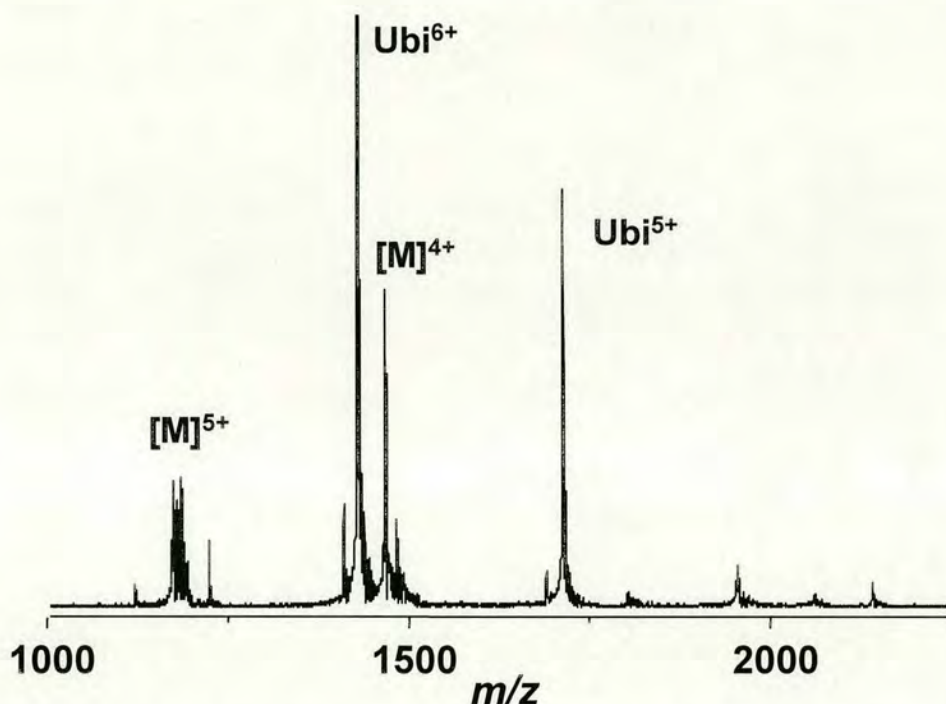
Figure 5.3 Raw ESI mass spectrum of Zn_4SmtA (10 μM) and bovine ubiquitin (5 μM) in 5 mM NH_4Ac water/methanol/formic acid (70:30:0.05, v/v/v). No losses are observed in this spectrum. The pH of the solution was measured to be 5.5.
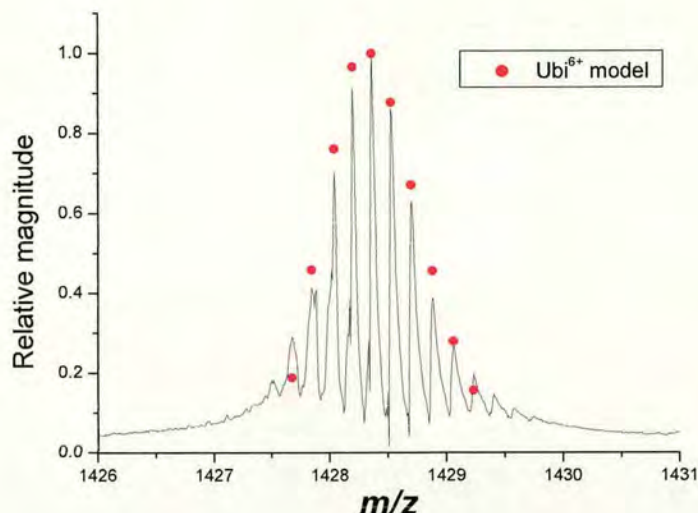
Figure 5.4 Isotopic distribution of the 6+ charge state of bovine ubiquitin, Ubi^{6+}, centred around m/z 1428. The theoretical isotope distribution calculated using the Xmass™ software is shown as red dots. The good fit between experiment and model validates the use of Ubi^{6+} as a suitable standard for $[M]^{4+}$.

The theoretical isotope distribution of Ubi^{6+} was again calculated using the Xmass™ software, as described previously, based on the elemental composition of bovine ubiquitin.[6] The isotopic distribution of Ubi^{6+} in Figure 5.4 shows little distortion, on the more intense isotope peaks. This confirms that Ubi^{6+} is an appropriate standard to check that no isotope distortion is present in the mass spectrum.

For the purpose of determining the mass increase that occurs when substituting the natural Zn isotope abundances in Zn_4SmtA with ^{67}Zn (93 % enriched), the mass raw spectra were deconvoluted to the neutral charge mass spectra. This was carried out with the deconvolution algorithm [25] in the Xmass™ software. The deconvolution algorithm assumes that charges on a protein are due to protonation (for cations) and to deprotonation (for anions). In a metalloprotein charge is also carried by the metal centres. The oxidation state of some metals (e.g. Fe) can vary, thus making the interpretation of their mass spectra more complicated [20]. In the example of Fe, the algorithm would underestimate the neutral charge

---

[6] Elemental composition of neutral bovine ubiquitin: $C_{378}H_{629}N_{105}O_{118}S_1$

mass for a Fe^{3+} containing metalloprotein by the mass of 3 protons, whereas for a Fe^{2+} containing protein the neutral mass would be underestimated by the mass of 2 protons, even if the protein itself is the same. In order to be consistent, Johnson *et al.* therefore defined the 'apparent mass' as the mass determined by the deconvolution algorithm, which then needs to be corrected by the charge carried by the metal centres (i.e. increased by the number of protons that the charge on the metal centres equals to). For Zn there is no ambiguity concerning the oxidation state, which is always Zn^{2+}. Hence there is no need to correct for the fact that charge is carried by the metal centres. In fact, in order to obtain the neutral charge species of Zn_4SmtA from for example $[M]^{4+}$, the mass of 4 protons is substracted, which is exactly what the deconvolution algorithm does. It is for this reason that the 4+ charge state is labelled $[M]^{4+}$ and not $[M+4H]^{4+}$. The neutral elemental composition for Zn_4SmtA^4 is determined from the amino acid sequence of SmtA (with the acidic groups protonated and the basic groups deprotonated – i.e. neutral); the mass of 4 Zn atoms is added and 8 hydrogens are removed since Zn (Zn^{2+}) is the charge carrier.

The deconvoluted neutral mass spectra for Zn_4SmtA with a natural isotope distribution (Figure 5.5 A) and the reconstituted $^{67}Zn_4SmtA$ (Figure 5.5 B) show that the most intense isotope peak increases from 5862.95 Da (Figure 5.5 A) to 5868.81 Da (Figure 5.5 B). This compares to calculated theoretical values of 5863.00 Da and 5869.00 Da, deviations of 8.5 and 32 ppm respectively. The substitution of all 4 Zn by ^{67}Zn (93% enriched) causes a mass increase of the most intense peak by 6 Da and the isotope distribution becomes more narrow, due to a decrease in the abundances of the various Zn isotopes. The theoretical isotope distributions are in fairly good agreement with the experimental ones. The most intense peaks derived experimentally and from the model match. Some distortion in the isotope distribution is observed and deviation from model to experiment is generally more pronounced for the lower intensity isotope peaks. This is also the case for Ubi^{6+} in Figure 5.4.
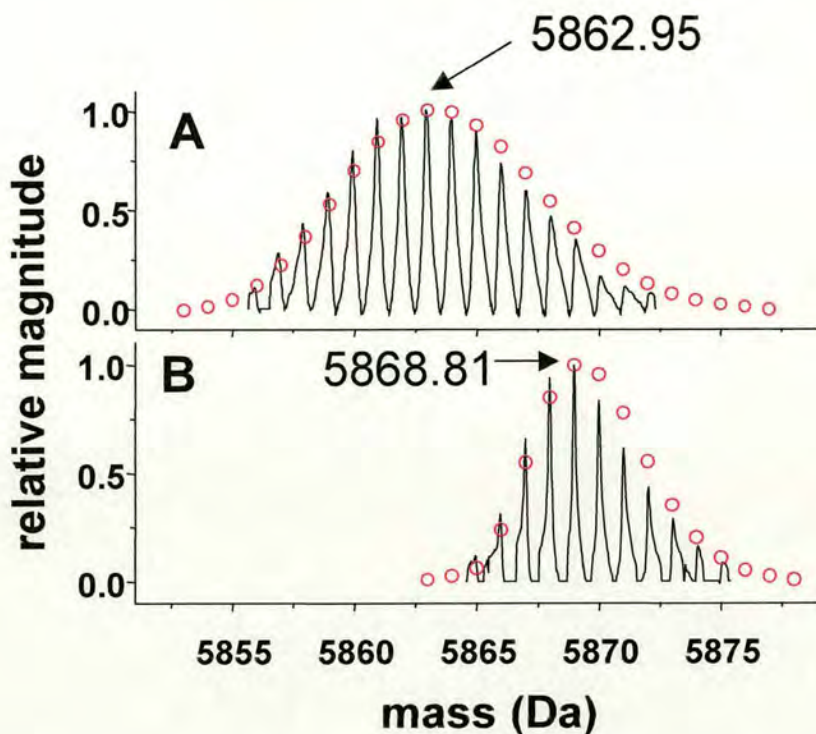
Figure 5.5 Deconvoluted neutral charge mass ESI mass spectra and modelled isotope distributions (red circles) of (A) natural abundance Zn_4SmtA and (B) 93% enriched $^{67}Zn_4SmtA$.

5.5 Zn isotope exchange

Given that the exchange of all 4 Zn atoms in Zn_4SmtA gives rise to such a pronounced mass increase, it should be possible to determine the number of Zn atoms exchanged at any point in between 0 and 4 and hence to investigate the kinetic exchange of Zn in this protein. Zn_4SmtA was incubated with $^{67}ZnCl_2$ (93% ^{67}Zn enriched) for different time intervals (at 310 K) and to then determine which of the modelled isotope envelopes for $Zn_{4-x}^{67}Zn_xSmtA$ (x = 1-4, in 0.25 intervals) fitted the experimental data best (see publication, Appendix 5).

When comparing the experimental to the theoretical isotope distribution, one could compare the most intense isotope peaks, as was done above for $[M]^{5+}$ in Figure 5.2. However, there is a substantial risk that a distortion in the isotope envelope could lead to a mislabelling of the mass. This would especially be the case for higher

mass compounds (> 15 kDa), where the isotope envelopes are much wider and the most abundant isotope peak is not readily identifiable. Instead, Senko *et al.* proposed the 'chi squared test', which takes the whole isotope distribution into account and statistically compares it to the theoretical one [17]. He *et al.* proposed a method which uses the standard deviation as measure of how well the experimental and theoretical isotope distributions match [19]. The 'standard deviation method' has the advantage that it is easier to compare the quality of fits for different charge states.

Here the 'standard deviation method' is used as well, since it is easy to implement and allows a simple comparison between different fits. By statistically fitting the experimental to the theoretical isotope distributions, the number of ^{67}Zn exchanged *on average* can be determined much more accurately. To maximise the accuracy of this fitting, only the 7 most abundant isotope peaks are included in this analysis, since lower intensity isotope peaks are more likely to deviate from the model and would thus give rise to distortions. The 4+ charge state of Zn_4SmtA, $[M]^{4+}$ is chosen as the model system, given that it is the most abundant charge state in the 'soft' ionisation conditions used and therefore probably the best representation of the charge state in solution. As stated, the internal standard Ubi^{6+} is very close in mass-to-charge and hence one can assume that if little isotope distribution distortion is observed for Ubi^{6+}, the same would hold true for $[M]^{4+}$.

The procedure for calculating the relative isotope abundances for a different number of x in $Zn_{4-x}{}^{67}Zn_xSmtA$, which is the number of Zn atoms that have exchanged *on average*, is shown in Table 5.1. In this procedure the assumption is made that only ^{67}Zn from the 10-fold excess $^{67}ZnCl_2$ exchanges for Zn atoms in the metallothionein, since for example the exchange of ^{64}Zn from the solution for any other isotope in the protein is negligible. The relative abundance of each Zn isotope for different values of x is then used to generate a theoretical isotope distribution. The most abundant isotope peak of the experimental data and of the theoretical distribution are normalised to 1. Standard deviations of the 7 most intense isotope peaks are calculated for various $Zn_{4-x}{}^{67}Zn_xSmtA$ possibilities (in x=0.25 intervals) (see Figure 5.6). A least squares second order polynomial fitting is carried out to determine the minimum standard deviation which was retained as the best fit. Figure 5.7 shows the isotope distributions of $[M]^{4+}$ for different time points and the

corresponding best fits (nearest by 0.25 ^{67}Zn intervals) as determined by the 'standard deviation method'.
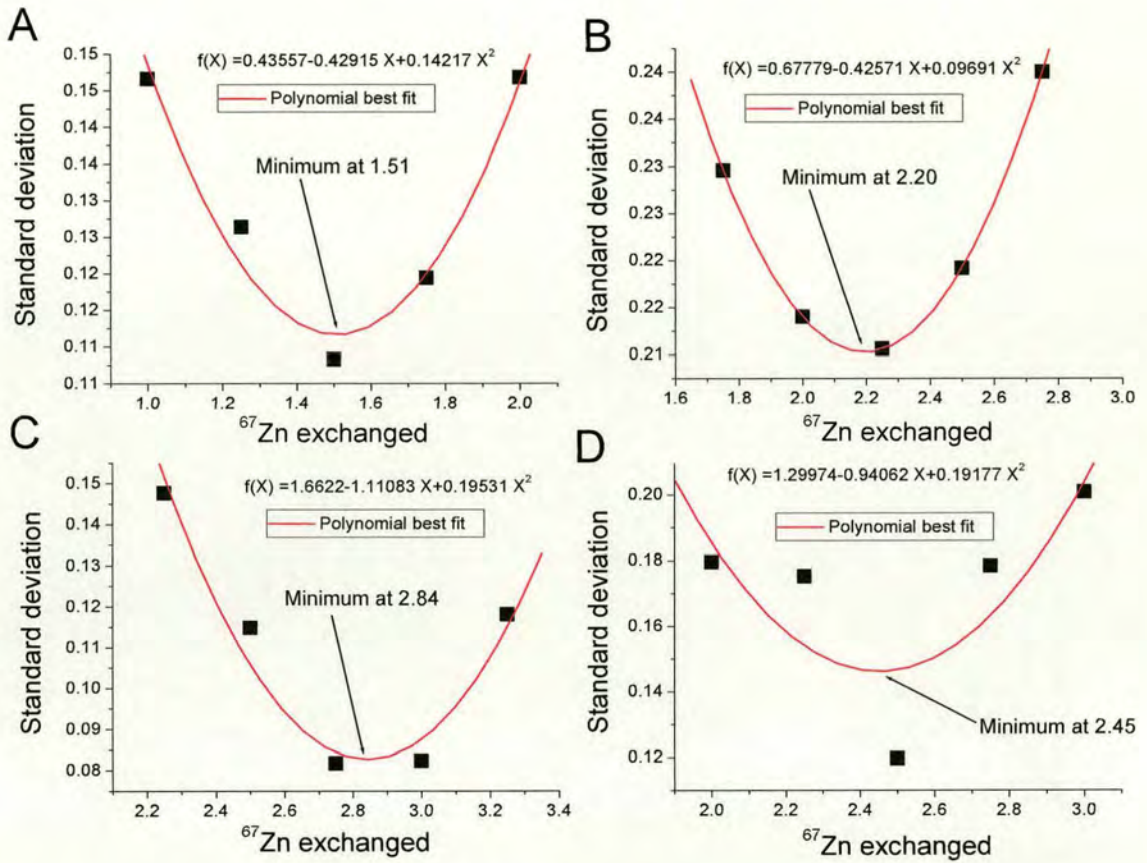

Figure 5.6 Fitting of the number of ^{67}Zn exchanged on average (x) for different time points using the 'standard deviation method'. (A) 1 hour, (B) 7.5 hours, (C) 99 hours and (D) 168 hours. The theoretical isotope distribution for different values of x in Zn$_{4-x}$^{67}Zn$_x$SmtA (in x=0.25 intervals) was compared to the experimental isotope distribution measured at a particular time point of incubation by calculating the standard deviation for the 7 most intense isotope peaks. A least squares second order polynomial fit was performed to determine the point of ^{67}Zn exchanged where the standard deviation is minimised for each time point. The expression of each polynomial best fit is given, from which the minimum was determined analytically.

No. Zn exchanged	Zn isotope	Normal Zn distribution	Zn exchanged	Zn not exchanged	Zn site not exchanging	Zn isotope abundances
1	64	0.486	0.11312865	0.00837135	0.3645	0.37287135
	66	0.279	0.064944225	0.004805775	0.20925	0.214055775
	67	0.041	0.009543775	0.000706225	0.03075	0.264231225
	68	0.188	0.0437617	0.0032383	0.141	0.1442383
	70	0.006	0.00139665	0.00010335	0.0045	0.00460335
		sum	0.232775		check	1

Table 5.1 Screen-shot of MS Excel™ script used to calculate Zn isotope abundances for value of x equal to 1 in $Zn_{4-x}{}^{67}Zn_xSmtA$. The heading 'Normal Zn distribution' shows the natural abundance of each isotope. In the heading 'Zn exchanged', of the 1 Zn out of 4 that is exchanged *on average*, the amount of each isotope that is exchanged for ^{67}Zn is calculated (e.g. for ^{64}Zn: (1/4)*0.486*0.9311=0.113). The sum of all Zn that has exchanged for ^{67}Zn is calculated. In the heading 'Zn not exchanged', the amount of Zn (of the 1 Zn out of 4 *on average*) that has not exchanged for ^{67}Zn is calculated (e.g. for ^{64}Zn: (1/4)*0.486-0.113=0.008) (This arises since ^{67}Zn is only enriched to 93.11%, <u>not</u> 100%). In the heading 'Zn site not exchanging', the amount of each Zn isotope that is not involved in ^{67}Zn exchange is calculated (e.g. for ^{64}Zn: (1-(1/4))*0.486=0.3645). In the heading 'Zn isotope abundances', the relative amount of each isotope is given for any value of x. For ^{64}Zn, this means summing the values in columns 'Zn not exchanged' and 'Zn site not exchanging'. For ^{67}Zn these two columns are summed, as well as the total amount of Zn that has exchanged for ^{67}Zn (sum = 0.233). The check box should always have the value of 1 to indicate that the relative amounts of Zn isotopes add up correctly.

For the time points of 1 hour (Figure 5.7 A) and especially for 7.5 hours incubation (Figure 5.7 B) the best fits are not centred around the experimental isotope distribution. It appears that the 'standard deviation method' has not given the best fit for these distributions. Visually fitting the isotope distributions could have led to a more accurate result! The reason for this poor fitting using the 'standard deviation method' is that the abundances of the isotope peaks in the experimental isotope distribution are suppressed relative to the most intense isotope peak; this is particularly the case for Figure 5.7 B. As a result of this, the sum of the differences squared, which is calculated in the 'standard deviation method', can be minimised, even though the theoretical isotope distribution is not symmetrically centred around the experimental one.
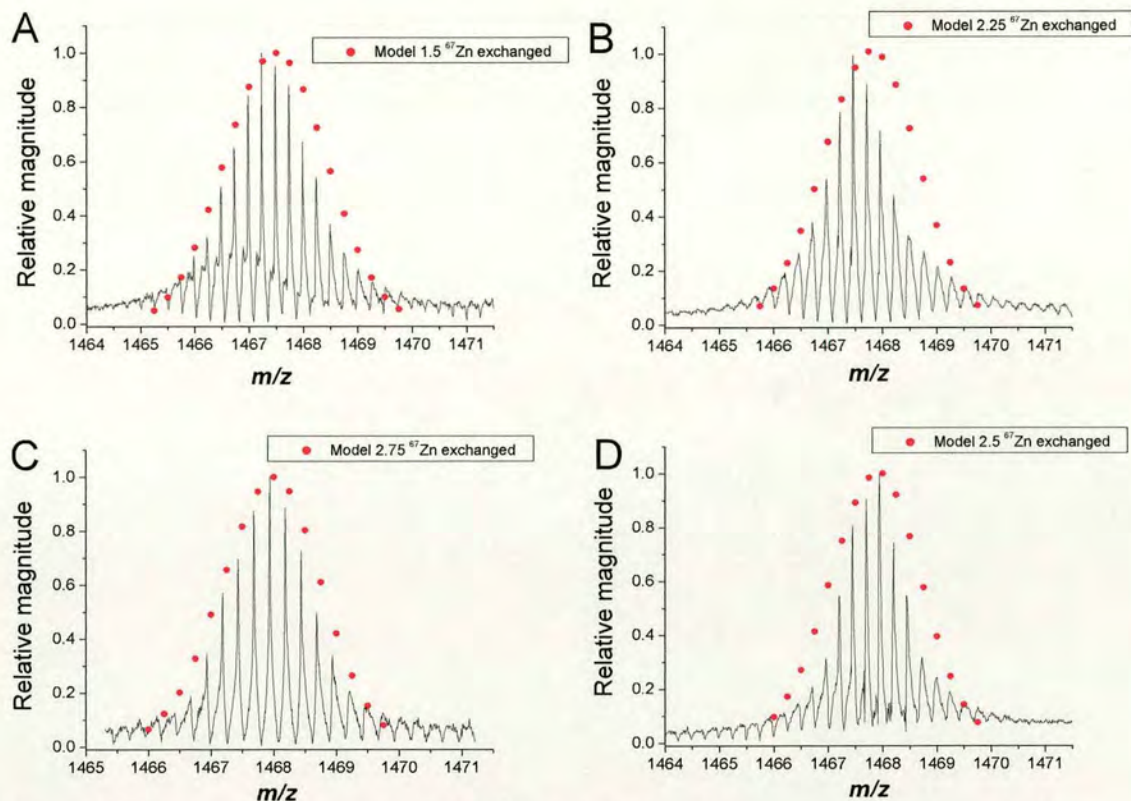
Figure 5.7 Isotope distributions of $[M]^{4+}$ for different time points, and corresponding nearest best theoretical fits (by 0.25 ^{67}Zn interval) determined by 'standard deviation method' for (A) 1 hour, (B) 7.5 hours, (C) 99 hours and (D) 168 hours.

