Negative Ion Mass Spectrometry of Peptides

An Aid to Structure Determination

A thesis submitted for the degree of Doctor of Philosophy (Ph.D.)

by

Daniel Bilusich

B.Sc. (Hons.)

Department of Chemistry

THE UNIVERSITY OF ADELAIDE
AUSTRALIA

May 2006

Contents

Contents ..... **ii**
Abstract ..... **v**
Statement of Originality ..... **vi**
Acknowledgements ..... **vii**
20 Common Amino Acids ..... **viii**

Chapter 1: Introduction to Amphibians and their Peptides **1**
1.1 Preface 1
1.2 Australian Amphibians 1
1.3 Peptide Biosynthesis 2
1.4 Amphibian Peptide Activity 4
1.5 Collection of Frog Secretions 7
1.6 High Performance Liquid Chromatography 9
1.7 Peptide Sequencing 10
1.7.1 Mass Spectrometry 10
1.7.2 Lys-C Digestion 11
1.7.3 Automated Edman Sequencing 12
1.7.4 C-Terminal End Group Determination 14

Chapter 2: Introduction to Positive and Negative Ion Mass Spectrometry **15**
2.1 The Mass Spectrometers 15
2.1.1 Q-Tof 2 Mass Spectrometer 15
2.1.2 LCQ Mass Spectrometer 17
2.2 Electrospray Ionisation 19
2.3 Positive Ion Mass Spectrometry 21
2.4 Negative Ion Mass Spectrometry 24
2.4.1 Amide Backbone Fragmentations 25
2.4.2 The $\beta^{|}$ Backbone Cleavage 29
2.4.3 Characteristic Side Chain Fragmentations 32
2.4.4 Side Chain Induced Backbone Cleavages 34
2.4.5 γ and δ Cleavages of Glu and Gln 37
2.4.6 Ser and Thr Side Chain and Backbone Cleavages 41
2.5 Theoretical Calculations 45

Chapter 3: Negative Ion Cleavages of the Cys Residue 48
3.1 Introduction 48
3.2 Results and Discussion 51
3.2.1 The Side Chain Loss of H_2S From Cys 51
3.2.2 Ser Side Chain Loss Re-Investigated 53
3.2.3 Fragmentation of C-Terminal Cys 56
3.2.4 γ and δ Cleavages of Cys 62
3.3 Summary and Conclusions 68
3.4 Experimental 70
3.4.1 Peptide Synthesis 70
3.4.2 Mass Spectrometry 70
3.4.3 Theoretical Calculations 70

Chapter 4: Negative Ion Cleavages of Intramolecular Disulfide Bridged Peptides 72
4.1 Introduction 72
4.1.1 Intramolecular Disulfide Bridged Peptide Sequencing by Enzymatic Cleavage Followed by Positive Ion Mass Spectrometry 76
4.1.2 Intramolecular Disulfide Bridged Peptide Sequencing by Disulfide Reduction/Alkylation Followed by Positive Ion Mass Spectrometry 77
4.2 Results and Discussion 80
4.2.1 Loss of H_2S_2 from Intramolecular Disulfide Bridged Peptides 80
4.2.2 Loss of H_2S_2 and CO_2 from C-Terminal Intramolecular Disulfide Bridged Peptides 88
4.2.3 Multiple Disulfide Bridged Peptides 103
4.3 Summary and Conclusions 107
4.4 Experimental 109
4.4.1 Peptide Synthesis 109
4.4.2 Mass Spectrometry 109
4.4.3 Theoretical Calculations 110

Chapter 5: Internal Loss of Val from Citropin 1.1 and some Synthetic Analogues..111
5.1 Introduction....111
5.1.1 Ile21/Glu23 Cyclisation and Backbone Cleavage....112
5.1.2 Internal Loss of Val from Citropin Peptides....116
5.2 Results and Discussion....117
5.3 Summary and Conclusions....130
5.4 Experimental....131
5.4.1 Peptide Synthesis....131
5.4.2 Mass Spectrometry....131
5.4.3 Theoretical Calculations....132