Hence, it was decided to employ a different method of statistical fitting, the 'statistical deviation method' which is a measure of the degree of symmetry between the experimental and theoretical isotope distributions (see Figure 5.8). The minimum 'symmetry deviation' was again determined by a least squares second order polynomial fitting procedure. The nearest best fits (by 0.25 ^{67}Zn intervals) as determined by this method are shown in Figure 5.9.
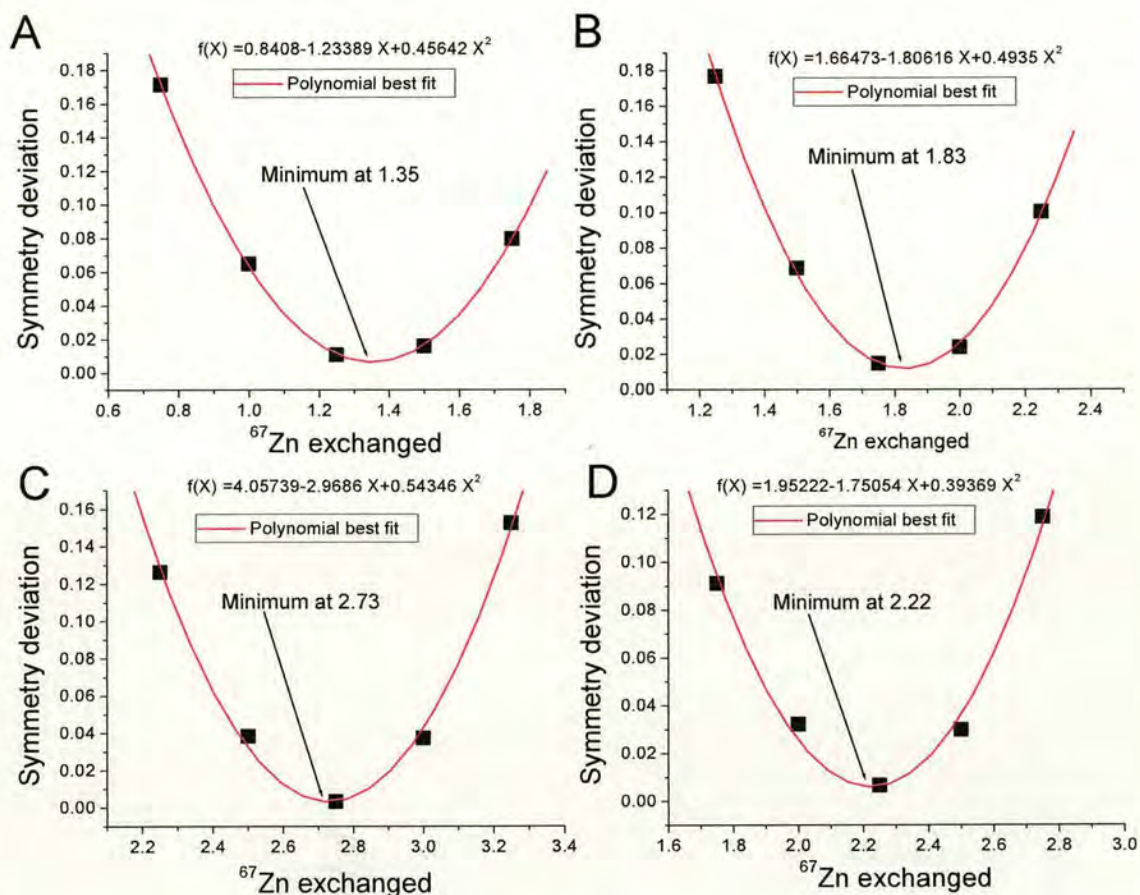
Figure 5.8 Fitting of the number of ^{67}Zn exchanged on average (x) for different time points using the 'symmetry deviation method'. (A) 1 hour, (B) 7.5 hours, (C) 99 hours and (D) 168 hours. The theoretical isotope distribution for different values of x in $Zn_{4-x}{}^{67}Zn_xSmtA$ (in 0.25 intervals) was compared to the experimental isotope distribution measured at a particular time point of incubation by calculating the symmetry deviation for the 7 most intense isotope peaks. The symmetry deviation consists in the following expression for the 7 most intense peaks: $\sigma_S = [(I_{model,1} - I_{exp,1}) - (I_{model,7} - I_{exp,7})]^2 + [(I_{model,2} - I_{exp,2}) - (I_{model,6} - I_{exp,6})]^2 + [(I_{model,3} - I_{exp,3}) - (I_{model,5} - I_{exp,5})]^2 + [(I_{model,4} - I_{exp,4})]^2$, where $I_{model,1}$ is the normalised intensity of peak number 1 of the modelled isotope distribution and $I_{exp,1}$ is the normalised intensity of peak number 1 in the experimental isotope distribution. When the symmetry deviation is minimised the theoretical fit is centred around the experimental isotope distribution. A least squares second order polynomial fit was carried out to determine the minimum symmetry deviation.

Not surprisingly, the theoretical fits in Figure 5.9 are centred around the experimental isotope distributions completely symmetrically. The advantage of the 'statistical deviation method', as opposed to the 'standard deviation method', is that a distortion in the isotope distribution such as a suppression of the lower isotope peaks

is less likely to result in an inaccurate fitting. For lower magnetic fields, such as our 3 Tesla magnet, distortions in the isotope distributions are much more common than for high-field magnets (> 7 T). Distortions in the isotope distribution were minimal for the work by He *et al.* on their 9.4 Tesla instrument [19]. The fitting for 168 hours incubation (Figure 5.7 D) gave a best fit of 2.25 ^{67}Zn exchanged. The most intense isotope peaks of experiment and model do not coincide, which may be due to a suppression of the isotope peaks on the right-hand side of the isotope envelope. Given the possibility of such a suppression effect, this implies that the extent of ^{67}Zn exchanged at any time point can only be determined to an accuracy of about ± 0.25 ^{67}Zn.
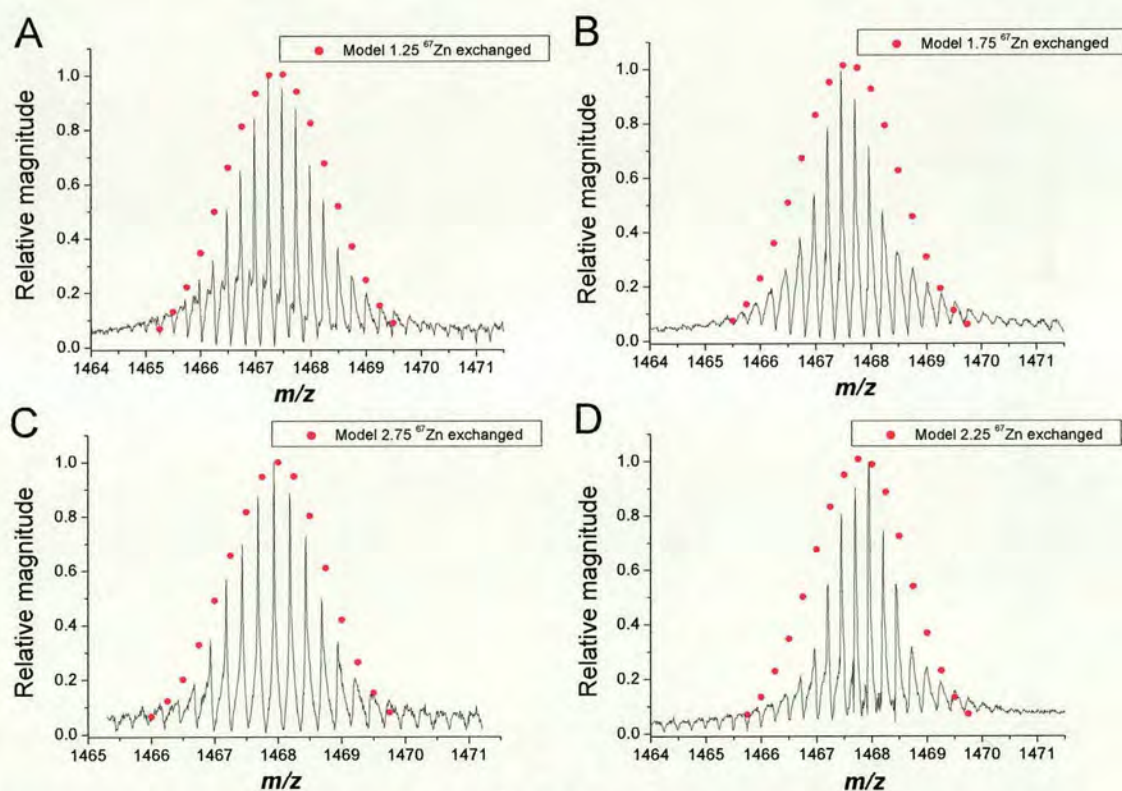
Figure 5.9 Isotope distributions of [M]$^{4+}$ for different time points, and corresponding nearest best theoretical fits (by 0.25 ^{67}Zn interval) determined by 'symmetry deviation method' for (A) 1 hour, (B) 7.5 hours, (C) 99 hours and (D) 168 hours.

The best exact fits for the different time points as determined by the 'symmetry deviation method' are summarised in Figure 5.10. Following an initial fast exchange of about 1.35 Zn after 1 hour, the exchange kinetics progressively decrease. A maximum incorporation of 2.73 ^{67}Zn is observed after 99 hours and no further increase occurs after this (value for 168 hours lower, but within experimental error). In terms of the kinetics of the four zinc sites, this means that one of the Zn atoms exchanges quickly (< 1 hour) and a second Zn is exchanged over a period of ~ 10 hours. A third Zn atom is exchanged much more slowly (100 hours), whereas the fourth Zn atom does not exchange over the period of one week. The reason for the lower measured incorporation of ^{67}Zn for 168 hours incorporation probably also has to do with the unusually large extent of Na adduct formation in the mass spectrum (see Figure 5.11). Possibly, the Na$^+$ concentration in solution increased over time, thereby affecting the Zn exchange study. Given that the 93% enriched ^{67}Zn medium is at a 10-fold excess, the maximum number of Zn atoms that can exchange when 3 of the Zn sites exchange is a value of x=2.7. Hence, it appears that at maximum 3 of the 4 Zn sites are able to exchange for Zn in solution, suggesting that one of Zn sites is kinetically inert.
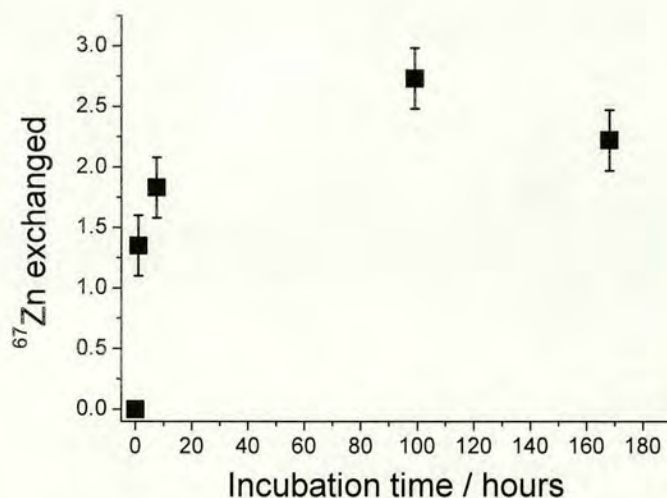
Figure 5.10 Time course of the reaction of Zn$_4$SmtA with a 10-fold excess of ^{67}ZnCl$_2$ showing the number of ^{67}Zn atoms exchanged on average (x). Error bars x ±0.25 ^{67}Zn.
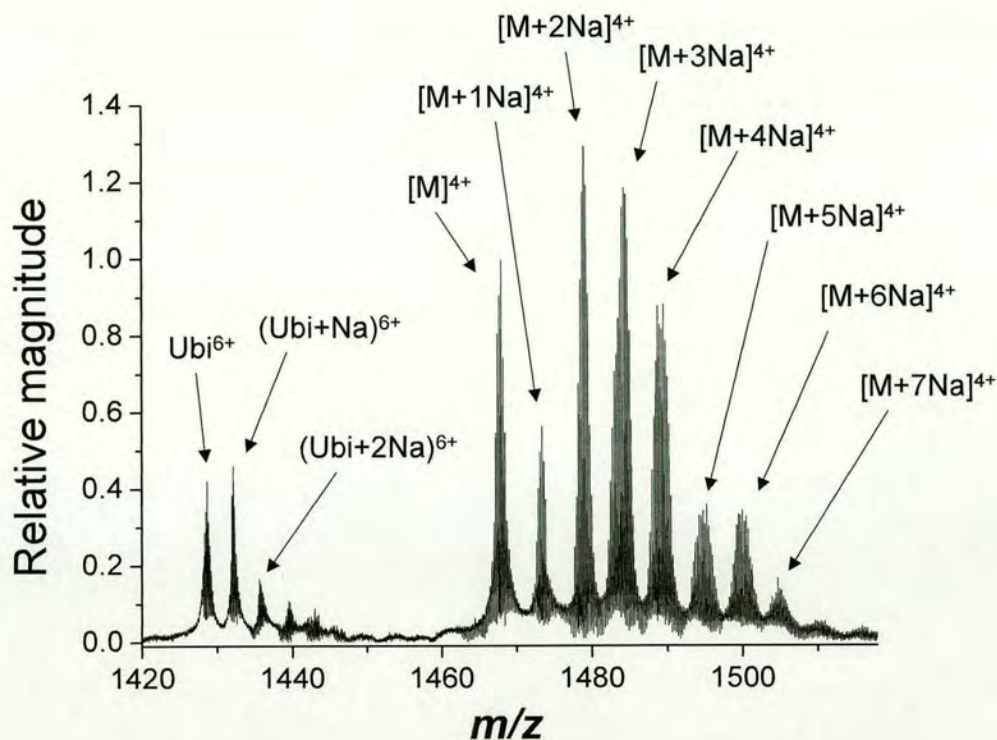
Figure 5.11 Raw ESI mass spectrum for Zn$_4$SmtA incubated for 168 hours showing high extent of Na adduct formation.

Previous NMR experiments of the reaction of ^{111}Cd^{2+} with Zn$_4$SmtA had shown that Zn site A is kinetically inert [11]. The space-filling representation in Figure 5.12 shows that whereas site B-D are surface accessible, site A is completely buried inside the molecule.
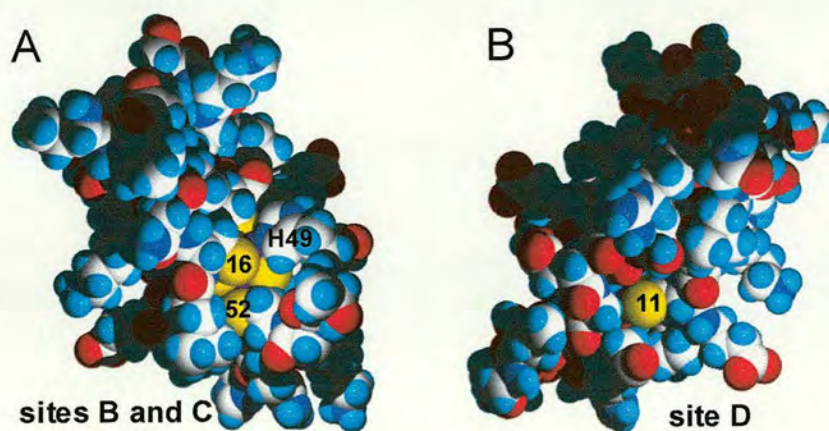

Figure 5.12 Space-filling model showing availability of Zn sites B, C and D. Site A (not shown) is buried inside the molecule.

It appears that the elements of secondary structure surrounding site A render it unavailable for exchange. This occurs in spite of the fact that site A is bound to 4 cysteines; in most cases, all cysteine binding sites exhibit a much stronger binding to Cd^{2+} than Zn^{2+}, and complete substitution occurs [26,27]. This Zn isotope exchange study has confirmed that one of the Zn sites is kinetically inert. Importantly, FT-ICR MS allows direct determination of the metal-protein ratio, while at the same time conserving the metal –protein complex intact. Alternative techniques, such as radioisotope exchange require a separate detection of the metal and protein. Zinc as an element is not readily detected by other techniques. The d^{10} electronic configuration of Zn^{2+} precludes d-d excitations and therefore makes electronic spectroscopy of Zn^{2+} less informative. Since all electrons are paired, EPR cannot be used to investigate Zn^{2+}. The NMR spectroscopic properties of zinc are highly unfavourable: low magnetic moment, implying low sensitivity, modest resonance frequency and a nuclear spin of 5/2, implying a high quadrupole moment. Usually, zinc metalloproteins have been studied by substituting Zn for a 'spin-spy' (typically ^{113}Cd) with more favourable spectroscopic properties [28,29]. It is important to note that NMR cannot directly observe the native Zn metalloproteins. Hence, the contribution of mass spectrometry in the study of Zn metalloproteins should not be underestimated.

Given the wide range of transition metal isotopes, this metal exchange study could be extended to a variety of other metal-protein systems. The high mass resolution and accuracy of FT-ICR MS enable an accurate determination of the extent of isotope exchange. Higher field magnets, where distortion of the isotope distribution is less of a problem, can play an important role in metal exchange studies of proteins. Further, the rapidity of mass spectrometric analysis would allow the kinetic study of this exchange on a much faster time scale (i.e. minutes), thus approaching faster biological time scales.

References

(1)   Klug, A.; Rhodes, D. *Trends Biochem. Sci.* **1987**, *12*, 464-469.

(2)   Udom, A. O.; Brady, F. O. *Proc. North Dakota Acad. Sci.* **1979**, *33*, 5.

(3)   Coyle, P.; Philcox, J. C.; Rofe, A. M. *Div. Clin. Biochem.* **1995**, *309*, 25-31.

(4)   Hill, H. A. O., Sadler, P. J., Thompson, A. J. *Metal Sites in Proteins and Models*; Springer-Verlag: Berlin, Heidelberg, 1999.

(5)   Babini, E.; Borsari, M.; Capozzi, F.; Eltis, L. D.; Luchinat, C. J. *J. Biol. Inorg. Chem.* **1999**, *4*, 692-700.

(6)   Gray, H. B.; Malmstrom, B. G.; Williams, R. J. P. *J. Biol. Inorg. Chem.* **2000**, *5*, 551-559.

(7)   O'Halloran, T. V.; Culotta, V. C. J. *J. Biol. Chem.* **2000**, *275*, 25057-25060.

(8)   Cobine, P. A. G., G. N.; Jones, C. E.; Wickramasinghe, W. A.; Solioz, M.; Dameron, C. T. *Biochemistry* **2002**, *41*, 5822-5829.

(9)   Romero-Isart, N.; Vasak, M. J. *J. Inorg. Biochem.* **2002**, *88*, 388-396.

(10)   Blindauer, C. A. H. M., D.; Robinson A. K.; Parkinson, J. A.; Bowness, P. W.; Sadler Peter, J.; Robinson Nigel, J. *J. Molec. Microbiol.* **2002**, *45*, 1421-1432.

(11)   Blindauer, C. A.; Harrison, M. D.; Parkinson, J. A.; Robinson, A. K.; Cavet, J. S.; Robinson, N. J.; Sadler, P. J. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9593-9598.

(12)   Gehrig, P. M.; You, C.; Dallinger, R.; Gruber, C.; Brouwer, M.; Kagi, J. H. R.; Hunziker, P. E. *Prot. Sci.* **2000**, *9*, 395-402.

(13)   Hathout, Y.; Fabris, D.; Fenselau, C. *Int. J. Mass Spectrom.* **2001**, *204*, 1-6.

(14)   Jensen, L. T.; Peltier, J. M.; Winge, D. R. *J. Biol. Inorg. Chem.* **1998**, *3*, 627-631.

(15)   Merrifield, M. E.; Huang, Z.; Kille, P.; Stillman, M. J. *J. Inorg. Biochem.* **2002**, *88*, 153-172.

(16)   Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325-1337.

(17)   Senko, M. W.; Beu, S. C.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 229-233.

(18)   Johnson, K. A.; Shira, B. A.; Anderson, J. L.; Amster, I. J. *Anal. Chem.* **2001**, *73*, 803-808.

(19)   He, F.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 120-126.

(20)   Johnson, K. A.; Verhagen, M. F. J. M.; Brereton, P. S.; Adams, M. W. W.; Amster, I. J. *Anal. Chem.* **2000**, *72*, 1410-1418.

(21)   Vasak, M. *Methods Enzymol.* **1991**, *205*, 452-458.

(22)   *CRC Handbook of Chemistry and Physics*; CRC Press Inc.: Boca Raton, USA, 1991; Vol. 71.

(23)   Rockwood, A. L.; Van Orden, S. L.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 54-59.

(24)   Audi, G.; Wapstra, A. H. *Nucl. Phys.* **1995**, *595*, 409-480.

(25)   Zhang, Z.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 225-233.

(26)   Krizek, B. A.; Merkle, D. L.; Berg, J. M. *Inorg. Chem.* **1993**, *32*, 937-940.

(27)   Kägi, J. H. R. In *Metallothionein III*; Suzuki, K. T., Imura, N., Kimura, M., Eds.; Birkhäuser Verlag: Basel, Boston, Berlin, 1993; pp 29-56.

(28)   Vallee, B. L.; Galdes, A.; Auld, D. S.; Riordan, J. F. In *Metal Ions in Biology*; Spiro, T. G., Ed.; John Wiley & Sons: New York, 1983; Vol. 5, pp 25-75.

(29)   McAteer, K.; Lipton, A. S.; Ellis, P. D. In *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; John Wiley & Sons: New York, 1996; Vol. 2, pp 1085-1091.

Chapter 6
Further avenues in the study of the ECD mechanism

6.1 Introduction

The study of electron capture dissociation in Chapter 3 limited itself to particular aspects such as neutral side-chain losses, the ease of hydrogen atom transfer from the protonated site, 'normal' alpha-cleavage and 'reverse' alpha-cleavage of the backbone and the relative abundances of c and z fragments. These results were attempted to be rationalised in terms of a *non-ergodic* mechanism in ECD and the proposed *hot hydrogen re-arrangment* mechanism [1]. Chapter 4 showed putative gas-phase structures of some of the peptides dissociated with ECD, which were modelled using the Amber molecular mechanics force field [2,3]. In some cases ([LHRH_HR]$^{2+}$, [BOMB_NH$_2$]$^{2+}$) a correlation seemed to exist between preferential hydrogen bonding of the electron-capturing hydrogen to a particular backbone carbonyl (as predicted by the low-energy structures) and higher ECD cleavage adjacent to that carbonyl. Conversely, for [LHRH_HR_NH$_2$]$^{2+}$ no such correlation was found for the low-energy structures.

In this Chapter, ECD mass spectra of some synthetic peptides[1] are shown, which were generated on a 9.4 T APEX III FT-ICR mass spectrometer equipped with an indirectly heated dispenser cathode (see Chapter 2, section 2.3.3). In this light the neutral losses will be re-visited (section 6.2.1). The link between the fixation of the charge and the resulting fragments is further studied: a truly 'fixed' charge group peptide is subjected to ECD (section 6.2.2) and a small peptide with no arginine residues is irradiated with electrons (section 6.2.3).

The second part of this Chapter (section 6.3) assumes a much slower time-scale of the ECD mechanism. In Chapter 4 the assumption was made that the ECD mechanism is ultra-fast (< ps) and that therefore the gas-phase structure of the ion

---

[1] These peptides were obtained from the group of Prof. Vicki Wysocki from the University of Arizona (Tucson, Arizona, USA).

prior to electron capture determines the dissociation pattern. While electron capture itself is likely to be a very fast process (< ps), the intra-molecular hydrogen atom transfer could take place on a much slower time-scale. The gas-phase structure of an ion that has captured an electron on the arginine side-chain is attempted using the Amber molecular mechanics force field (section 6.3).

6.2 ECD of synthetic peptides

6.2.1 Another look at neutral losses

The doubly protonated form of the peptide RVYIH(Bz)PF (for structure of H(Bz) see Figure 6.1) was irradiated with electrons giving the mass spectrum in Figure 6.2. The insert in Figure 6.2 shows a high neutral loss of 91.053 Da from the reduced species $[M+2H]^{+\bullet}$, which matches the mass of $C_7H_7^{\bullet}$ (mass =91.054 Da). The loss of this group can be associated with the benzyl group on the histidine side-chain. The neutral fragment $C_7H_7^{\bullet}$ does not have a hydrogen atom incorporated. The loss of the whole histidine(Bz) side-chain ($mass_{theoretical}$ = 173.108 Da, $mass_{observed}$ = 173.121 Da) is of low abundance. Further, the loss of 44 Da linked with the arginine side-chain is of low abundance. These results suggest that:

- either the hydrogen atom is transferred to N_τ on the histidine side-chain giving rise to radical cleavage of the benzyl group (see Figure 6.3)

- or that N_τ is already protonated

- or that this radical cleavage occurs without hydrogen atom re-arrangement

The first hypothesis could be tested by determining the hydrogen atom affinity of the nitrogen atom on the histidine side-chain by *ab initio* calculations. If the hydrogen atom affinity of N_τ is high, then the mechanism shown in Figure 6.3 is possible. The abundant loss of $C_7H_7^{\bullet}$ is likely to be due to resonance stabilisation of this fragment (see Figure 6.3).

Another feature in this mass spectrum is the appearance of an a_5^+ fragment ion. The corresponding c_5^+ fragment cannot be observed due to the proline residue C-

214

terminal to H(Bz). No other *a* fragments are seen in this spectrum. Thus, it appears as if the molecule somehow 'knew' that it had to cleave the C_α-C bond (to give *a* fragment) as opposed to the N-C_α bond (to give c fragment) to result in backbone fragments. The labelled $(a_5-17)^+$ peak could correspond to an a_5^+ fragment that has lost its N-terminal NH_3. This suggests that two cleavages result from one electron capture, an observation that has also been made by O'Connor and co-workers [4].
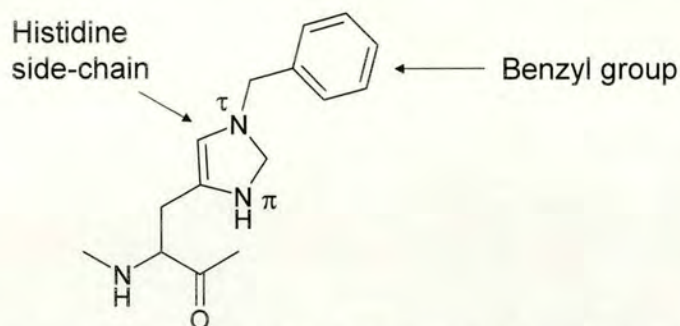
Figure 6.1 Structure (stick structure) of histidine(Bz) residue, showing N_π and N_τ, which carries an additional benzyl group.
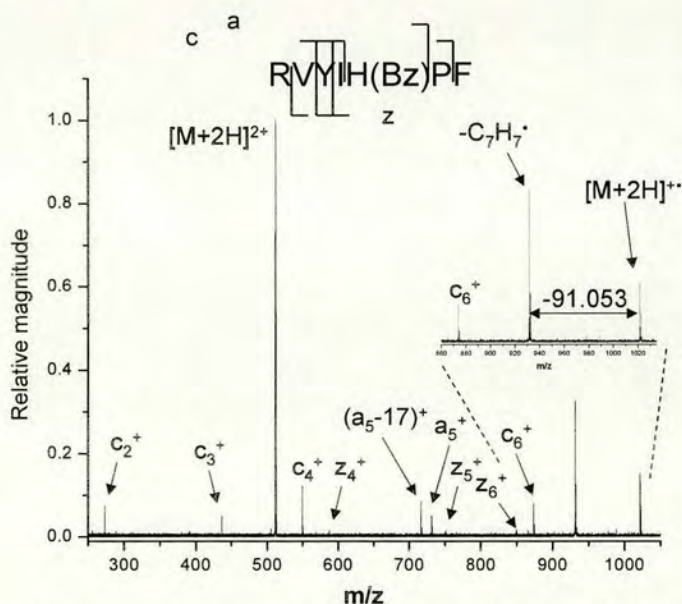
Figure 6.2 ECD mass spectrum of doubly protonated peptide RVYIH(Bz)PF. (200 time-domain transients, hexapole accumulation = 0.6 s, electron irradiation = 25 ms).
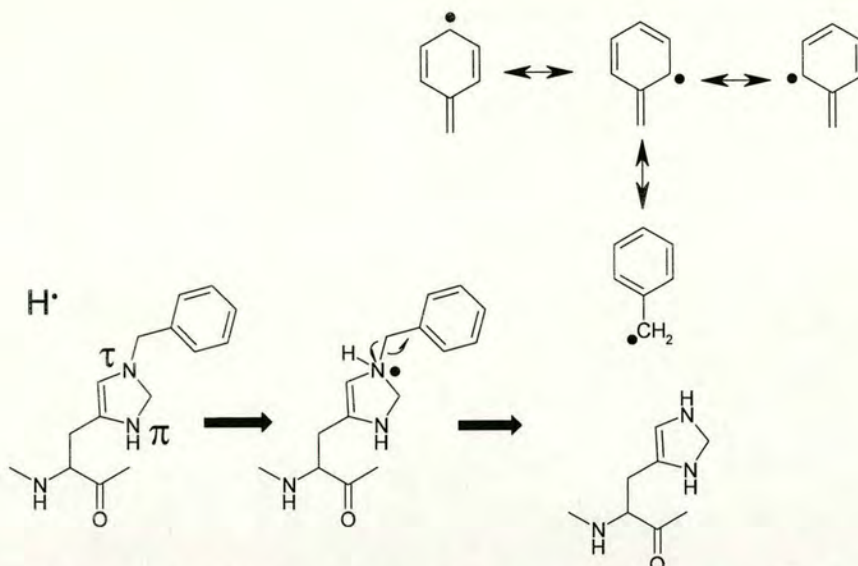
Figure 6.3 Schematic of proposed mechanism of hydrogen atom transfer to N_τ on the histidine side-chain, inducing radical cleavage of benzyl group. This fragment is likely to be energetically favoured due to resonance stabilisation (Note that arrows refer to single-electron movement).

6.2.2 ECD of peptide with fixed charge group

The doubly charged (but only singly protonated) peptide $\Phi_3P^+CH_2C(O)$-VYIHPF (for structure of fixed charge group $\Phi_3P^+CH_2C(O)$- see Figure 6.4) was irradiated with electrons to give the ECD mass spectrum in Figure 6.5. As expected, all charged backbone fragments include the $\Phi_3P^+CH_2C(O)$- fixed charge group, which was also observed in collisionally-induced dissociation (CID) of this peptide [5]. The most likely site of protonation is the side-chain of histidine, which is further supported by abundant loss of 82 Da.
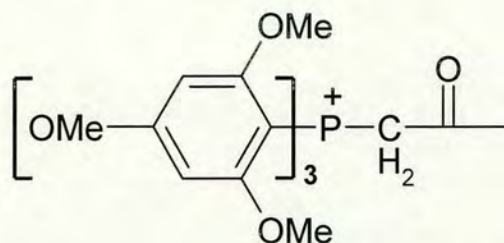
Figure 6.4 Structure (stick structure) of $\Phi_3P^+CH_2C(O)$- group, showing (nominal) positive charge on phosphorous atom.
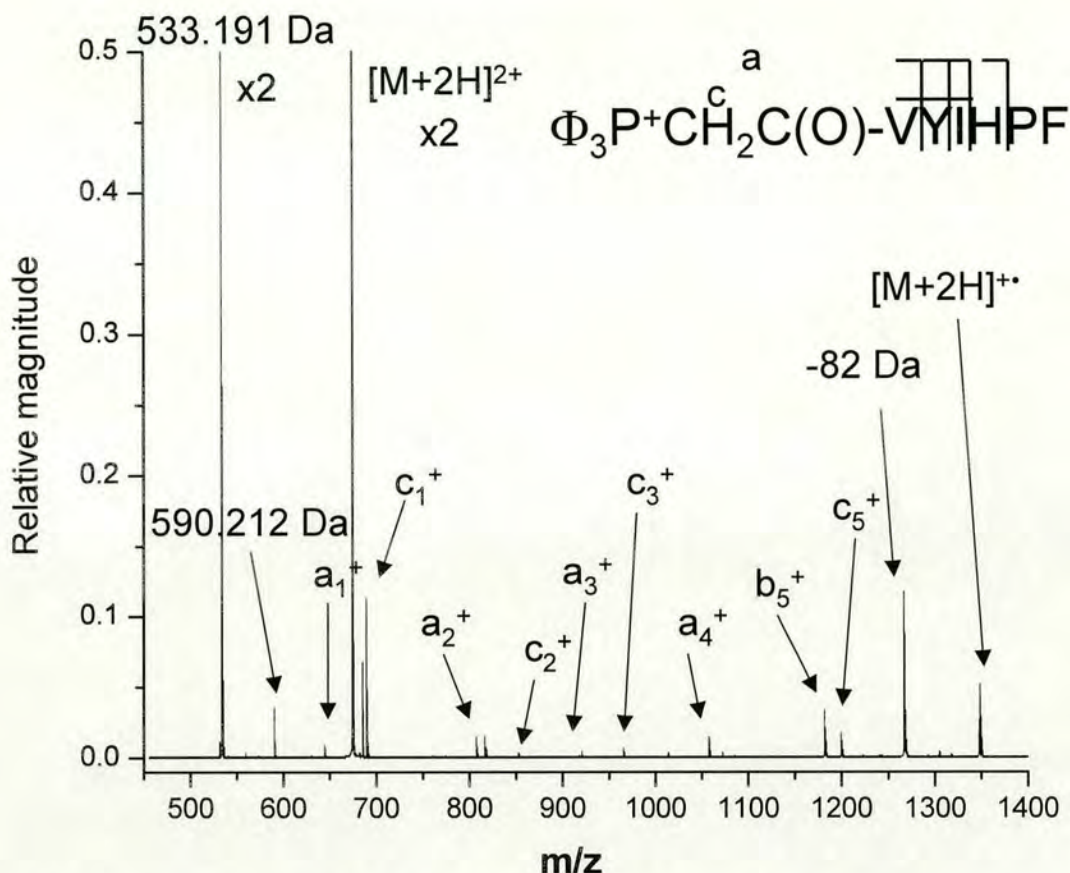
Figure 6.5 ECD mass spectrum of the doubly protonated peptide $\Phi_3P^+CH_2C(O)$-VYIHPF. (200 time-domain transients, hexapole accumulation = 0.6 sec, electron irradiation = 30 ms).

The peaks labelled 533.191 Da and 590.212 Da are lower in mass than the fragment ion a_1^+. One would assume that the fixed charge group is included in both of these fragments. In this light the elemental formula matches for these are $C_{27}H_{34}O_9P_1^+$ (monoisotopic mass = 533.1935 Da) and $C_{29}H_{37}O_{10}N_1P_1^+$ (monoisotopic mass = 590.2150 Da). For the fragment at m/z 533.191 this implies hydrogen atom incorporation into the fragment. Figure 6.6 shows a possible mechanism that could explain hydrogen atom incorporation into this fragment. The fact that the peak at m/z 533.191 Da is by far the highest abundance fragment implies that hydrogen atom transfer to this group is particularly efficient. To test this assumption the hydrogen atom affinity of P^+ would need to be calculated.

The peak at m/z 590.212 can be rationalised by hydrogen atom transfer to the carbonyl group of the fixed charge group, giving rise to a c-type fragment.

These results show that for a fixed charge group in a doubly charged peptide all charged ECD fragments include the fixed charge group. The very high abundance of the fragment at m/z 533.191 can be rationalised by hydrogen atom to the phosphorous atom on the fixed charge group as shown in Figure 6.6. These results appear to imply that *hydrogen atom re-arrangement* in the ECD mechanism does take place, since it would be difficult to explain the incorporation of a hydrogen atom into this fragment otherwise.
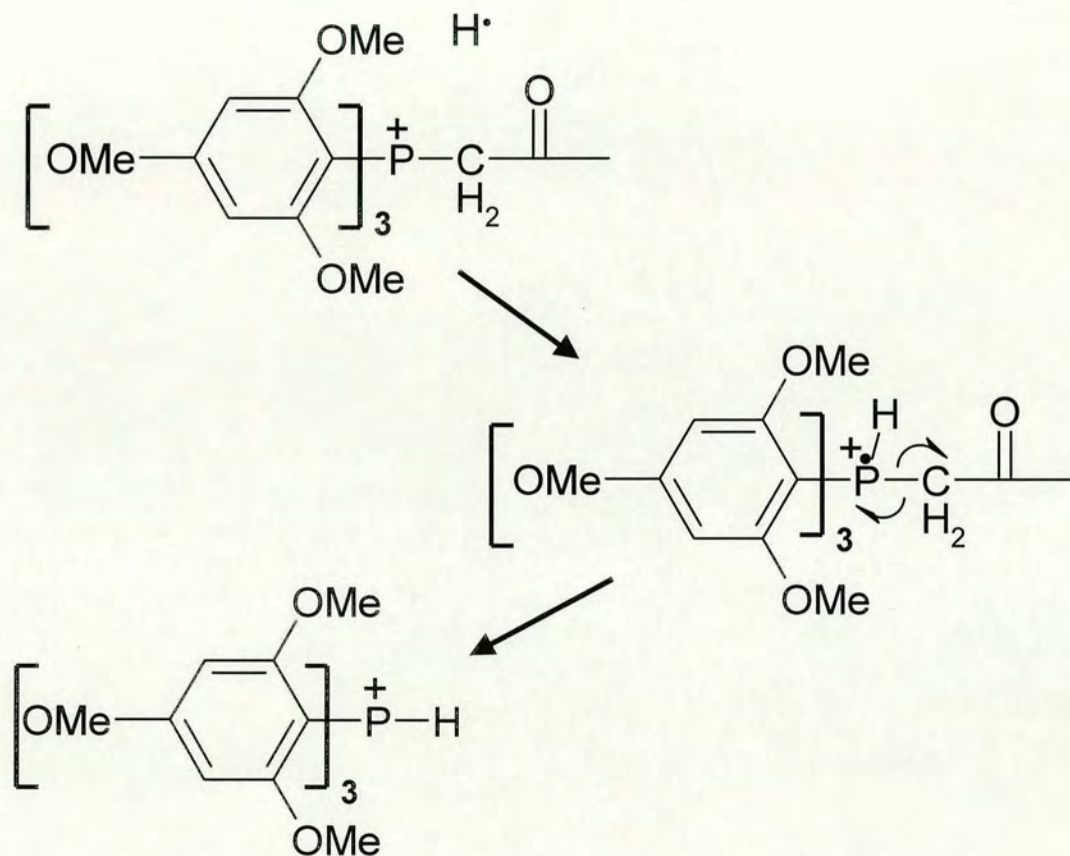
Figure 6.6 Schematic for hydrogen atom capture at P$^+$ causing radical cleavage. (Note that arrows refer to single-electron movement).

6.2.2 Peptide without arginine residue

Figure 6.7 shows the ECD mass spectrum of the doubly protonated peptide with sequence VYIHPF. The fragmentation pattern is very unpredictable and multiple processes seem to be taking place. Many of the fragment peaks are unidentified and may be due to internal fragments[2]. The occurrence of b fragments are likely to be due to vibrational excitation processes. Since VYIHPF is a small peptide and the proton affinity of the least basic site is low, the resulting recombination energy from electron capture is particularly large (see Figure 1.13, Chapter 1). Hence, ECD may result in highly energetic fragments that further fragment by vibrational excitation dissociation resulting in b and y fragments, as well as internal fragments.

All identified fragments (apart from x_1^+ and b_2^+) include the most basic residue in the molecule, namely histidine. This once again confirms the correlation between the higher gas-phase basicity of the protonated residue and the probability that this residue is included in a charged ECD fragment. The insert in Figure 6.7 shows the detail of the 'reduced' species and hydrogen atom desorption. The 'reduced' species ($[M+2H]^{+\bullet}$) is not present in this case since hydrogen atom desorption (peak labelled $[M+H]^+$) is much more prominent. The high hydrogen atom desorption may be due to the lower gas-phase basicities of the protonated sites in this molecule. Surprisingly, no neutral side-chain losses are observed in this spectrum.

From this result it appears that non-specific protonation (i.e. more protonated sites than basic sites) leads to more unpredictable fragmentation. Conversely, the presence of two basic sites in a doubly protonated molecule, and in particular one higher gas-phase basicity site such as arginine, leads to more predictable ECD spectra. These results are analogous to the work by Wysocki and co-workers on surface-induced dissociation (SID) [6], who showed that the number of ionising protons relative to basic sites influences the dissociation behaviour of the peptide [7].

---

[2] An internal fragment is due to two backbone cleavages, one N-terminal and one C-terminal, thus resulting in a fragment ion that does not include either terminus.
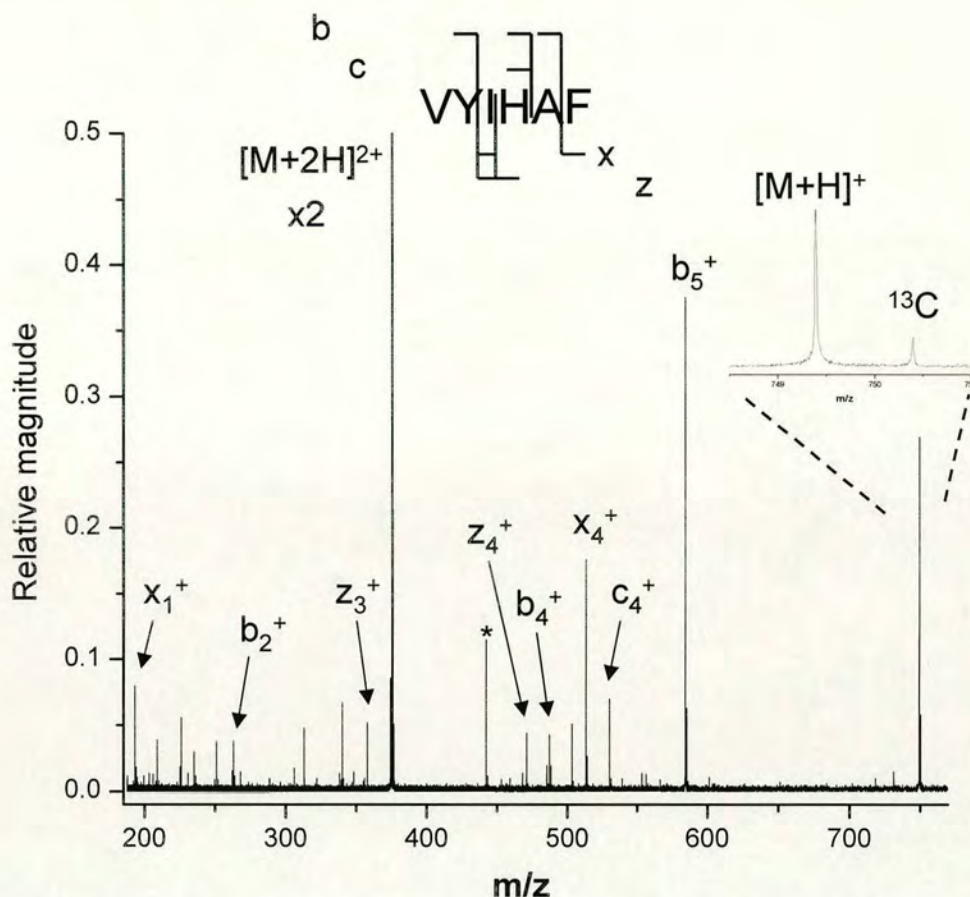
Figure 6.7 ECD mass spectrum of doubly protonated peptide VYIHPF. (200 time-domain transients, hexapole accumulation = 0.6 sec, electron irradiation = 30 ms).