Chapter 6: Studies of the Dermal Secretions of *Litoria peronii*....133
6.1 Introduction....133
6.2 Results and Discussion....136
6.2.1 Isolation of Peptides....136
6.2.2 Peptide Sequence Determination....139
6.2.3 Peptide Activity....153
6.3 Summary and Conclusions....156
6.4 Experimental....157
6.4.1 Collection of Glandular Secretions....157
6.4.2 HPLC Separations of Peptide Secretions....157
6.4.3 Mass Spectrometry....158
6.4.4 Lys-C Enzymatic Digestions....158
6.4.5 Automated Edman Sequencing....158
6.4.6 C-Terminal End Group Determination....159
6.4.7 Synthetic Peptides....159
6.5.8 Smooth Muscle Activity Testing....159

References....161
Appendix....172
Publications....178

Abstract

Amphibians contain a rich chemical arsenal in their skin glands, vital for their defence against predators, both large and small. The peptides secreted by the frogs have a range of biological activities. These include both antibacterial and anticancer activity, others are neuropeptides, while some exhibit antifungal and antimalarial activity.

Peptides are usually sequenced using positive ion mass spectrometry (MS). However, negative ion MS can also provide valuable sequencing information. Under negative ion MS conditions, the presence of a Cys residue is immediately identified by the facile side chain loss of H_2S. The position of the Cys residue is determined by the formation of a side chain induced backbone cleavage ion. When a Cys residue is in the C-terminal position of a peptide, the spectrum is dominated by the loss of both H_2S and CO_2.

Negative ion MS can also be used to identify and sequence intramolecular disulfide bridged peptides. The disulfide bridge is immediately identified by the facile loss of H_2S_2 from the parent anion. Once the disulfide bridge is cleaved, further amide cleavages provide much of the sequence of the peptide, including the residues originally within the disulfide link. When one of the disulfide bridged Cys residues is in the C-terminal position, the major fragmentation is the loss of H_2S_2 and CO_2 from the parent ion.

The negative ion mass spectra of citropin 1.1 and synthetically modified analogues show an unusual loss of an internal Val residue from the $(M-H)^-$ parent ion. This rearrangement requires the decomposing anion to have an α-helical structure.

The skin secretions of *Litoria peronii* or Peron's Tree Frog contain five novel peptides which have been named peroneins. Four pro-peroneins are present in the summer secretions only. The biologically active peptides caerulein 1.1, caerin 1.1 and caerin 2.1 are also present in the glandular secretions.

Statement of Originality

This thesis contains no material which has been accepted for the award of any other degree, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference is made in the text.

I give my consent to this copy of my thesis, when deposited in the University Library, to be available for loan and photocopying.

Daniel Bilusich Date

Acknowledgements

Firstly I would like to thank my supervisor Prof. John H. Bowie for all his guidance and assistance throughout the course of my Ph.D. I particularly appreciate his enthusiasm and dedication to research, and also for allowing me to be involved in such an interesting project.

Next I would like to thank the past and present members of the Bowie group whom have helped me immensely over the past few years. Particular mention goes to Dr. Andrew M. McAnoy, Dr. Craig S. Brinkworth, Dr. Mark Fitzgerald, Tara L. Pukala, Vita M. Maselli, Micheal J. Maclean and Rebecca J. Jackway.