6.3 Slow hydrogen atom transfer

The previous results have shown that the arginine residue may play an important part in ECD by making ECD spectra more predictable. An arginine residue that should resemble a neutralised protonated arginine side-chain (i.e. a protonated arginine side-chain that has captured an electron) was built using the procedure in Chapter 4 (section 4.2.1). The RESP fit charges for this residue (ARE) are shown in Figure 6.8. The ESP charge fitting procedure was performed on the trimer p-Glu-Arg-Gly-OH, since charged termini (protonated N-terminus, de-protonated C-terminus) gave rise to a distortion of the charges (data not shown). The RESP derived charges for the protonated arginine residue (ARG) contained in AMBER 7 [8] are

shown for reference in Figure 6.9. While the backbone charges (N, C_α, C and O) are not identical for both residues, they do not differ significantly. The sum of the charges for ARE $\Sigma_{charges}$ = 0.016, which is close an approximation to zero. Comparing the charges for both residues on the guanidine group (-NH-C-(NH$_2$)$_2$) (ARE $\Sigma_{charges}$ = -0.181 and ARG $\Sigma_{charges}$ = 0.6898) shows that most of the charge of the electron is delocalised over this group.

The assumption is made here that electron capture at the guanidine group of the protonated arginine side-chain could effectively result in keeping the electron localised at this site for a period of time. Since the guanidine group has a high degree of charge delocalisation (which explains the high proton affinity of this group) it is conceivable that the additional electron density to the ion resulting from electron capture could be mainly localised over the guanidine group.

The peptide [LHRH_HR]$^{+\bullet}$ [3] was built in LeAP and a topology file was generated. The coordinate file for [LHRH_HR]$^{2+}$ min73 for the molecular dynamics simulation at 300 K (structure 300K.1000ps) was taken as the coordinate input file for the molecular dynamics of [LHRH_HR]$^{+\bullet}$ at 300 K and 500 K. Figure 6.10 shows the interatomic distance [{2}(His)N-HD1...N(Arg){8}] at 300 K and 500 K dynamics, where (His)N-HD1 is a nitrogen-bound hydrogen on the histidine side-chain and N(Arg) is one of the amino nitrogens on the arginine side-chain.

---

[3] [LHRH_HR]$^{+\bullet}$ differs from [LHRH_HR]$^{2+}$ by the fact that the residue ARG is replaced with the new residue ARE.
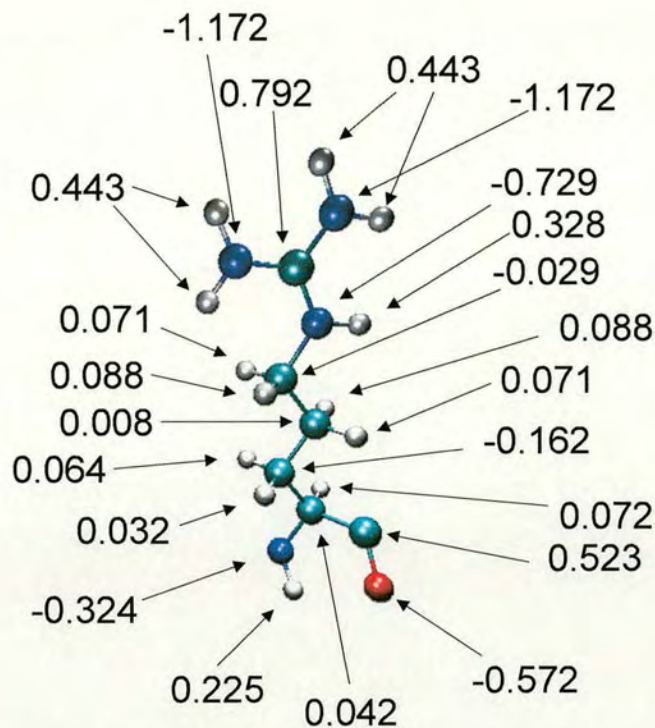
Figure 6.8 RESP fit charges on protonated arginine residue that has captured electron on side-chain (ARE). $\Sigma_{\text{charges}} = 0.016$; $\Sigma_{\text{charges(guanidine)}} = -0.181$.
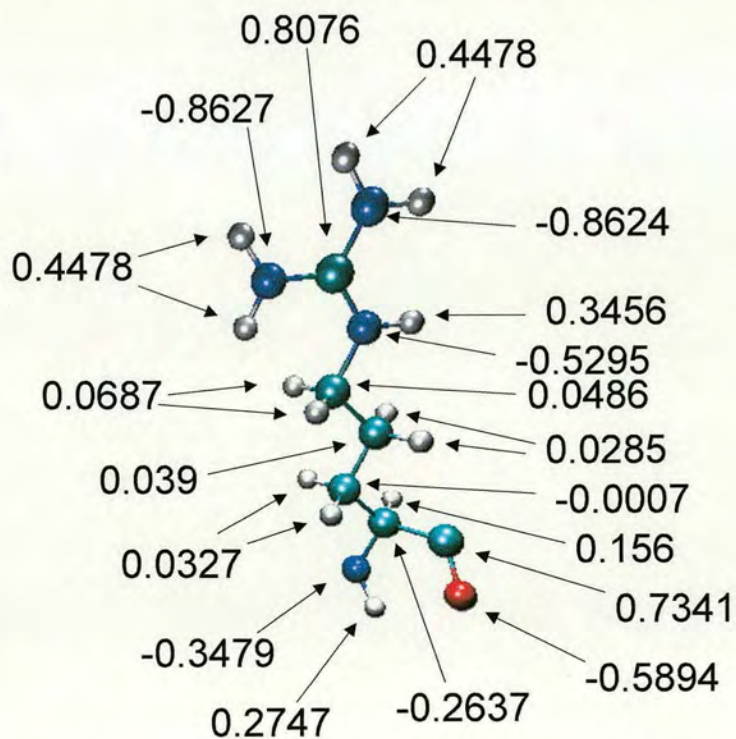

Figure 6.9 RESP fit charges on protonated arginine residue (ARG) contained in LeAP (AMBER program). $\Sigma_{\text{charges}} = 1$; $\Sigma_{\text{charges(guanidine)}} = 0.6898$. Charges from Cornell *et al.* [9].
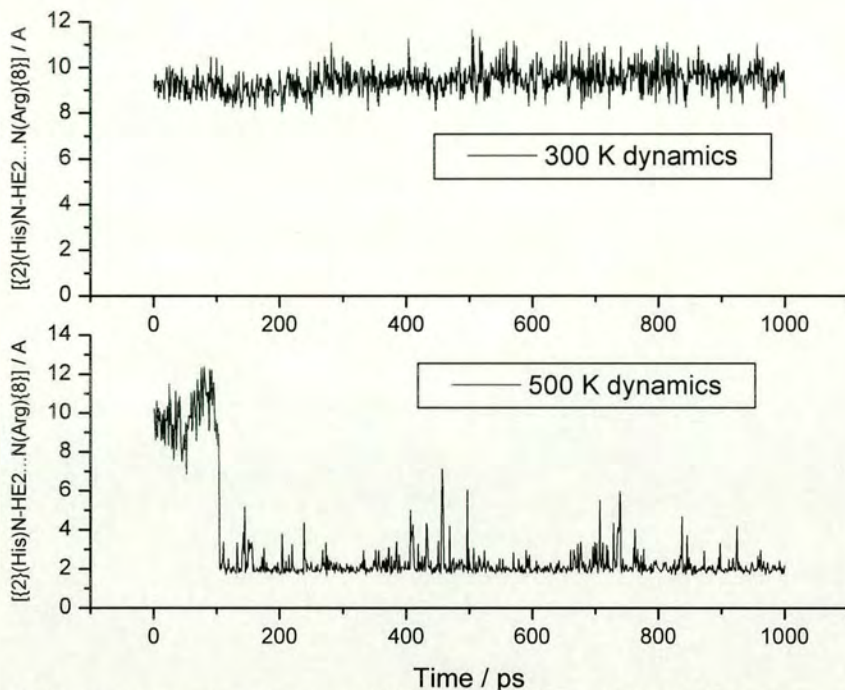
Figure 6.10 Interatomic distance [{2}(His)N-HE2…N(Arg){8}] for molecular dynamics of [LHRH_HR]$^{+\bullet}$ at (A) 300 K and (B) 500 K.
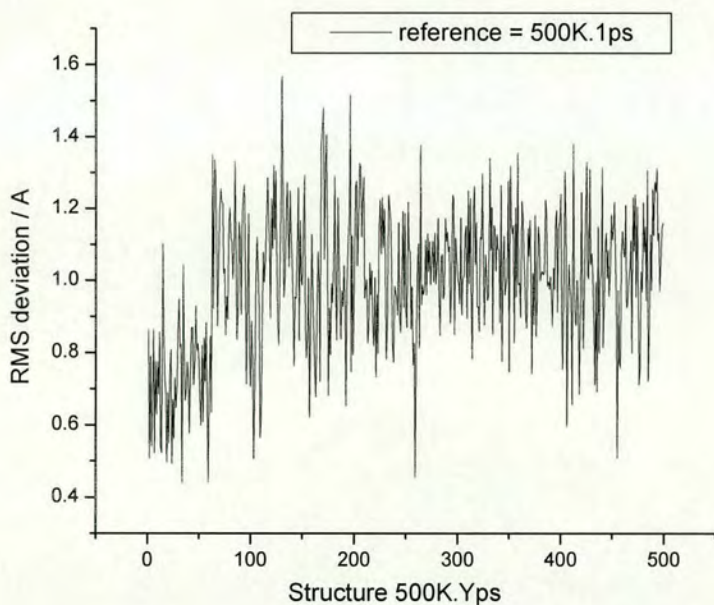
Figure 6.11 RMS (in Angstrom) for backbone (N, C$_\alpha$, C, and O) of [LHRH_HR]$^{+\bullet}$ for dynamics at 500 K (reference = 500K.1ps).
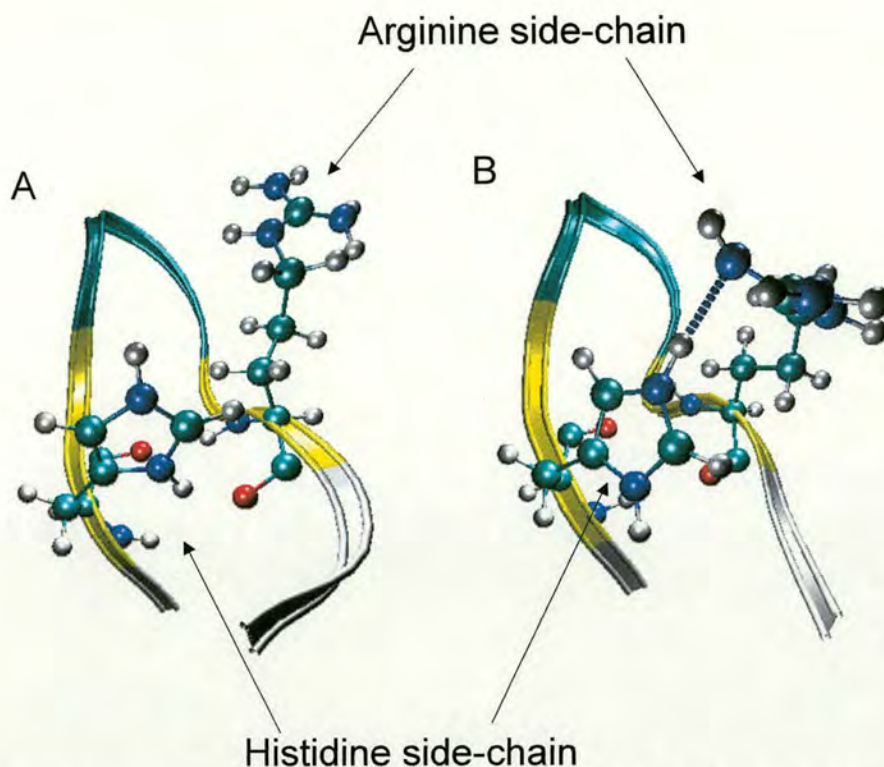
Figure 6.12 Representative structures of [LHRH_HR]$^{+\bullet}$ for dynamics at 500 K (A) 500K.50ps and (B) 500K.100ps. β-turn for residues 4-6 represented in cyan and β-sheet represented in yellow.

Modelling of this peptide has shown that the side-chains of histidine and arginine enter into hydrogen bonding contact after 100 ps dynamics at 500 K and that they remain in this state for the remainder of the simulation (see Figure 6.11). The increase in temperature from 300 to 500 K could be rationalised by the fact that electron capture is an exothermic process (releasing 4-7 eV energy). Importantly, the backbone structure of the peptide changes little over the length of this simulation (see RMS deviation Figure 6.11). A comparison to the RMS deviation for the 500 K dynamics simulation of [LHRH_HR]$^{2+}$ (see Figure 4.10, Chapter 4) shows little difference to the RMS deviation for the 500 K dynamics simulation of [LHRH_HR]$^{+\bullet}$ (see Figure 6.11). Hence, the interaction between the histidine and arginine (plus electron) side-chains arises mainly due to the difference in atomic charges on the guanidine group for ARE as opposed to ARG. Figure 6.12 shows representative gas-phase structures of [LHRH_HR]$^{+\bullet}$ for simulations at 500 K after 50 ps and 100 ps.

This result suggests that if the captured electron remains localised on the arginine side-chain the histidine and arginine side-chains interact. This could explain proton transfer from the less basic (i.e. histidine side-chain) to the more basic site (i.e. arginine side-chain) prior to or during backbone cleavage. In fact, this result could be interpreted in terms of a 'concerted' mechanism, where hydrogen atom transfer from the arginine side-chain to the backbone is concerted with proton transfer from the histidine side-chain to arginine (see Figure 6.13). This mechanism is consistent with the observation that in doubly protonated peptides the charged fragment ion always includes the most basic amino acid residue.
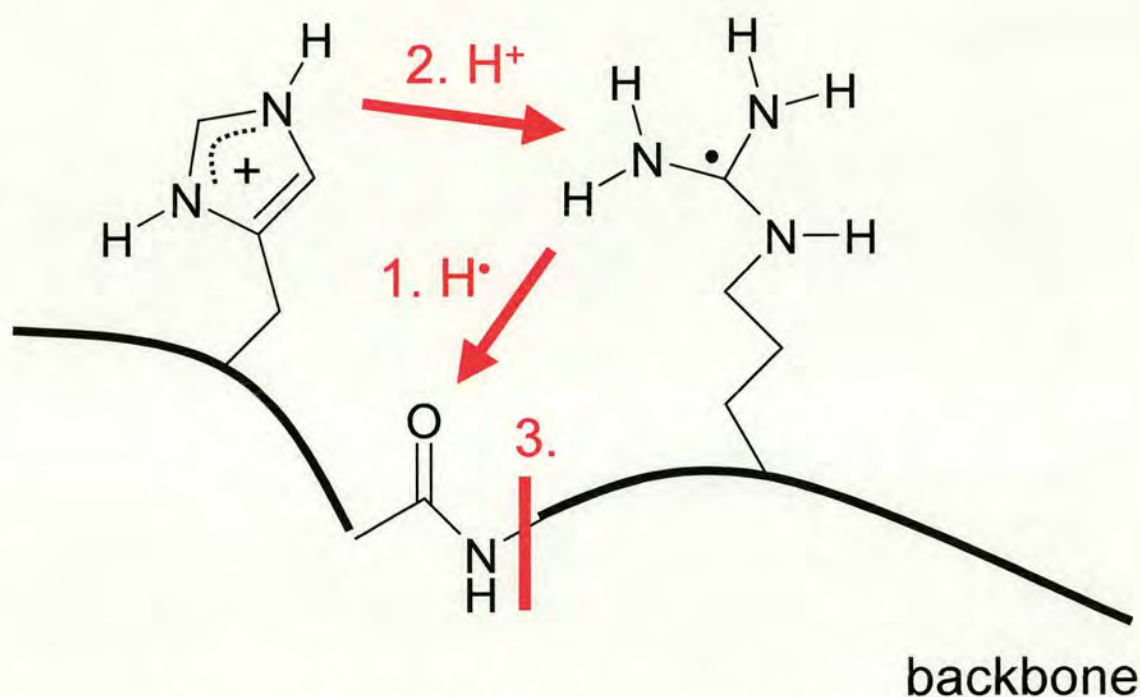
Figure 6.13 'Concerted' mechanism for [LHRH_HR]$^{+\bullet}$ where electron is captured on the arginine side-chain: 1. hydrogen atom transfer from arginine side-chain to backbone carbonyl, 2. proton transfer from histidine to arginine side-chain and 3. cleavage of backbone (for 'normal' backbone alpha-cleavage).

6.4 Summary and final conclusions

The ECD results in this thesis suggest that a *hydrogen atom re-arrangement* can take place in the ECD mechanism. The highest fragment peak in the peptide with the fixed charge group (Figure 6.5, section 6.2.2) can only be explained as due to the incorporation of a hydrogen atom. The occurrence of c fragments without incorporation of a hydrogen atom and z fragments which include an additional hydrogen atom, referred to as c$^\bullet$ and z' fragment ions, can be explained by a 'reverse' alpha-cleavage mechanism (see section 3.4, Chapter 3). Trends influencing the occurrence of the 'normal' and the 'reverse' cleavage are not yet clear. Nevertheless, these results seem to indicate that the location of charge on the precursor ion and the gas-phase conformation of the peptide may play a role.

This work has also shown that in a peptide with relatively unambiguous assignment of the protonation sites the most basic site is preferentially included in the charged fragments. Comparing the ECD spectra of [RVYIH(Bz)PF]$^{2+}$ (Figure 6.2) and [VYIHPF]$^{2+}$ (Figure 6.7), ECD of the latter is less clear. The trend for the studied peptides in this thesis is that the presence of an arginine residue and a residue of lower gas-phase basicity (such as histidine) gives the most predictable ECD spectra. In terms of the *non-ergodic* mechanism in ECD and the proposed *hot hydrogen re-arrangement* mechanism this could be explained by less efficient transfer of the hydrogen atom from the arginine residue to the backbone. In other words, the hydrogen atom from the arginine side-chain is less 'mobile' than the hydrogen atom from the histidine side-chain, in analogy to the *mobile proton* model by Wysocki and co-workers [10].

The *ECD/Hbond correlation* study (Chapters 3 and 4) for [LHRH_HR]$^{2+}$ and [BOMB]$^{2+}$ seemed to show agreement between preferential hydrogen bonding of the histidine side-chain to particular backbone carbonyls (as predicted by low-energy structures) and enhanced ECD cleavage at those sites. Conversely, no such correlation was found in the case of [LHRH_HR_NH$_2$]$^{2+}$.

The ECD results for [Ac_BK]$^{2+}$ are not easily explained in terms of the *hot hydrogen re-arrangement* mechanism. Nevertheless, the observation that the most basic site is preferentially included in the charged fragments is conserved. The fact

that the most basic residue is preferentially included in charged fragments implies that the mechanism here is not *non-ergodic*.

Whereas, electron capture is likely to be a very fast process (< ps), the mechanism involving hydrogen atom re-arrangement in the molecule culminating in backbone cleavage may take place on a slower time-scale (i.e. ps or ns). The modelling results of [LHRH_HR]$^{+\bullet}$ (section 6.3) are not robust, since molecular mechanics models are not accurate for radical species. Further, the electron is thought to be captured in a Rydberg state initially [11], which is not represented in the model. Nevertheless, the possibility of proton and hydrogen atom re-arrangement taking place before the molecule is broken apart may need to be taken into account.

Given the very long time-scale of FT-ICR experiments (100's ms) relative to the time-scale of the dissociation, it is difficult to say whether the mechanism really is *non-ergodic* or not. The estimated energy released from electron capture (4-7 eV, equivalent to approximately 400-700 kJ mol^{-1}) is easily sufficient to cause backbone cleavage. Recent *ab initio* calculations by Tureček have shown that it may be unnecessary to invoke the hypothesis of a *non-ergodic* mechanism [12].

The debate about the mechanism in electron capture dissociation will generate controversy for some time to come until a spectroscopic technique can prove that the mechanism is not *non-ergodic* (proving the opposite would be harder). Some of the work in this thesis may be of benefit for predicting ECD spectra, especially in relation to the sequencing of tryptic digest peptides. This thesis has raised many questions about the ECD mechanism. The crucial question on whether the gas-phase structure plays an important role in determining the relative fragmentation still remains to be answered. Hence, the applicability (and limitations) of electron capture dissociation as a tool to investigate the gas-phase structures of ions is yet to be verified. Hopefully, this thesis may provoke some discussion which will bring us closer to solving this scientific puzzle.

6.5 Further work

Some of the peptides in this study, and particularly the LHRH variants, should be characterised using ion mobility measurements. Such experiments should establish the preferential gas-phase conformations of these peptides. If the nature of the amino acid in position 6 for doubly charged LHRH peptides also changes the gas-phase conformation, this may give clues to the differences in 'normal' and 'reverse' alpha-cleavage for ECD of $[LHRH\_(l)W]^{2+}$ and $[LHRH\_(d)W]^{2+}$ in particular. Such a result would also imply that the relative ECD fragmentation pattern is not significantly affected by the gas-phase conformation.

Molecular mechanics modelling of some of these peptides could be carried out using the newer 2002 force field (ff02) in AMBER. This could show whether the 'loop' structure involving a β-turn is predicted by ff02 as a stable low-energy conformation for LHRH peptides in the gas-phase. In other words, the accuracy of the newer force field ff02 for predicting gas-phase conformation could be tested.

Molecular mechanics modelling of $[LHRH\_HR]^{+\bullet}$ (Figure 6.12 B) could be taken as a starting geometry for *ab initio* calculations to determine whether effective proton and hydrogen atom transfer could take place under these conditions.

Ab initio calculations should also be performed to determine the hydrogen atom affinity of the phosphorous atom in the fixed charge group (Figure 6.4). If the hydrogen atom affinity of this site was found to be exceptionally high, this would explain the high cleavage at that site. This would also add more weight to proving that the *hot hydrogen re-arrangement* mechanism takes place in ECD.

Finally, the protonation sites of peptides in the gas-phase made by electrospray ionisation need to be characterised. Higher gas-phase basicity sites are not necessarily protonated in favour of less basic sites if the alternative structures are more energetically favourable overall. In fact, protons may be 'mobile' in the ion, depending on the internal energy of the ion. Hence, a direct probe of protonation sites is required, which is most readily achieved with infrared spectroscopy. I hope to investigate some of these problems at the FOM institute (Rijnhuizen, Netherlands) using the free electron laser facility (FELIX).

References:

(1)   Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn David, M.; Carpenter Barry, K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.

(2)   Weiner, S. J.; Kollmann, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta Jr., S.; Weiner, P. J. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.

(3)   Weiner, S. J.; Kollmann, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230-252.

(4)   Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8949-8958.

(5)   Tsaprailis, G.; Nair, H.; Zhong, W.; Kuppannan, K.; Futrell, J. H.; Wysocki Vicki, H. *Anal. Chem.* **2003**, submitted.

(6)   Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *J. Mass Spectrom.* **1996**, *31*, 339-350.

(7)   Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 5142-5154.

(8)   Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K., M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R.; Duan, Y.; Pitera, J.; Massova, I. G.; Seibel, G. L.; Singh, U. C.; Weiner, P. J.; Kollmann, P. A. In; AMBER 7; University of California: San Francisco.

(9)   Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K., M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollmann, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

(10)   Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399-1406.

(11)   Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. *Eur. J. Mass Spectrom.* **2002**, *8*, 337-349.

(12)   Turecek, F. *J. Am. Chem. Soc.* **2003**, *125*, 5954-5963.

Appendix 1

Potential energy surface

Since the fundamentals of the SIMION program have been explained by David Dahl elsewhere (see literature referred to in section 1.3.3, Chapter 1), this appendix will only give a brief explanation of SIMION. Electrostatic fields can be modelled as boundary problem solutions of an elliptical differential equation called the Laplace equation:

$$DEL2V = DELxE = dEx/dx + dEy/dy + dEz/dz = 0 \qquad \textbf{Equation A1.1}$$

In the Laplace equation a zero charge volume density assumption is made, which means that it does not allow for space charge effects. Conversely, Poisson's equation does allow for such space charge effects to be taken into account. This is a limitation to SIMION, which needs to be taken into account especially when carrying out simulations on high density ion beams. Furthermore, given the limited space charge capacity of an ICR cell, such considerations should not be neglected.

In SIMION the Laplace equation is solved for each grid unit by using a finite difference technique called *over-relaxation*. In a 3D array, as was used in the practical simulations in this work, each array point has 6 nearest neighbours. Thus, each array point is calculated as the average of its nearest neighbours:

$$V = (P1+P2+P3+P4+P5+P6)/6 \qquad \textbf{Equation A1.2}$$

The software performs this by undertaking a number of iterations on each point until the value of each point reaches a set convergence limit.

Ion trajectory calculations

SIMION calculates the forces acting on a charged particle in each grid unit and expresses them in the form of volts per grid unit in the case of electrostatic fields and Gauss in the case of magnetic fields. The resulting acceleration on the particle is then added to the velocity that the ion has when entering the particular grid unit. The

position and velocity of the ion are calculated for each time increment. However, in order to minimise computing requirements, SIMION uses a self-adjusting time step method, which is a standard fourth order Runge-Kutta method. Hence in regions where the voltage gradient is almost linear, larger time steps are used, whereas smaller time steps are used in regions where steeper voltage gradients are present.

User programs

SIMION offers a very powerful tool to model ion traps, quadrupoles, collision cells and a range of other devices that rely on changing conditions (i.e. voltages, pressure) over time. Such user programs are contained in ASCII files with the extension .PRG and are based on a RPN (reverse Polish notation) software language. The user program can dynamically change electrode voltages at each time increment and hence significantly extend the capabilities of SIMION. In the light of ICR cells the applicability of user programs can be used to model the dipoar excitation of ions before detection, quadrupolar excitation axialisation (QEA), as well as sustained off-resonance excitation collisionally anduced dissociation (SORI CAD). Given the complex trajectories of ions in the ICR cell in such conditions, SIMION can provide a very powerful visualisation tool.

ICR open-ended cell

The ICR cell used in these simulations (see Figures A1.1 and A1.2) is modelled on a capacitively coupled open-ended cell, which refers to a detection cell where the dipolar excitation is also applied to the excitation trapping plates (i.e. electrodes 1, 4, 9 and 11). The cell consists of 3 elements, which are made of 4 sub-elements each. The cell entity is enclosed by an electrode box held at 0 V (i.e. ground). The reason for the outer electrode box is to prevent a distortion in the electrostatic energy surface, due to an infinite boundary problem. Whereas in Figure A1.1 this outer electrode is shown, Figure A1.2 shows the same instance without the outer box electrode and also the direction of the magnetic field as defined by the magnetic instance. The dimensions of the electrostatic array are x=67, y=67 and z=201, resulting in > 900000 grid units (1 grid unit = 1 mm^3). A magnetic instance

231

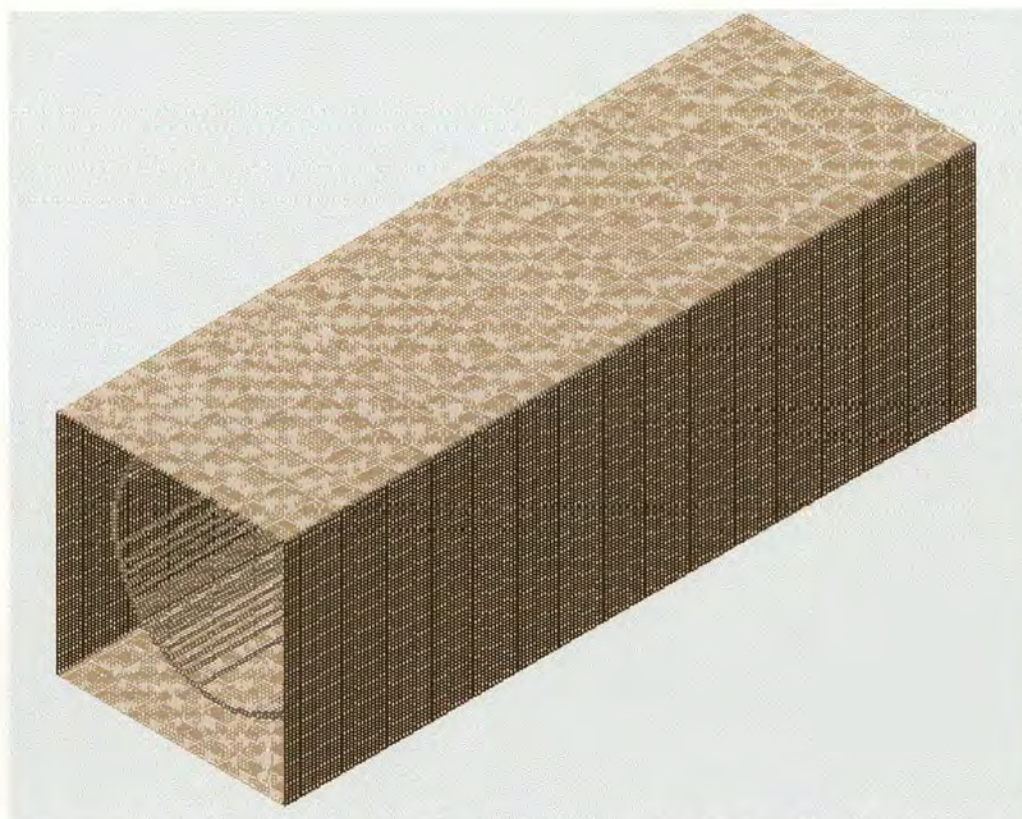(x=7, y=7, z=26) is scaled to overlapped with the electrostatic instance (1 grid unit = 1,000 mm^3).

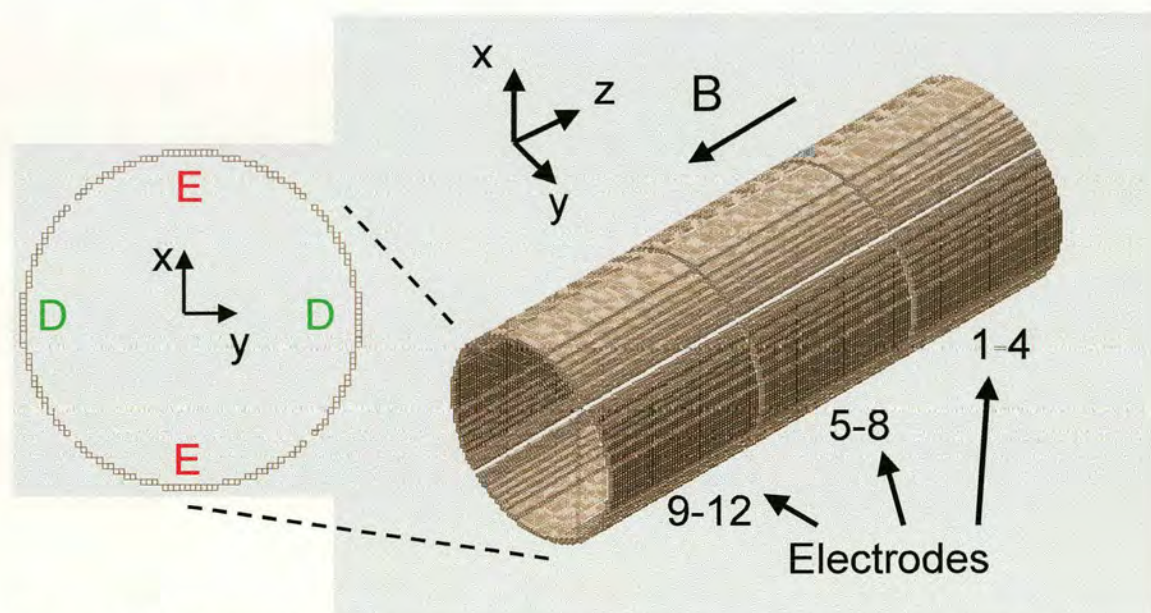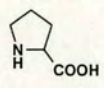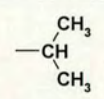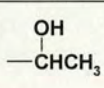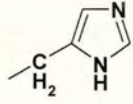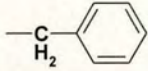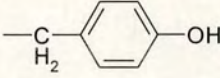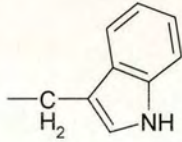**Figure AI.1** Electrostatic instance comprising the ICR cell and the surrounding
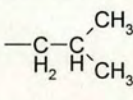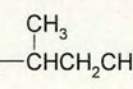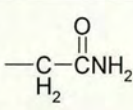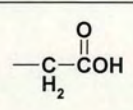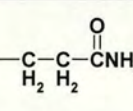
Figure AI.2 Electrostatic instance showing ICR cell with 12 electrodes. Insert shows xy cross-section of cell, indicating 4 electrodes, which consist of oppositely placed excitation electrodes (E) and detection electrodes (D).

Appendix 2

Amino Acid	One-letter code	Three-letter code	Structure of R group	Residue mass (Da)
Glycine	G	Gly	-H	57.02
Alanine	A	Ala	$-CH_3$	71.04
Serine	S	Ser	$-CH_2OH$	87.03
Proline[#]	P	Pro		97.05
Valine	V	Val		99.07
Threonine	T	Thr		101.05
Cysteine	C	Cys	$-CH_2SH$	103.01
Leucine	L	Leu		113.08
Isoleucine	I	Ile		113.08
Asparagine	N	Asn		114.04
Aspartate	D	Asp		115.03
Glutamine	Q	Gln		128.06

Lysine	K	Lys	-CH$_2$CH$_2$CH$_2$CH$_2$NH$_2$	128.09
Glutamate	E	Glu		129.04
Methionine	M	Met	-CH$_2$CH$_2$SCH$_3$	131.04
Histidine	H	His		137.06
Phenylalanine	F	Phe		147.07
Arginine	R	Arg		156.10
Tyrosine	Y	Tyr		163.06
Tryptophan	W	Trp		186.08

\# In case of proline, the whole amino acid structure is shown.

Table A2.1: Genetically encoded amino acids with structures and residue masses (mono-isotopic), i.e. amino acid masses after water loss in peptide bond formation.
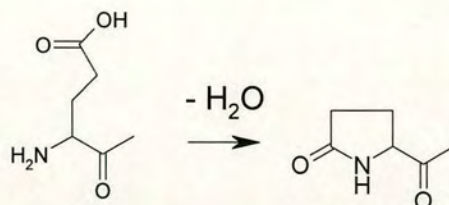

Figure A2.1 Formation of pyro-glutamic acid (pE) through elimination of water.

Amino acid	GB (kcal mol^{-1})	PA (kcal mol^{-1})^a	pKa Side-chain
Gly	202.7	210.5	
Cys	206.2	214.0	8.37 (sulfhydril)
Ala	206.4	214.2	
Ser	207.6	215.2	
Asp	208.6	216.4	3.90 (β-COOH)
Val	208.7	216.5	
Leu	209.6	217.4	
Ile	210.8	218.6	
Thr	211.7	219.5	
Phe	212.1	219.9	
Asn	212.8	220.6 (223.6)	
Tyr	213.1	220.9	10.46 (phenol)
Met	213.3	221.1 (224.1)	
Gln	214.2	222.0 (225.0)	
Pro	214.3	222.1	
Glu	215.6	223.4 (226.4)	4.07 (γ-COOH)
Trp	216.1	223.9 (226.9)	
Lys	221.8	235.6	10.54 (ε-NH$_3^+$)
His	223.7	231.5 (234.5)	6.04 (imidazole)
Arg	237.0	244.8 (247.8)	12.48 (guanidine)

Table A2.2 Evaluated gas-phase basicities (GB) and proton affinities (PA) (taken from Harrison, A. G. *Mass Spectrom. Rev.* **1997**, *16*, 201-217) and side-chain pKa values, where applicable, (taken from Dawson, R. M. C.; Elliott, D. C.; Elliott, W. H.; Jones, K. M. In *Data for Biochemical Research (3rd Ed.)*; Oxford Science Publications, 1986; pp 1-31). ^{a} Proton affinities estimated assuming $\Delta S^0_{rot} = 0$ except for lysine (Lys). Values in parantheses assume $\Delta S^0_{rot} = -10$ cal mol^{-1} K^{-1} due to intramolecular hydrogen bonding.
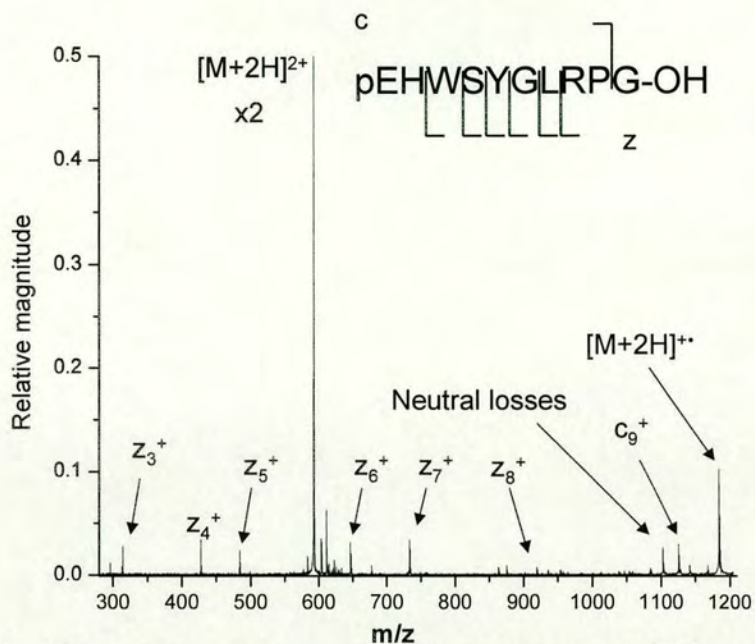
Appendix 3

Figure A3.1 ECD mass spectrum of doubly protonated pEHWSYGLRPG-OH. ([LHRH_HR]$^{2+}$) (100 time-domain transients) (3 T APEX II, electron filament, 500 ms electron irradiation)
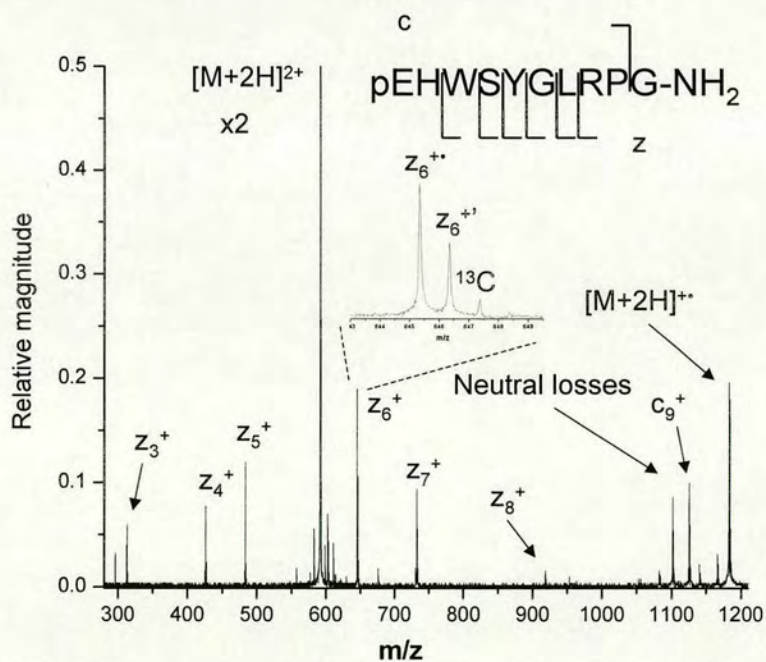

Figure A3.2 ECD mass spectrum of doubly protonated pEHWSYGLRPG-NH$_2$. (100 time-domain transients) (9.4 T APEX III, dispenser cathode, 30 ms electron irradiation)
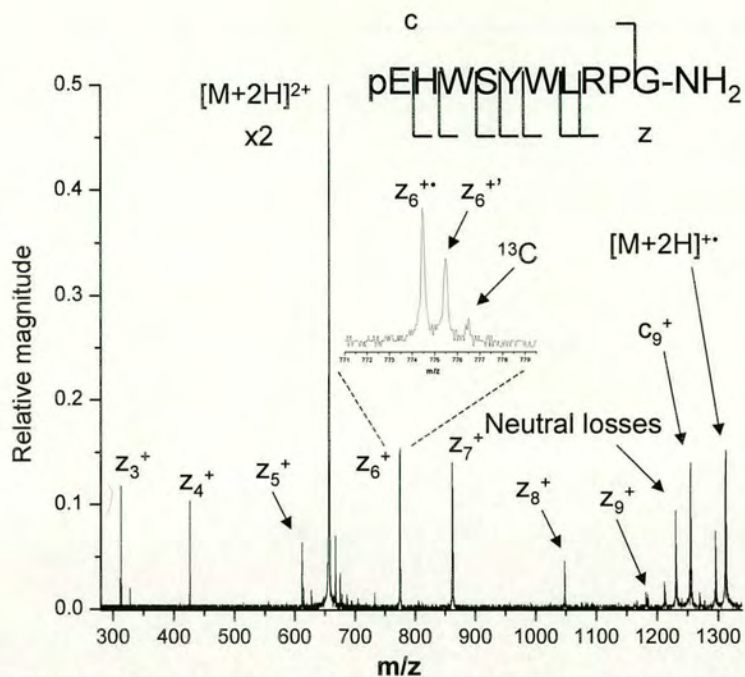
Figure A3.3 ECD mass spectrum of doubly protonated pEHWSYWLRPG-NH$_2$. (100 time-domain transients) (9.4 T APEX III, dispenser cathode, 35 ms electron irradiation)
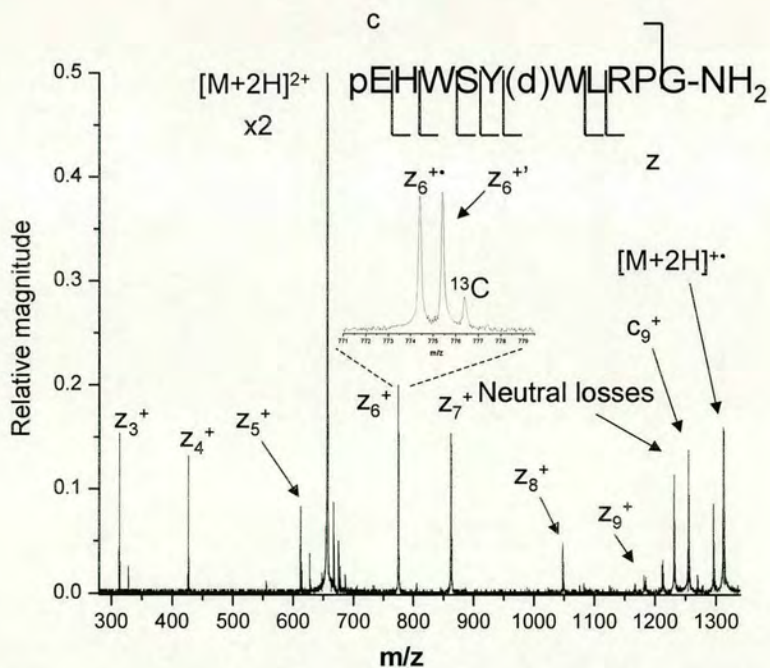