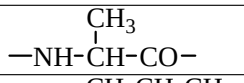
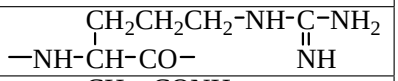
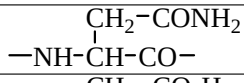
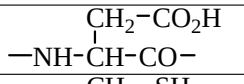
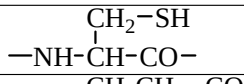
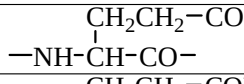
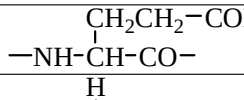
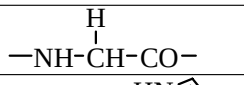
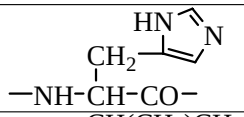
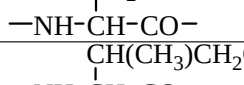
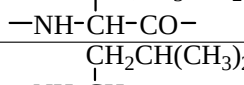
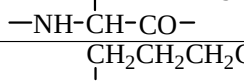
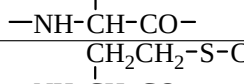
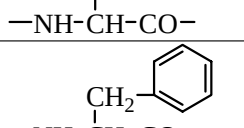
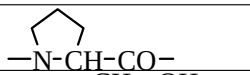
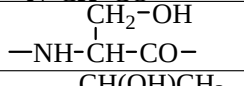
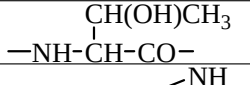
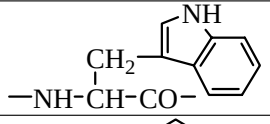
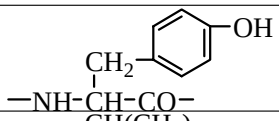
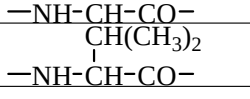
The proof reading of my thesis during its final stages by Prof. John H. Bowie, Dr. Simon M. Pyke and Tara L. Pukala is greatly appreciated, thank you.

I must also thank the academic, teaching and technical staff of the Department of Chemistry, The University of Adelaide, for their assistance over the past eight or so years during my bachelor, honours and Ph.D. degrees.

A thank you must also go to Prof. Albert T. Lebedev, Dr. Tatiana Samguina, Dr. Vladislav V. Lobodin and Konstantin A. Artemenko for being such great hosts during my stay at Moscow State University (Московский Государственный Университет).

Finally, a special thank you to my family and friends whom have shown a continuous interest in my research and its progress.

20 Common Amino Acids

Amino Acid	Abbreviation		Integral Mass	Structure
	3 Letter	1 Letter		
Alanine	Ala	A	71	
Arginine	Arg	R	156	
Asparagine	Asn	N	114	
Aspartic Acid	Asp	D	115	
Cysteine	Cys	C	103	
Glutamic Acid	Glu	E	129	
Glutamine	Gln	Q	128	
Glycine	Gly	G	57	
Histidine	His	H	137	
Isoleucine	Ile	I	113	
Leucine	Leu	L	113	
Lysine	Lys	K	128	
Methionine	Met	M	131	
Phenylalanine	Phe	F	147	
Proline	Pro	P	97	
Serine	Ser	S	87	
Threonine	Thr	T	101	
Tryptophan	Trp	W	186	
Tyrosine	Tyr	Y	163	
Valine	Val	V	99	

Nominal masses of peptides are used throughout this thesis. Nominal masses are obtained by summation of the integral masses of the amino acid residues.

Chapter 1:

Introduction to Amphibians and their Peptides

1.1 Preface

Pools of stagnant water are breeding grounds for bacteria and micro-organisms, but despite this, amphibians are able to thrive in environmental conditions such as these whilst remaining unaffected.[1] This is due to the protection provided by the array of peptides and toxins released by dermal glands dispersed over the anuran's skin.[2] With bacteria becoming increasingly resistant to currently available antibiotics[3] the search for novel agents to combat bacteria has increased in intensity over recent years.[4,5]

During the past 15 years, the dermal gland secretions of over 30 species of Australian anurans have been studied, with the isolation of over 200 different peptides.[6] Many of these peptides have biological activity in the form of narrow and wide spectrum antibiotics and anticancer agents. Some display smooth muscle activity, while others are analgesics thousands of times more potent than morphine.[6]

1.2 Australian Amphibians

Over 4000 species of frogs have been identified worldwide, on every continent except Antarctica. Of these, just over 200 species have been identified in Australia[7] which accounts to approximately 5% of the world's known species. The huge diversity among Australian amphibians is a result of the vastness of the Australian land mass and differing

climatic regions. Australia can be divided into four main zoogeographical areas as shown in Figure 1.1. The Torresian region contains many tropical rainforests, while the Eyrean region is predominantly desert. The South-western and Bassian areas have similar climates but are separated by a vast desert which has resulted in a diversification of species in these two regions. Many species of frogs are found in localised areas within these four regions. However, some species have dispersed over extensive regions of Australia, mainly as a result of the absence of natural barriers such as large mountain ranges and vast inland seas.[7]

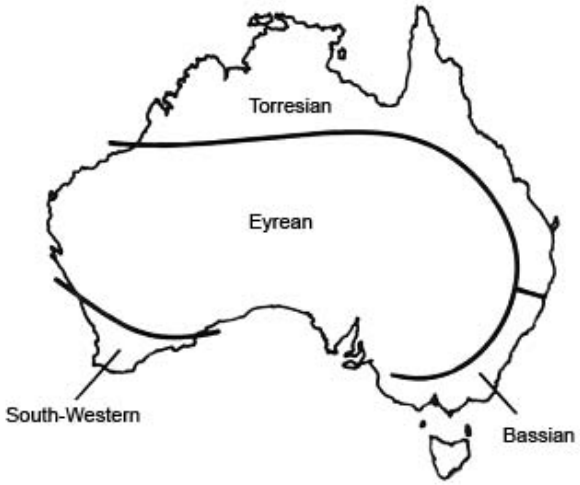


Figure 1.1. The four zoogeographic regions of Australia.[7]

1.3 Peptide Biosynthesis

The biologically active peptides amphibians secrete onto their skin are only a small portion of a much larger inactive peptide. The original translated inactive peptide is known as a prepropeptide (Figure 1.2).[8] It consists of three regions, the signal, spacer and active peptide. Once synthesised in the endocrine system, the signal portion directs the

prepropeptide to secretory glands. The pre portion is enzymatically removed and the inactive propeptide stored until required. Just prior to secretion from the dermal glands onto the skin surface, the pro portion is removed by an endoprotease. Further activation of some peptides by post-translational modifications are also common. Such modifications include the protection of the N-terminal end of the peptide, modifications of side groups such as sulfation of Tyr residues, conversion of L-amino acids to D-amino acids and amidation of the C-terminal end of the peptide. The active peptide is then secreted onto the skin surface via the secretory ducts.[9-11]

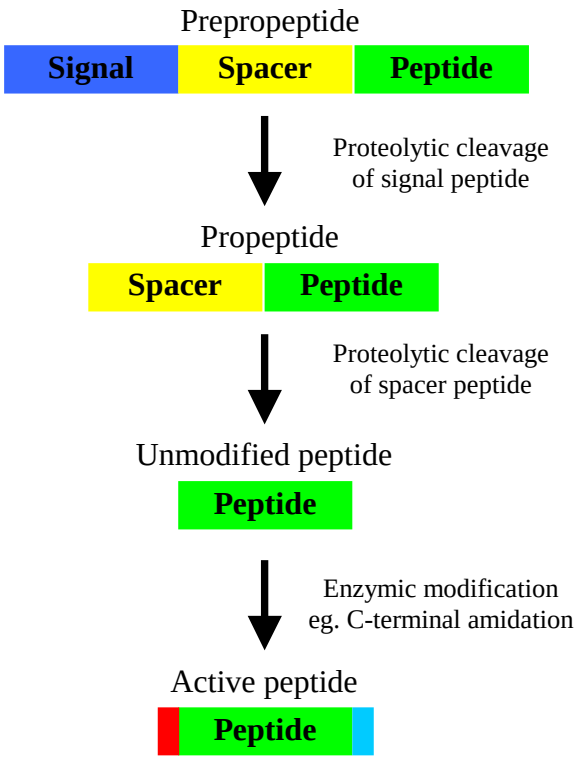


Figure 1.2. Synthesis of a bioactive peptide from the prepropeptide precursor.