Figure A3.4 ECD mass spectrum of doubly protonated pEHWSY(d)WLRPG-NH$_2$. (100 time-domain transients) (9.4 T APEX III, dispenser cathode, 35 ms electron irradiation)
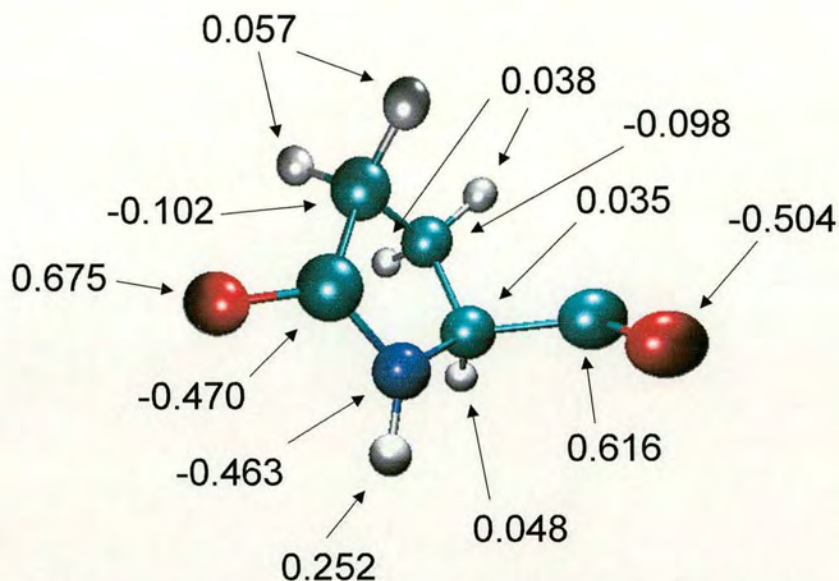
Appendix 4

Figure A4.1 Structure of pyro-glutamic acid residue with annotated RESP derived charges. $\Sigma_{charges} = 0.179$.[1]
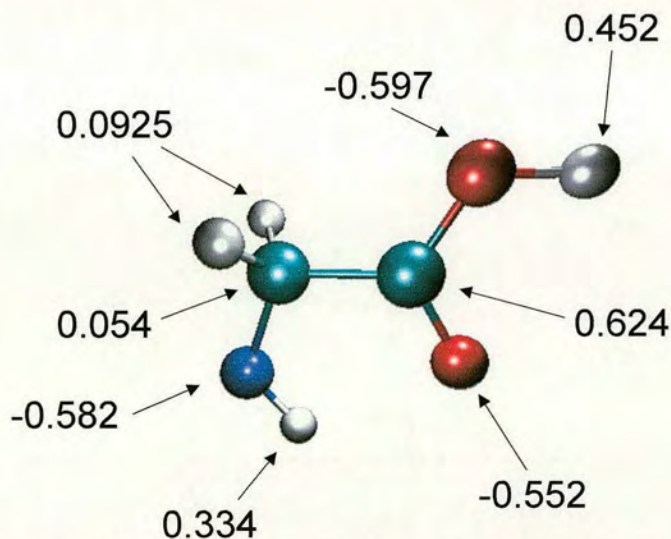
Figure A4.2 Structure of neutral C-terminal glycine with annotated RESP derived charges. $\Sigma_{charges} = -0.082$.[1]

---

[1] Residues made by Dr. Perdita E. Barran

Figure A4.3 Structure of neutral amidated C-terminal glycine with annotated RESP derived charges. $\Sigma_{charges}$ = -0.1173.[1]

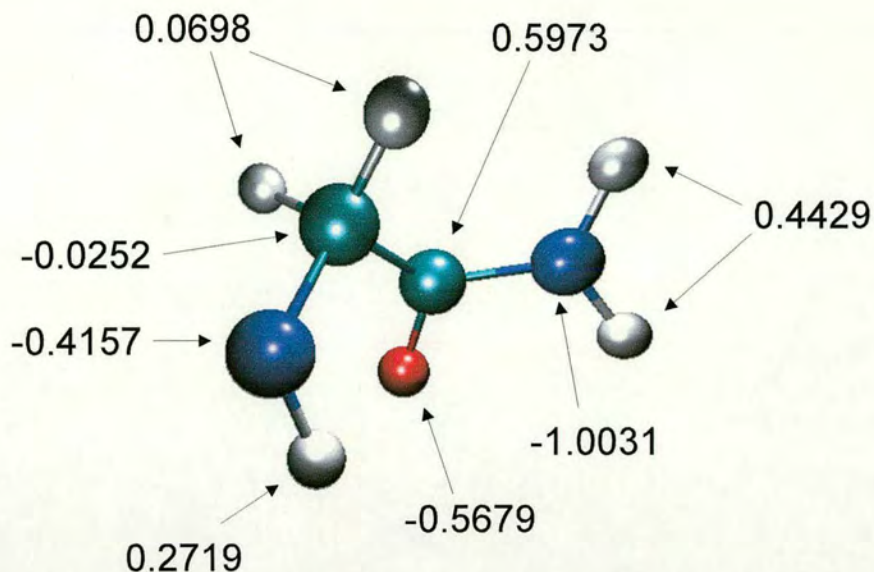**Figure A4.4** Structure of protonated amidated C-terminal glycine with annotated RESP derived charges. $\Sigma_{charges}$ = 1.001.
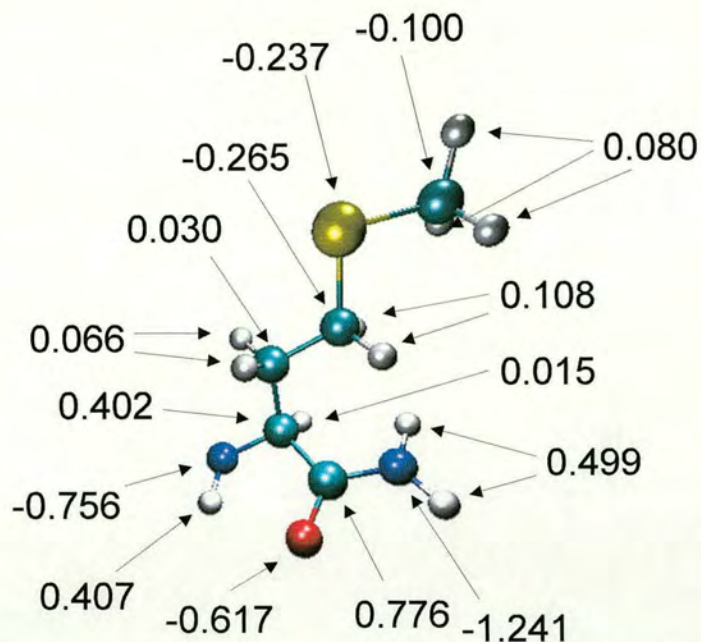
Figure A4.5 Structure of neutral amidated C-terminal methionine with annotated RESP derived charges. $\Sigma_{charges} = 0$. Charge on α-carbon (0.402) higher than expected.

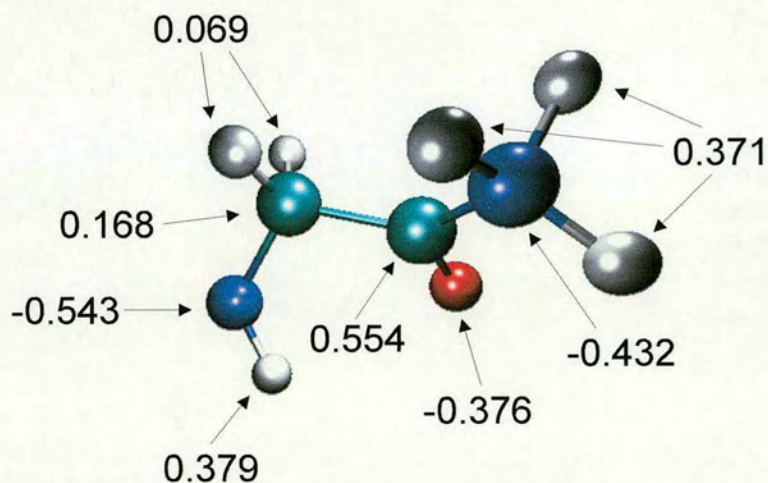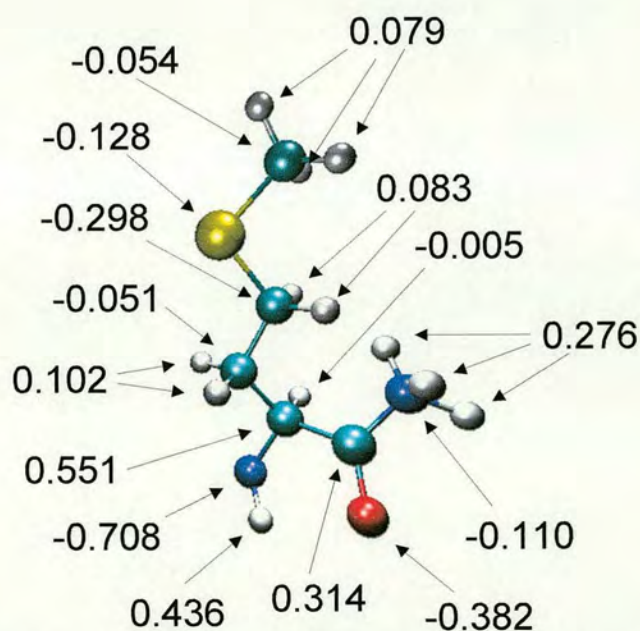**Figure A4.6** Structure of protonated amidated C-terminal methionine with annotated RESP derived charges. $\Sigma_{charges} = 1$. Charge on α-carbon (0.551) higher than expected.

Appendix 5

Articles included as part of this thesis

Polfer, N. C.; Haselmann, K. F.; Zubarev, R. A.; Langridge-Smith, P. R. R. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 936-943.

Blindauer, C. A.; Polfer, N. C.; Keiper, S. E.; Harrison, M. D.; Robinson, N. J.; Langridge-Smith, P. R. R.; Sadler, P. J. *J. Am. Chem. Soc.* **2003**, *125*, 3226-3227.

RCM

Electron capture dissociation of polypeptides using a 3 Tesla Fourier transform ion cyclotron resonance mass spectrometer

Nicolas C. Polfer[1], Kim F. Haselmann[1], Roman A. Zubarev[2] and Pat R. R. Langridge-Smith[1]*

[1]Department of Chemistry, The University of Edinburgh, King's Buildings, West Main Road, Edinburgh EH9 3JJ, Scotland, UK
[2]Department of Chemistry, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

Electron capture dissociation (ECD) of polypeptides has been demonstrated using a commercially available 3 Tesla Fourier transform ion cyclotron resonance (FTICR) instrument. A conventional rhenium filament, designed for high-energy electron impact ionisation, was used to effect ECD of substance P, bee venom melittin and bovine insulin, oxidised B chain. A retarding field analysis of the effective electron kinetic energy distribution entering the ICR cell suggests that one of the most important parameters governing ECD for this particular instrument is the need to employ low trapping plate voltages. This is shown to maximise the abundance of low-energy electrons. The demonstration of ECD at this relatively low magnetic field strength could offer the prospect of more routine ECD analysis for the wider research community, given the reduced cost of such magnets and (at least theoretically) the greater ease of electron/ion cloud overlap at lower field. Copyright © 2002 John Wiley & Sons, Ltd.

Electron capture dissociation (ECD)[1] has proven to be a valuable research tool for polypeptide sequencing and the localisation of post-translational modifications.[2-6] This fragmentation technique is complementary to the more established *heating* techniques, such as collisionally-activated dissociation (CAD)[7] and infrared multi-photon dissociation (IRMPD).[8] In combination with these other dissociation techniques, ECD can play an important role in *de novo* sequencing, where complete sequence coverage is required.

To date ECD has only been applied on Fourier transform ion cyclotron resonance mass spectrometers equipped with super-conducting magnets of 4.7 Tesla, or higher, field strength. The obvious advantages of higher field strength, B, are higher mass resolution (proportional to B) and the ability to trap more ions (proportional to B^2). However, at higher magnetic field, the ions and electrons are more tightly focused, potentially making their efficient overlap more difficult. The advantages of lower field magnets are reduced cost and greater magnetic field stability. The use of FTICR mass spectrometers equipped with lower field magnets could enable ECD mass spectra to be obtained more cost-effectively for lower mass compounds, such as tryptic digests of proteins, where high mass resolution is not as essential.

Correspondence to: P. R. R. Langridge-Smith, Department of Chemistry, The University of Edinburgh, King's Building, West Main Road, Edinburgh EH9 3JJ, Scotland, UK.
E-mail: prrls@ed.ac.uk

Since the electron capture cross-section increases with the charge squared (z^2) of the cationic species of interest, a requirement for ECD is an ionisation technique capable of generating multiply charged cations, such as electrospray ionisation (ESI). In order to increase the probability of electron capture, low-energy electrons (<0.2 eV) are also required.[1] This contrasts with the usual function of standard filament-based electron sources, which are designed for ionisation of neutral molecules by high-energy (e.g. 70 eV) electron impact.

A number of approaches have been proposed in order to improve ECD. McLafferty and co-workers[9] have used a design of ICR cell, consisting of an open-ended cylindrical cell with additional flat trapping plate electrodes, 'the nested ion cell geometry', which has enabled the simultaneous trapping of positively charged ions and electrons. They have reported ECD mass spectra for bovine ubiquitin showing extensive sequence coverage. However, the use of multiple electron irradiation (3 s) and storage (6 s) cycles (overall 10 × 9 s) increases the overall experimental time and can preclude the use of ECD in conjunction with on-line liquid chromatographic (LC) separation techniques.

In contrast, Tsybin *et al.*[10] and Haselmann *et al.*[11] have used indirectly heated high-surface area dispenser cathodes, in the former case fitted with a tungsten sponge impregnated with Ba-Ca aluminate, and in the latter a low work function BaO surface. Both groups have reported substantially reduced electron irradiation times, as short as 1 ms electron
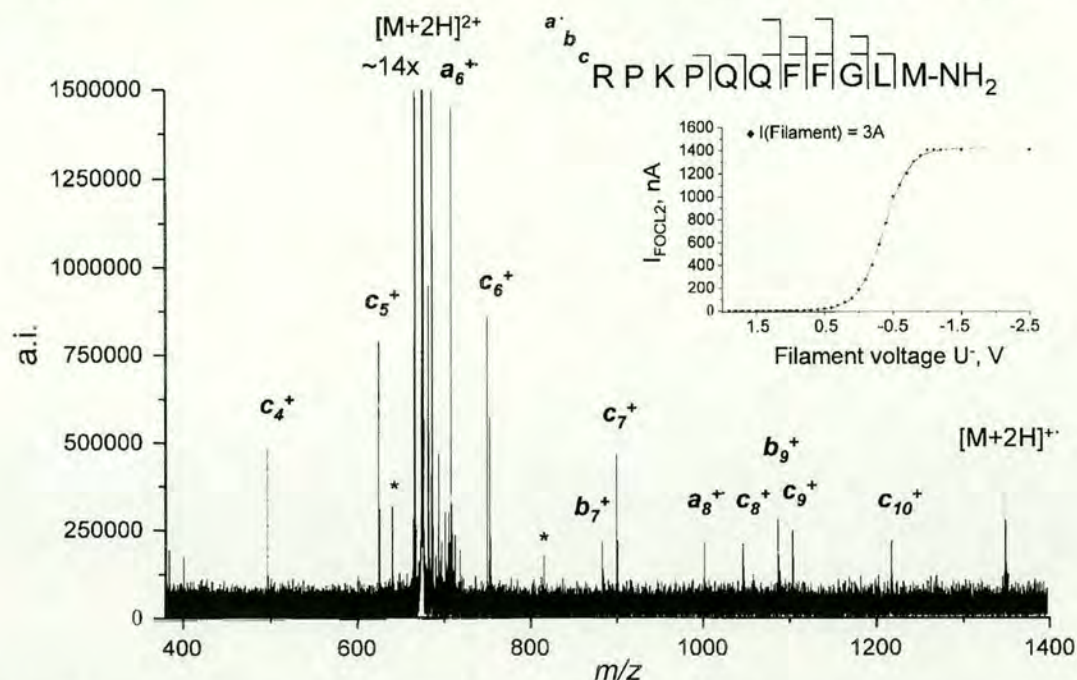
Figure 1. ECD mass spectrum of substance P in the 2+ charge state (40 scans). Electron irradiation time = 1 s. Asterisk denotes artefact peaks. The insert shows the electron current measured on the ion optic FOCL2, the rear element of the last Einzel lens that the ions pass through prior to entering the ICR cell, as a function of filament voltage U⁻.

irradiation,[10,11] which are more compatible with those required for on-line LC.

The use of dispenser cathodes has been shown to markedly improve the rate of ECD. Primarily, this is due to the higher emission current and the larger emitting surface area, which result in emission of electrons with a larger beam diameter. This not only increases overlap of the electron beam with the ion cloud, but it has also been suggested[10] that the electron beam itself is large enough to confine the ion cloud, thereby creating a potential well that traps the ions in the radial direction in conjunction with the magnetic field. Nevertheless, these dispenser cathodes do present practical problems, such as increased blackbody radiation, which can lead to increased pressure within the ICR cell, as well as concerns regarding the possibility of contamination of the ICR cell with Ba sputtered from the cathode surface.

These different approaches to ECD have only been demonstrated by a select few research groups and are not generally available to the wider research community. Most commercially available FTICR instruments already are equipped with standard tungsten or rhenium filaments, which in principle should allow ECD to be performed. In this paper we report results on ECD experiments which were carried out on a commercial FTICR mass spectrometer (Bruker Daltonics APEX II) equipped with a 3 Tesla superconducting magnet and a standard rhenium filament as the electron source. Characterisation of the electron kinetic energy distribution from the rhenium filament was carried out in an attempt to optimise the efficiency of the ECD process.

EXPERIMENTAL

The polypeptides studied in this work were substance P (MW 1347.6), bee venom melittin (MW 2846.5), bovine insulin, oxidised B chain (MW 3495.9), and bovine ubiquitin (MW 8564.9), which were all obtained from Sigma and used without further purification. The compounds were dissolved in a water/methanol/acetic acid solution (49:49:2, v/v/v) to concentrations of between 10–100 µM. Electrospray ionisation (ESI) mass spectra were generated using an Analytica ion source (Branford, CT, USA), which is equipped with a standard glass capillary (end-caps metallised) and a storage hexapole. The solutions were infused using a syringe pump (Cole Parmer, IL, USA), at flow rates of between 80–120 µL/h. Mass spectra were recorded using a passively shielded 3 Tesla FTICR instrument (APEX II) manufactured by Bruker Daltonics (Billerica, MA, USA). The cations produced by ESI were accumulated for 500–1000 ms in the hexapole of the ion source before being injected into the Infinity© ICR cell. The ions were trapped at low trapping plate voltages, typically 0.5 V, which were maintained during the entire experiment. No Sidekick© or pulsed gas was used to assist trapping of the ions. Precursor ion isolation was achieved using a correlated sweep isolation of the most intense charge state. In the case of substance P, the doubly protonated precursor ion was isolated, whereas for both melittin and B chain insulin the quadruply protonated precursor ion was isolated. For ubiquitin no isolation was carried out. In order to induce ECD inside the ICR cell the cations were irradiated, for a
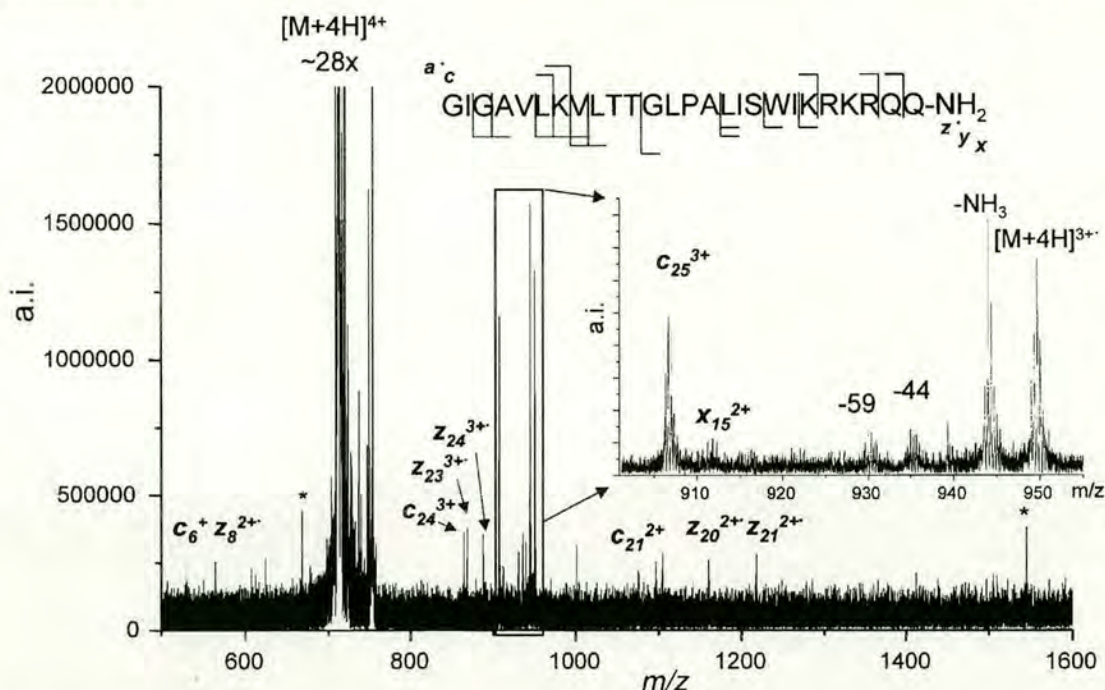
Figure 2. ECD mass spectrum of melittin in the 4+ charge state (100 scans). Electron irradiation time = 1.5 s. The insert shows an expansion of the *m/z* 950 region of the reduced species $[M + 4H]^{3+\cdot}$. Asterisk denotes artefact peaks.

period of 1–5 s, by low-energy electrons from a heated rhenium filament located 13 cm behind the ICR cell.