1.4 Amphibian Peptide Activity

The peptides produced by amphibians have a wide range of activities. Some have antifungal and antimalarial activities,[12] while splendipherin is a male sex pheromone isolated from the anuran, *Litoria splendida*.[13]

Antibacterial peptides are essential for anurans to be able to survive in and around stagnant pools of water; ideal breeding grounds for bacteria.[1] These peptides work by disrupting the cell membrane of bacterial cells.[8] There are two main mechanisms by which this is achieved. An amphipathic α-helical secondary structure of the peptide is essential for both cases. The first mechanism, known as the barrel-stave model, is where the peptides aggregate on the membrane surface (see Figure 1.3). This occurs due to the electrostatic attraction between charged residues and the ionic sites on the lipid bilayer. Peptides then insert through the membrane to form a trans-membrane pore.[14-16]

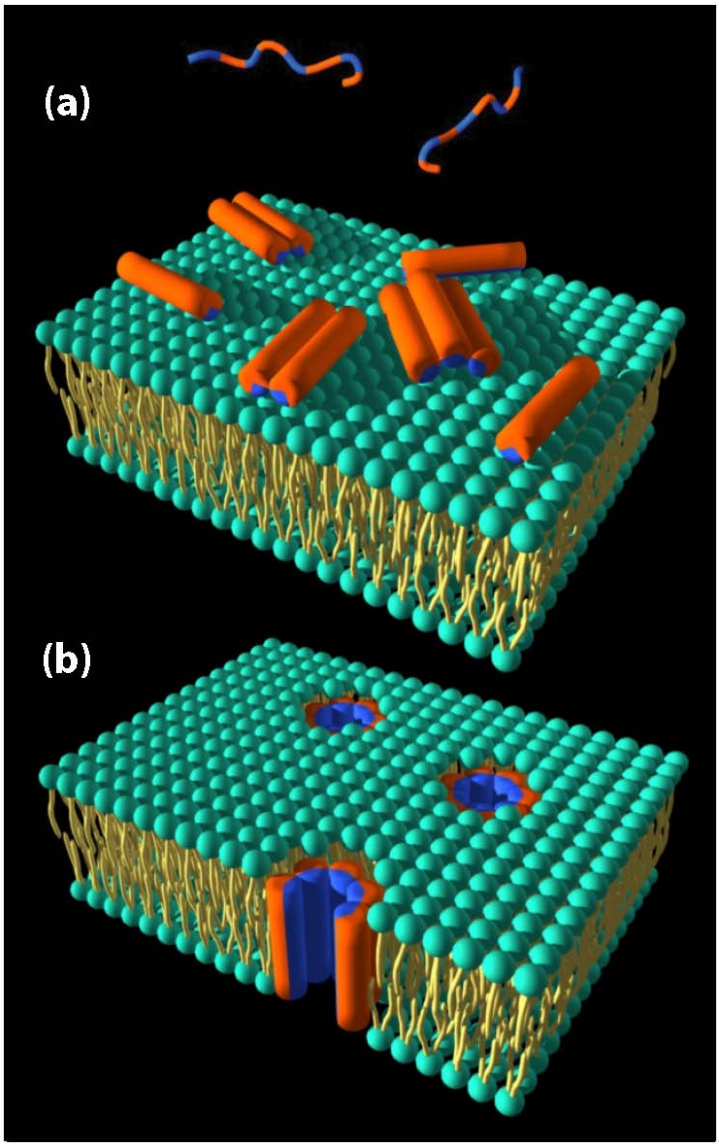


Figure 1.3. The barrel-stave mechanism: (a) aggregation of peptides on the membrane surface and (b) insertion of peptides through the bilayer to form trans-membrane pores. Peptides in α-helical form are represented by cylinders with hydrophilic and hydrophobic regions blue and red respectively.[17]

The carpet mechanism on the other hand involves peptides binding with their α-helical axis parallel with the membrane to form a carpet-like monolayer (Figure 1.4). Once the concentration of peptide reaches a critical point, transient holes are produced due to the strain on the bilayer curvature.[18-20]