The rhenium filament (ca. 5 mm long Re ribbon, 0.038 × 0.51 mm, resistance 0.09959 Ohm/cm) was operated at a current of 3 A. During electron irradiation a voltage U^- (range −0.48 to +0.65 V) was applied to one side of the filament. As a consequence of the potential drop across the filament, a voltage U^+ ($U^+ = U^- + 2$ V) was measured on the other side of the filament. The emitted electron current was measured by using a picoammeter (model 610C; Keithley Instruments, Munich, Germany) at the Einzel lens element closest to the ICR cell (FOCL2). In the case of this particular instrument the filament voltage U^- at which no electron emission occurred was found to be +3.5 V. Detection of the precursor ions and their fragments was carried out in broadband mode using 512 K data points. No apodisation or extra zero-fill processing was carried out prior to fast Fourier transformation of the data. For all ECD mass spectra presented, between 40 and 200 scans were accumulated, unless otherwise stated.

RESULTS AND DISCUSSION

A prerequisite for ECD is good overlap of the ion cloud, trapped in the ICR cell, and the irradiating electron beam. The electron beam, however, has a small radius due to the effect of the magnetic field present and the limited surface area of the filament. For this design of ICR cell the trapping plates have small annular apertures (6 mm dia.). Trapping of ions on or very close to the centreline axis of the cell is therefore crucial. In order to trap ions at 0.5 V without

Sidekick® trapping or the use of a pulsed gas, the ion source and tuning parameters of the ion transfer optics had to be very carefully adjusted. The precursor ion magnitude had first to be maximised since electron irradiation was found to reduce the initial ion abundance by approximately two-thirds. Larger precursor ion magnitudes could be achieved by using higher trapping voltages. However, it was found that ECD was markedly reduced. The reason for this is discussed later.

The insert in Fig. 1 shows the current measured on the last Einzel lens element (FOCL2) that the ions pass through prior to entering the ICR cell, as a function of the filament voltage U^-. For experimentally favourable ECD conditions the optimum filament voltage U^- was found to be near +0.5 V; this corresponds to the onset of the steep rise in the emitted current. Another important parameter is the electron irradiation time. This must be long enough to permit sufficient electrons to enter the ICR cell, but not to lengthen excessively the acquisition time or to cause significant ion loss. Irradiation times of between 1 and 5 s were found to be sufficient. Since the ECD fragments are of low abundance the accumulation of several scans (S) is essential, thereby increasing the signal-to-noise (S/N) ratio by a factor of $\sqrt{S}$. This is a further reason for keeping the acquisition time for each single scan as short as possible. In the case of substance P, for example, 40 scans were found to be sufficient to generate ECD mass spectra with adequate sequence coverage.

ECD mass spectra

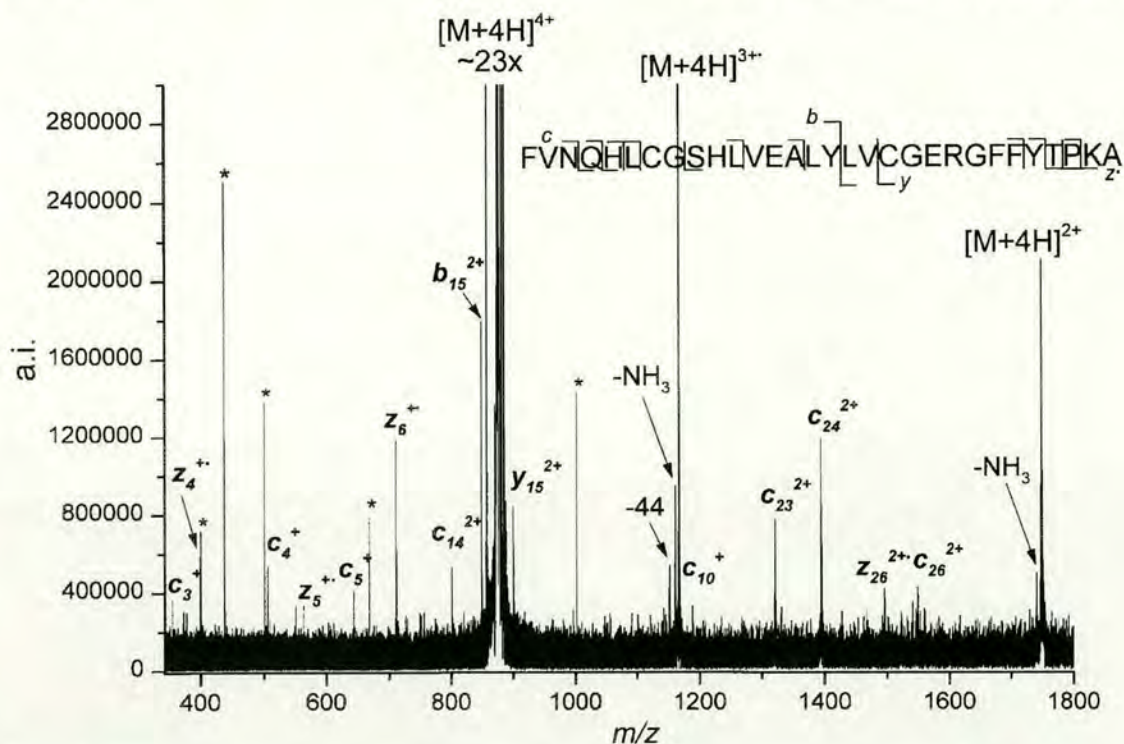The ECD mass spectrum of the neuromodulator substance P

Figure 3. ECD mass spectrum of bovine insulin, oxidised B-chain, in the 4+ charge state (200 scans). Electron irradiation time = 5 s. Asterisk denotes artefact peaks.

is shown in Fig. 1. A 1-s electron irradiation time was used and 40 time-domain transients were accumulated. The N-terminal c-ion fragments starting from c_4^+ to c_{10}^+ are visible in the mass spectrum. The c_8^+ fragment is of lowest abundance, with a S/N ratio of 10. This mass spectrum is representative of previously reported ECD mass spectra for this compound.[8] In the mass spectrum shown in Fig. 1, two b-ions are present, suggesting that other activation mechanisms, as well as electron capture, took place. Vibrational excitation due to blackbody infrared irradiation from the heated filament could explain such fragment peaks. In addition, two odd electron a-ions can also be seen in the mass spectrum.

For ECD of bee venom melittin, isolation of the quadruply protonated precursor ion was performed. The electron irradiation time was increased to 1.5 s. The optimum filament voltage U^-, where the reduced triply protonated species was just visible in a single scan, was found to be +0.65 V. The optimum filament voltage will probably vary from instrument to instrument, since the filament position, the ICR cell geometry and, most importantly the filament itself, will be different.

The ECD mass spectrum obtained for melittin is shown in Fig. 2. In this case, 100 scans were accumulated and a total of 17 ion fragments were observed, including four c-ions, eight z-ions, one a-ions, three y-ions and one x-ion. Fragment peaks arising from the triply protonated reduced species ($[M + 4H]^{3+\cdot}$ at 949.59 Th) due to the typical small neutral losses expected in ECD, such as hydrogen desorption,

ammonia and arginine side-chain losses, can also be seen. The mass resolution on the quadruply protonated precursor ion peak was approximately 18000. The individual peaks in the isotopic distribution had a symmetrical shape and were fully baseline resolved. For the reduced species the mass resolution dropped to 8000 and the peak shapes deteriorated slightly. Nevertheless, the mass resolving power of this 3 Tesla instrument was adequate since all the fragment peaks could be isotopically resolved.

In other experiments the 5+ charge state of melittin was isolated and subjected to ECD. However, no further structural information was gained. This was rather surprising, since the efficiency of electron capture increases with the square of the charge (z^2). The most likely explanation is due to the fact that the initial signal magnitude, after isolation, of the 5+ charge state of the precursor ion was only half that of the 4+ charge state. Addition of m-nitrobenzyl alcohol (m-NBA) was tried, in order to increase the signal magnitude at higher charge states,[9] but, although the relative abundance was shifted towards these higher charge states, the overall signal magnitude was suppressed.

SORI-CAD and IRMPD mass spectra were also recorded for the 4+ charge state of melittin. The typical b- and y-ion fragment pattern was observed (data not shown), which is complementary to the ECD fragmentation pattern. By combining the fragmentation data from the ECD mass spectrum as well as the SORI-CAD and IRMPD mass spectra, a total of 21 out of a possible 25 backbone bond cleavages could be identified.
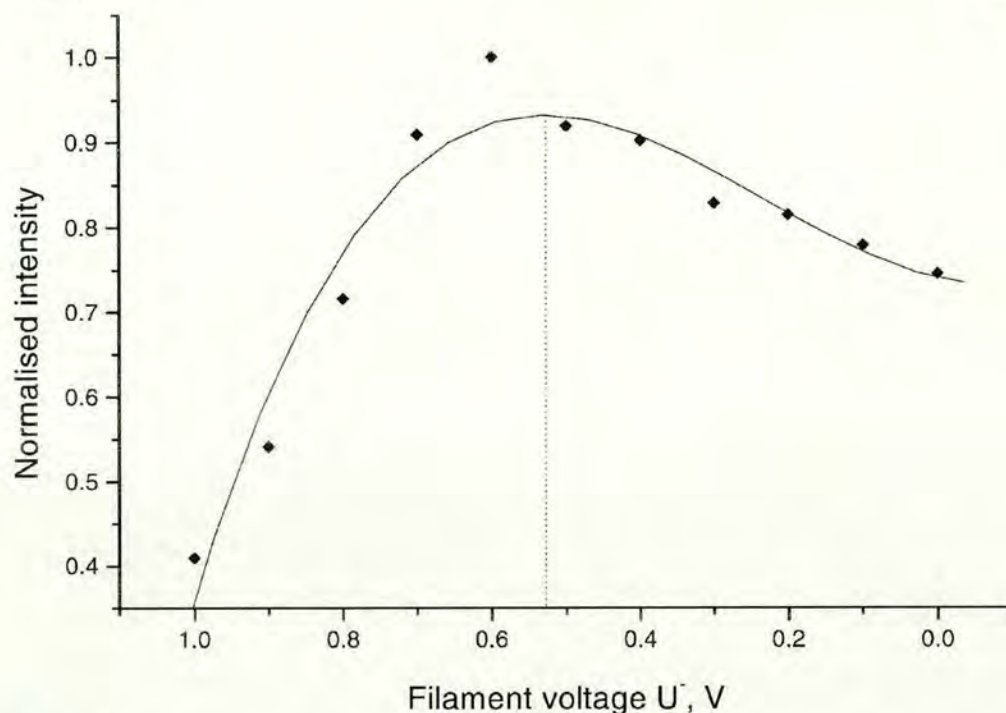
Figure 4. ECD fragment efficiency for substance P as a function of filament voltage U⁻. The ECD fragment efficiency was calculated by adding the abundance of the ECD fragment peaks, c_4^+ to c_{10}^+, and normalising with respect to the mass spectrum with the highest ECD fragment efficiency.

The ECD mass spectrum for the 4+ charge state of bovine insulin, oxidised B chain, is shown in Fig. 3. This mass spectrum was generated by averaging 200 time-domain transients and employing a 5-s electron irradiation time. This translates into a total spectral accumulation time of slightly over 23 min, which is about 5 times longer than for melittin. The backbone bond cleavage was reduced to 13 out of a possible 29 bonds, equivalent to 45% sequence coverage. The B chain of insulin is modified by oxidation of the two cystein residues at position 7 and 19, resulting in an overall mass increase of 95.97 Da. However, no ambiguities were encounted in assigning the peaks nor did any loss of modification occur. In spite of the 5-s electron irradiation time employed, the mass resolution on the precursor ion peak was almost 40000, indicating that the electron flux could be further increased without compromising the mass resolution or peak shapes.

Not surprisingly, the interresidue bond cleavage efficiency was observed to decrease, from substance P to bovine insulin, with increasing molecular weight due to the higher number of decomposition pathways and the larger degree of tertiary structure. Isolation of the 5+ charge state of bovine insulin, oxidised B-chain, was also attempted, but the signal magnitude was insufficient for assigning fragment ions.

To explore the limits of this 3 Tesla instrument, ECD was also attempted on bovine ubiquitin. In this case no mass isolation was carried out. The charge state distribution for this compound extended over the range 8+ to 13+. Despite averaging 200 time-domain transients, the resulting ECD

mass spectrum only showed very limited fragmentation (data not shown). The reduced species, charge state 7+, was visible, as well as a number of fragment ions, including c_8^+, c_{24}^{3+}, c_{27}^{3+}, c_{31}^{4+}, z_{17}^{2+}, z_{17}^{3+}, z_{26}^{4+}, y_{18}^{2+}, y_{18}^{3+}, y_{24}^{4+} and b_3^+. Furthermore, peaks due the expected small neutral losses accompanying ECD were also seen in the mass spectrum. Whilst some degree of ECD seemed to have occurred, full interpretation of the mass spectrum is not possible due to insufficient mass resolving power to enable all of the fragment ions to be identified. Therefore, this size of peptide appears to be beyond the capability of this particular 3 Tesla FTICRMS instrument.

Characterisation of the rhenium filament

In order to gain a better understanding of the ECD process, measurements of the ECD fragment efficiency as well as the effective kinetic energy distribution of the electrons when entering the ICR cell as a function of the filament voltage U⁻ were carried out. The ECD efficiency was measured for substance P. The fragment efficiency was obtained by adding the abundance of the ECD fragment peaks c_4^+ to c_{10}^+ in each spectrum (of 40 scans). The data were then normalised with respect to the mass spectrum with highest fragment efficiency. The initial, non-fragmented, precursor ion magnitude was approximately the same in all spectra. Figure 4 shows the ECD fragment efficiency for the 2+ charge state of substance P as a function of the filament voltage U⁻, confirming that maximum electron capture is observed at a filament voltage U⁻ just above +0.5 V. A third-
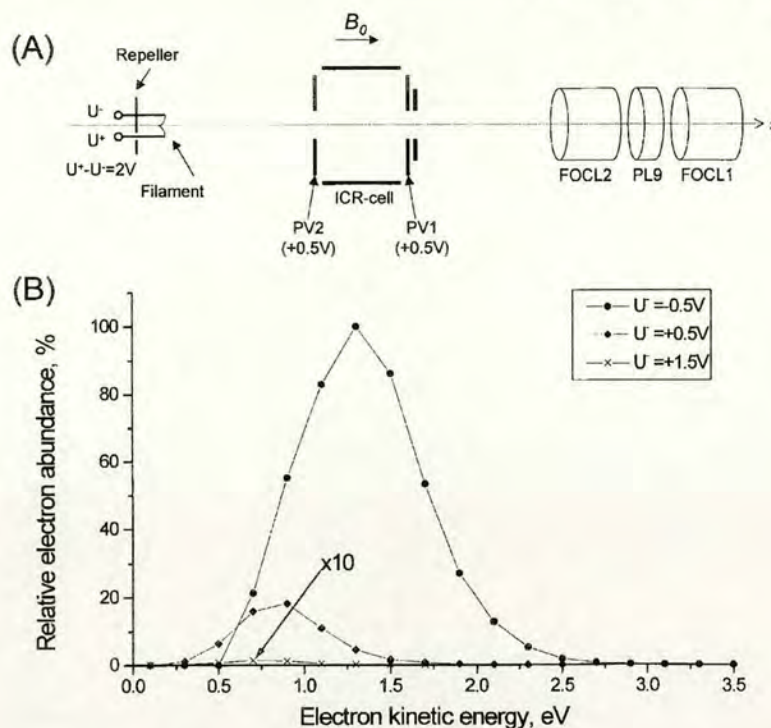
Figure 5. (A) Schematic diagram of the rhenium filament, ICR cell and the Einzel lens elements FOCL2, PL9 and FOCL1 on the APEX II instrument. The diagram is not drawn to scale. (B) The effective electron kinetic energy distribution at the rear trapping plate PV2 (held at +0.5 V during ECD experiments) for different filament voltages U⁻. In the case of the distribution shown for a filament voltage $U^- = +1.5$ V, the y-axis intensity has been multiplied by a factor of 10 for clarity.

order polynomial expression was fitted to the data to estimate the voltage corresponding to maximum electron capture (+0.52 V). When the filament was turned on in advance and allowed sufficient time to reach operating temperature, this parameter (the voltage corresponding to maximum electron capture) was observed to be highly reproducible (±0.1 V). Furthermore, baking out the ICR cell with the filament on at regular intervals maintained the reproducibility.

The use of low trapping plate voltages (+0.5 V) on this particular instrument was observed to increase ECD. Conversely, increasing the voltage on the trapping plates gave rise to a reduction in ECD; for example, no ECD was observed at trapping plate voltages of 1 V. This could be due to a combination of factors, including higher on-axis velocities of both the ions and electrons, as well as increased magnetron motion of the ions. Another possible reason could be higher loss of electrons to the rear trapping plate (PV2), when more positive trapping plate voltages are applied. Experimentally, however, the emitted electron current measured on PV2 was found to be very low (<50 pA), indicating that the electron beam was tightly focused and that few electrons were lost to this trapping plate. Furthermore, when measuring the emission current at the Einzel lens element FOCL2, higher currents were observed

at higher trapping plate voltages, suggesting higher extraction efficiency of the electrons.

One of the most important parameters for ECD is to employ low-energy electrons (<0.2 eV), which increases the electron capture cross-section. Since there is a potential energy difference between the filament, where the electrons are emitted, and the rear trapping plate,[14] this could be used as a means to alter the kinetic energy of the electrons entering the ICR cell in order to effect ECD.

Figure 5(A) shows a schematic diagram of the salient details of the APEX II instrument used in this work, showing the rhenium filament, ICR cell and the Einzel lens components FOCL2, PL9 and FOCL1. One side of the filament is held at a voltage U⁻ as well as the repeller plate. As a consequence of the potential drop across the filament, a voltage U⁺ is measured on the other side of the filament, which is 2 V higher than U⁻. The repeller plate is, therefore, negatively biased relative to U⁺. Some of the emitted electrons travel in the z-direction (the same direction as the magnetic field B) through the ICR cell to the Einzel lens shown. The filament and the ICR cell are in the bore of the superconducting magnet, while the Einzel lens is situated outside the bore, in the fringing magnetic field. The magnetic field lines diverge substantially in the fringing field region, enabling the emitted electrons to be detected on the Einzel

lens. Experimentally, an electron current could only be measured on the element FOCL2 of the Einzel lens. An attempt was made to increase the number of electrons striking the element FOCL2 by employing a 'reverse' Sidekick$^{\circledR}$. However, this had no effect on the measured electron current. This would imply that most of the emitted electrons from the filament that passed through the ICR cell could be reliably detected at the element FOCL2.

Since it was the aim of this study to gain an understanding of the energetics of the electrons involved in ECD, only those electrons that passed through the ICR cell were of interest. Therefore, only the velocity component of the emitted electrons in the z-direction is of interest. While it is known that electrons are emitted in a range of angles from a thermionic electron emitter,[15] the resulting x- or y-velocity components of the electrons are likely to be small due to axialisation in the strong magnetic field. No attempt was made in this work to measure the actual emitted electron kinetic energy profile from the filament. Since ECD takes place in the ICR cell, it is only the kinetic energy distribution of the electrons entering the ICR cell that is of relevance for electron capture.

The effective kinetic energy distributions of the electrons entering the cell are shown in Fig. 5(B). These distributions were obtained by measuring the total electron current detected on the Einzel lens element FOCL2, which was held at ground potential, whilst varying the potential difference between the filament and the trapping plate PV2 (the trapping plate closest to the filament). This potential energy ramp provided a means of 'retarding field' analysis. For these measurements, the voltage on the trapping plate PV2 was initially set to a sufficiently negative value. At this voltage there was no transmission of emitted electrons through the ICR cell to the Einzel lens element FOCL2. The trapping plate voltage was then incrementally increased to 0 V, and then to positive values until there was no further increase in the measured electron current. For a filament voltage U^- of −0.5 V, a voltage of −2.4 V on the rear trapping plate PV2 was found to be just sufficient to retard the electrons sufficiently so that no current was observed. This implies that even the most energetic electrons had an effective kinetic energy at PV2 of 0 eV. When using a voltage of +0.5 V on the trapping plate PV2, as during an ECD experiment, such an electron would have an effective kinetic energy of 2.9 eV [0.5 −(−2.4)] when entering the ICR cell. The experimentally measured current profiles, I_{FOCL2} versus the retarding field voltage, were differentiated. The data from these differentiated current profiles was then plotted as a function of the effective kinetic energy of the electrons at PV2. This kinetic energy is the difference between the voltage applied to the rear trapping plate PV2 for the retarding analysis and the actual voltage applied to the trapping plate during an ECD experiment (+0.5 V). Figure 5(B) shows effective kinetic energy distributions for different values of U^-, and hence U^+.

As can be seen in Fig. 5(B), at a filament voltage U^- of −0.5 V the most abundant electrons have an effective kinetic energy of 1.3 eV and, at a filament voltage U^- of +0.5 V, the most abundant electrons have an effective kinetic energy of 0.9 eV. The effective kinetic energy of the most abundant

electrons is higher when lower (i.e. more negative) filament voltages are used. Whilst the total emitted electron current is much larger at a filament voltage U^- of −0.5 V, compared to that at $U^- = +0.5$ V, the proportion of low-energy electrons is much greater for the kinetic energy distribution corresponding to a filament voltage of $U^- = +0.5$ V.

Currently, the abundance of the electrons in the low-energy tail of the kinetic energy distribution is not fully optimised. The only way to generate more low-energy electrons with the current approach is to reduce the trapping plate voltage, so that the kinetic energy of the electrons is effectively lowered since the retarding field is greater. Given that low trapping plate voltages were already employed (i.e. +0.5 V), there is not much scope left to further reducing the voltage on the trapping plates without compromising trapping of the ions.

CONCLUSIONS AND FUTURE WORK

It has been shown that a commercial 3 Tesla FTICR instrument (Bruker Daltonics APEX II) equipped with a conventional electron filament is capable of performing ECD on peptides with molecular weights up to 3.5 kDa. It has also been shown that in order to achieve ECD using this instrument one of the most important parameters is to employ low trapping plate voltages. The current electron source design, which is optimised for high-energy electron impact ionisation, limits the efficiency of ECD that can be achieved.

In future work the use of a modified filament design, incorporating a thorium coating, will be investigated. This should permit increased emitted electron flux from the filament at lower operating temperatures, which should decrease the achievable average electron kinetic energy. With this modified design it should be feasible, even on a 3 Tesla FTICR instrument, to generate sufficient ECD sequence information on peptides with molecular masses as high as 5 kDa. This would permit studies of post-translational modifications on most tryptic digest fragments, without the need to employ FTICR instruments equipped with larger, and more expensive, superconducting magnets.

REFERENCES

1. Zubarev RA, Kelleher NL, McLafferty FW. *J. Am. Chem. Soc.* 1998; **120**: 3265.
2. Kelleher NL, Zubarev RA, Bush K, Furie B, Furie BC, McLafferty FW, Walsh CT. *Anal. Chem.* 1999; **71**: 4250.
3. Mirgorodskaya E, Roepstorff P, Zubarev RA. *Anal. Chem.* 1999; **71**: 4431.
4. Stensballe A, Jensen ON, Olsen JV, Haselmann KF, Zubarev RA. *Rapid Commun. Mass Spectrom.* 2000; **14**: 1793.

Copyright © 2002 John Wiley & Sons, Ltd.

Rapid Commun. Mass Spectrom. 2002; **16**: 936–943

5. Shi SD-H, Hemling ME, Carr SA, Horn DM, Lindh I, McLafferty FW. *Anal. Chem.* 2001; **73**: 19.
6. Håkansson K, Cooper HJ, Emmett MR, Costello CE, Marshall AG, Nilsson CL. *Anal Chem.* 2001; **73**: 4530.
7. Loo JA, Udseth HR, Smith RD. *Rapid Commun. Mass Spectrom.* 1988; **2**: 207.
8. Little DP, Speir JP, Senko MW, O'Conner PB, McLafferty FW. *Anal. Chem.* 1994; **66**: 2809.
9. Zubarev RA, Horn DM, Fridriksson EK, Kelleher NL, Kruger NA, Lewis MA, Carpenter BK, McLafferty FW. *Anal. Chem.* 2000; **72**: 563.
10. Tsybin YO, Håkansson P, Budnik BA, Haselmann KF, Kjeldsen F, Gorshkov M, Zubarev RA. *Rapid Commun. Mass Spectrom.* 2001; **15**: 1849.
11. Haselmann KF, Budnik BA, Olsen JV, Nielsen ML, Reis CA, Clausen H, Johnsen AH, Zubarev RA. *Anal. Chem.* 2001; **73**: 2998.
12. Axelsson J, Palmblad M, Håkansson K, Håkansson P. *Rapid Commun. Mass Spectrom.* 1999; **13**: 474.
13. Lavarone AT, Jurchen JC, Williams ER. *Anal. Chem.* 2001; **73**: 1455.
14. Håkansson K, Emmett MR, Hendrickson CL, Marshall AG. *Anal. Chem.* 2001; **73**: 3605.
15. Hutson AR. *Phys. Rev.* 1955; **98**: 889.

Inert Site in a Protein Zinc Cluster: Isotope Exchange by High Resolution Mass Spectrometry

Claudia A. Blindauer,[†] Nick C. Polfer,[†] Stella E. Keiper,[†] Mark D. Harrison,[‡] Nigel J. Robinson,[‡]
Pat R. R. Langridge-Smith,[†] and Peter J. Sadler*,[†]

*School of Chemistry, University of Edinburgh, West Mains Road, Edinburgh, EH9 3 JJ, U.K., and
Department of Biosciences, University of Newcastle, Newcastle, NE2 4HH, U.K.*

Received September 6, 2002 ; E-mail: p.j.sadler@ed.ac.uk

It is well known that proteins control the local environment of bound metal ions,[1] and hence their thermodynamic and kinetic properties, for example, redox potentials[2,3] and transfer rates.[4] Metallothioneins (MTs) appear to play an important role in Zn homeostasis and the zinc buffer/distribution system.[5] Mammalian MTs contain Zn_3Cys_9 and Zn_4Cys_{11} clusters,[6] and metal exchange reactions for MTs are usually fast.[7] Bacterial MTs possess only a single zinc cluster,[8] $Zn_4Cys_9His_2$ in the case of the cyanobacterial MT SmtA (Figure 1).[9,10] We have investigated Zn exchange reactions of Zn_4-SmtA by a new method using stable isotope labeling combined with Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS). We show that the Zn_4 cluster of SmtA, in contrast to the structurally analogous cluster of mammalian MT, contains a kinetically inert Zn site, a feature which can be related to its secondary and tertiary structure, and which is of potential importance to its biological function.

Gentle ionization by electrospray (ESI) has previously been exploited for MTs,[11] and in combination with FT-ICR-MS[12] it is a powerful tool for the analysis of metalloproteins.[13] Deconvoluted ESI-FT-ICR spectra of Zn_4-SmtA containing Zn isotopes in natural abundance[14] and with 93% enrichment[15] with ^{67}Zn are compared in Figure 2. The observed experimental masses of the most intense isotopic peaks in Figure 2A and B (5862.95 and 5868.81 Da, respectively) are in good agreement with calculated values (5863.00 and 5869.00 Da: deviations of 8.5 and 32 ppm, respectively).

The effects of isotope enrichment are pronounced. Exchange of all four Zn atoms causes an increase in mass of the most abundant peak by 6 Da, and the isotopic envelope becomes much narrower (Figure 2B).

To investigate Zn exchange behavior, we incubated natural abundance Zn_4-SmtA with $^{67}ZnCl_2$ for various time intervals at 310 K, removed unbound Zn^{2+} by rapid gel filtration (ca. 3 min), and analyzed the product by FT-ICR-MS.[16] The amount of exchanged Zn at each time point was determined by comparing the experimental data to modeled isotope envelopes for $Zn_{x-4}{}^{67}Zn_xSmtA$ ($x = 1-4$, in 0.25 Zn intervals), taking into account the isotopic compositions of both natural abundance Zn and the ^{67}Zn-enriched $^{67}ZnCl_2$ used[14,15] (see Figure S2).

Crucially, the FT-ICR-MS measurements allow direct determination of the metal:protein ratio, without the need for separate measurements of metal and protein concentrations, as is necessary in radioisotope studies, while simultaneously confirming the identity of the intact metal−protein complex.

We find that initial Zn exchange is fast (ca. 1.4 Zn exchanged after 1 h, see Figure S3) as expected for metallothioneins,[7a] but most interesting is the extent of ^{67}Zn incorporation at equilibrium. If exchange occurred at all four Zn sites, the maximum achievable
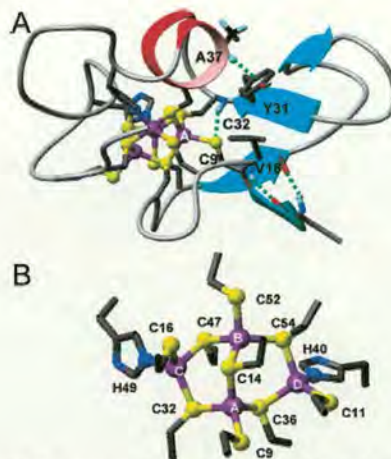

Figure 1. (A) 3D structure of Zn_4-SmtA (PDB 1JJD) showing elements of secondary structure around site A. The amide proton of Cys32 forms an H-bond to the sulfur of Cys9, which accounts for the extraordinary low-field shift of its ^1H resonance (10.01 ppm). The tertiary arrangement of the helix and sheet is further stabilized by the CH−π interaction between Ala37 and Tyr31. (B) The $Zn_4Cys_9His_2$ cluster of Zn_4-SmtA.
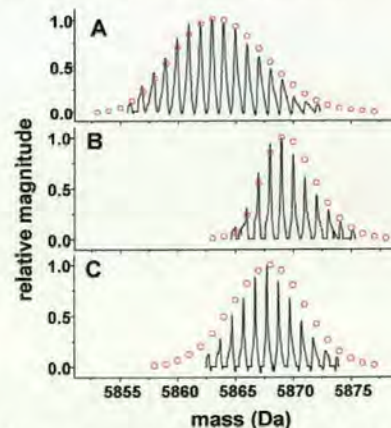

Figure 2. Deconvoluted ESI-FT-ICR mass spectra and modeled mass envelopes (red circles) of (A) natural abundance Zn_4-SmtA (first model circle is 1 Da below monoisotopic peak), (B) 93%-enriched $^{67}Zn_4$-SmtA, and (C) Zn_4-SmtA reacted with a 10-fold molar excess (with respect to Zn) of $^{67}ZnCl_2$ for 99 h at 310 K, and model for exchange of 2.75 Zn.

incorporation of ^{67}Zn with a 10-fold excess of 93% enriched ^{67}Zn would be 3.6 ^{67}Zn per mol SmtA. We observed a maximum incorporation of 2.75 ^{67}Zn (Figure 2C) after 99 h of incubation, close to the value of 2.7 ^{67}Zn calculated for exchange at only three sites. This implies that one of the four sites in the Zn_4 cluster (Figure 1B) is inert to exchange, a result consistent with our previous ^{111}Cd
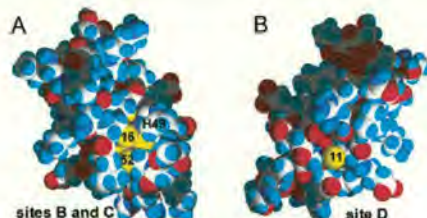
Figure 3. Space-filling models showing accessibility of the terminal Cys and His ligands in sites B, C, and D (Cys S yellow, N blue, O red, C white, H cyan). Site A is completely buried.
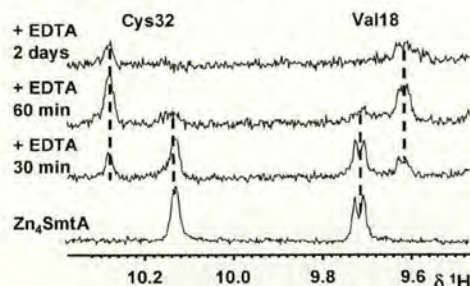

Figure 4. Low-field region (NH peaks) of the ^1H NMR spectrum of Zn_4-SmtA, and 30 and 60 min after reaction with ca. 6 mol equiv of EDTA, and 2 days after reaction with a further 10 mol equiv of EDTA. Zn removal was indicated by the decrease in intensity of the ethylene singlet of free EDTA ($\delta = 3.256$) and appearance of the analogous singlet for $[Zn(EDTA)]^{2-}$ ($\delta = 2.873$; Figure S5).

NMR observations[9] of reactions of $^{111}Cd^{2+}$ with Zn_4-SmtA, in which Zn(A) is inert. The formation of Cd_3Zn-SmtA is substantiated by ICP-AES and ESI-MS data (Figure S4). Normally, all-cysteine sites are thermodynamically stronger binding sites for Cd^{2+} as compared to Zn^{2+} (e.g., $10^2\times$ stronger for Cys_4 zinc fingers,[17] and $10^4\times$ stronger for mammalian MTs[18]), and Zn^{2+} replacement by Cd^{2+} in MTs is fast and stoichiometric.[19]

Zn exchange is likely to involve attack of $^{67}Zn^{2+}$ on an accessible ligand atom: S of Cys or N of His. It can be seen in Figure 3A and B that the terminal sulfurs in sites B, C, and D are accessible from the protein surface, but this is not the case for site A, for which all of the ligands (Cys 9, 14, 32, and 36) are buried.

Site A is surrounded by elements of secondary structure, an α-helix and two short antiparallel β-sheets, structural features which are not found in mammalian MTs, and these give rise to an H-bond between the S of Cys9, a ligand in site A, and the backbone NH of Cys32, a ligand in sites A and C (Figure 1A). Such an arrangement probably prevents intramolecular metal exchange into site A. The secondary structure around Zn(A) appears to be maintained during removal of Zn^{2+} from Zn_4SmtA by EDTA,[20] as indicated by the behavior of the low-field shifted NH resonances of Cys32 and Val18 (Figure 4; see also Figure 1A). It seems likely therefore that Zn-(B,C,D) are removed by EDTA more rapidly than Zn(A).

These findings show that the Zn_4 cluster in bacterial metallothionein confers novel properties on the protein. Site A and the surrounding secondary structure constitute a zinc finger fold of the kind found in GATA[21] and LIM[22] proteins which recognize other zinc finger proteins and DNA. Metal exchange and transfer reactions of proteins[4,23] are currently presenting important and challenging questions, and it is clear that FT-ICR-MS can make a major contribution to studies of both their thermodynamics and exchange dynamics.

Supporting Information Available: Experimental details, Figures S1—S5: FT-ICR-MS raw data, examples of fitted MS data, time-dependent zinc exchange, ESI mass spectra of Zn_4-SmtA before and after reaction with $CdCl_2$, ^1H NMR spectra of EDTA reaction (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) See, for example: *Metal Sites in Proteins and Models*; Hill, H. A. O., Sadler, P. J., Thomson, A. J., Eds.; Springer-Verlag: Berlin, Heidelberg, 1999.

(2) Babini, E.; Borsari, M.; Capozzi, F.; Eltis, L. D.; Luchinat, C. *J. Biol. Inorg. Chem.* **1999**, *4*, 692—700.

(3) Gray, H. B.; Malmström, B. G.; Williams, R. J. P. *J. Biol. Inorg. Chem.* **2000**, *5*, 551—559.

(4) (a) O'Halloran, T. V.; Culotta, V. C. *J. Biol. Chem.* **2000**, *275*, 25057—25060. (b) Cobine, P. A.; George, G. N.; Jones, C. E.; Wickramasinghe, W. A.; Solioz, M.; Dameron, C. T. *Biochemistry* **2002**, *41*, 5822—5829. (c) Drum, D. E.; Li, T.-K.; Vallee, B. L. *Biochemistry* **1969**, *8*, 3792—3797.

(5) Palmiter, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8428—8430.

(6) Romero-Isart, N.; Vašák, M. *J. Inorg. Biochem.* **2002**, *88*, 388—396.

(7) (a) Maret, W.; Larsen, K. S.; Vallee, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2233—2237. (b) Ejnik, J.; Munoz, A.; Gan, T.; Shaw, C. F., III; Petering, D. H. *J. Biol. Inorg. Chem.* **1999**, *4*, 784—790. (c) Otvos, J. D.; Liu, X.; Li, H.; Shen, G.; Basti, M. In *Metallothionein III*; Suzuki, K. T., Imura, N., Kimura, M., Eds.; Birkhäuser Verlag: Basel, Boston, Berlin, 1993; pp 57—109.

(8) Blindauer, C. A.; Harrison, M. D.; Robinson, A. K.; Parkinson, J. A.; Bowness, P. W.; Sadler, P. J.; Robinson, N. J. *Mol. Microbiol.* **2002**, *45*, 1421—1432.

(9) Blindauer, C. A.; Harrison, M. D.; Parkinson, J. A.; Robinson, A. K.; Cavet, J. S.; Robinson, N. J.; Sadler, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 9593—9598.

(10) Recombinant SmtA contains 55 amino acids (lacks N-terminal Met). Numbering scheme based on sequence deposited in Swissprot (www.expasy.ch). Similar zinc clusters are present in other cyanobacteria and pseudomonads (ref 8), and the key features of the inert Zn site are conserved in GatA from *E. coli*.

(11) (a) Gehrig, P. M.; You, C.; Dallinger, R.; Gruber, C.; Brouwer, M.; Kägi, J. H. R.; Hunziker, P. E. *Protein Sci.* **2000**, *9*, 395—402. (b) Hathout, Y.; Fabris, D.; Fenselau, C. *Int. J. Mass Spectrom.* **2001**, *204*, 1—6. (c) Jensen, L. T.; Peltier, J. M.; Winge, D. R. *J. Biol. Inorg. Chem.* **1998**, *3*, 627—631. (d) Merrifield, M. E.; Huang, Z.; Kille, P.; Stillman, M. J. *J. Inorg. Biochem.* **2002**, *88*, 153—172.

(12) Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325—1337.

(13) (a) Johnson, K. A.; Verhagen, M. F. J. M.; Brereton, P. S.; Adams, M. W. W.; Amster, I. J. *Anal. Chem.* **2000**, *72*, 1410—1418. (b) He, F.; Hendrickson, C. L.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 120—126. (c) Taylor, P. K.; Parks, B. Y. A.; Kurtz, D. M.; Amster, I. J. *J. Biol. Inorg. Chem.* **2001**, *6*, 201—206.

(14) Natural abundances of Zn isotopes: ^{64}Zn 48.6%, ^{66}Zn 27.9%, ^{67}Zn 4.1%, ^{68}Zn 18.8%, ^{70}Zn 0.6%.

(15) $^{67}Zn_4$-SmtA was prepared from apo-SmtA and $^{67}ZnCl_2$ (93.11% ^{67}Zn, remainder ^{64}Zn 1.37%, ^{66}Zn 2.58%, ^{68}Zn 2.89%, ^{70}Zn 0.05% supplied by Oak Ridge National Laboratory, TN, as ZnO and dissolved in HCl).

(16) Zn_4-SmtA (200 μM) in 10 mM ammonium acetate pH 7.4 was incubated with a 10-fold molar excess (with respect to Zn) of ^{67}Zn (8 mM) at 310 K. Unbound Zn was removed from 20 or 40 μL aliquots at various time intervals on a Pharmacia PD10 column using 10 mM NH_4Ac as eluant. The eluate (3.5 mL) was concentrated to ca. 23 μM in protein using an Amicon YM3 filter before FT-ICR-MS analysis. The final sample solutions (10 μM) also contained ubiquitin (5 μM), 30% MeOH, and 0.05% formic acid.

(17) Krizek, B. A.; Merkle, D. L.; Berg, J. M. *Inorg. Chem.* **1993**, *32*, 937—940.

(18) Kägi, J. H. R. In *Metallothionein III*; Suzuki, K. T., Imura, N., Kimura, M., Eds.; Birkhäuser Verlag: Basel, Boston, Berlin, 1993; pp 29—56.

(19) Nettesheim, D. G.; Engeseth, H. R.; Otvos, J. D. *Biochemistry* **1985**, *24*, 6744—6751.

(20) 0.5 mM Zn_4-SmtA in 50 mM Tris-HCl, 50 mM NaCl, pH 7.0 (degassed and saturated with N_2), was incubated with EDTA at 298 K.

(21) Omichinski, J. G.; Clore, G. M.; Schaad, O.; Felsenfeld, G.; Trainor, C.; Appella, E.; Stahl, S. J.; Gronenborn, A. M. *Science* **1993**, *261*, 438—446.

(22) Perez-Alvarado, G. C.; Miles, C.; Michelsen, J. W.; Louis, H. A.; Winge, D. R.; Beckerle, M. C.; Summers, M. F. *Nat. Struct. Biol.* **1994**, *1*, 388—398.

(23) Outten, C. E.; O'Halloran, T. V. *Science* **2001**, *292*, 2488—2492.

JA0284409