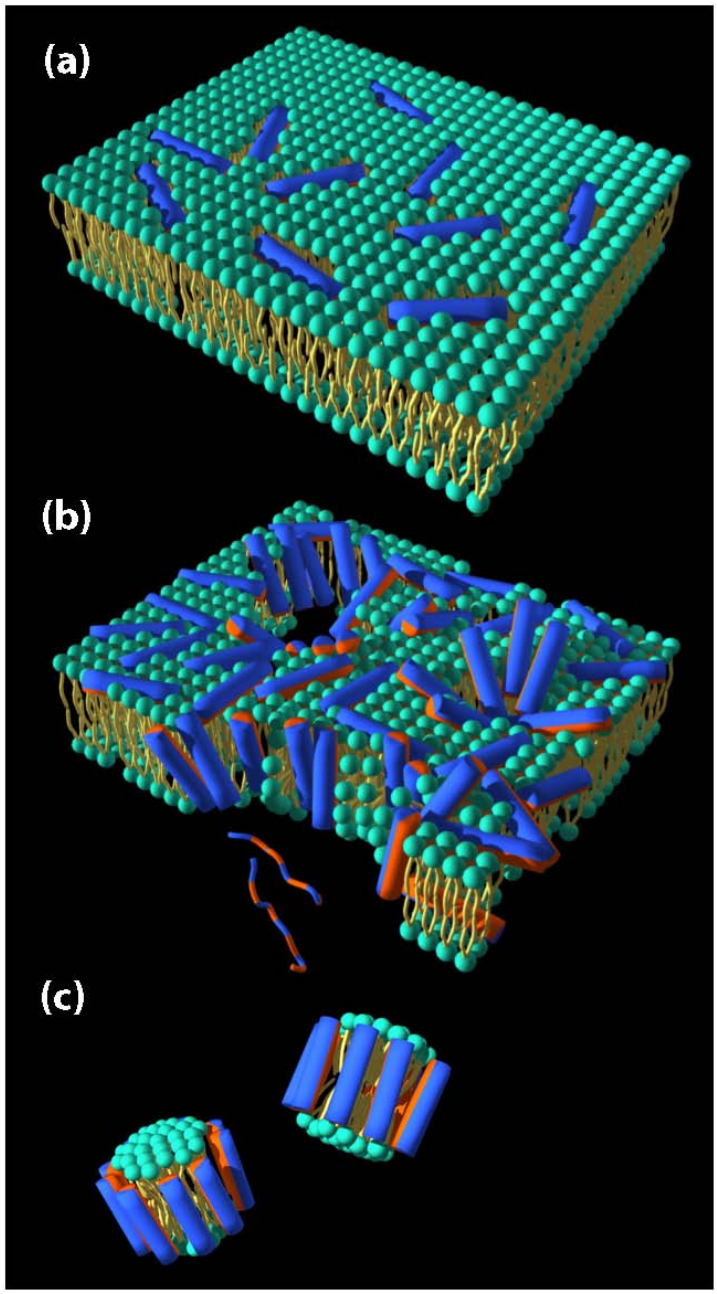


Figure 1.4. The carpet mechanism: (a) binding of peptides to the membrane to form a carpet-like monolayer. (b) Once the threshold concentration is reached, transient holes are produced with (c) the potential for membrane segments to break off as micelles. Peptides in α-helical form are represented by cylinders with hydrophilic and hydrophobic regions blue and red respectively.[17]

Either way, the disruption to the membrane results in a flux of ions and small molecules across the membrane, resulting in cell lysis. Many of these peptides also have anticancer properties due to similar modes of action in bacterial and cancer cells.[12] In fact, following secretion of these peptides onto the amphibian skin (5-30 min), an endoprotease degrades and deactivates these peptides to prevent them from destroying the anuran's cells.[21]

Many amphibians also produce neuropeptides. These peptides are responsible not only for defensive purposes but also for the regulation of dermal physiological action, enhancing blood circulation, thermoregulation and some are analgesics thousands of times more potent than morphine.[22-24] Other activities include the contraction of smooth muscle in predators[23] or inhibition of neuronal nitric oxide synthase (nNOS), the regulator of the cell signalling agent nitric oxide (NO).[12]

1.5 Collection of Frog Secretions

The original method for collecting the secretions of anurans involved the capture of many frogs which were then killed and their skins removed to be dried. Once dry, the bioactive material could be collected by extracting the skins with solvent.[25-27] A new method was developed in 1992 which involves a mild electrical current being applied to the frog's dorsal surface.[28] This method is known as the surface electrical stimulation (SES) method, and results in the release of glandular secretions onto the surface of the skin in the same manner as if the frog was defending itself against predators.

The SES method involves holding the frog by its hind legs and applying the electrical stimulus until the secretions appear on the dorsal surface (Figure 1.5). The secretions are washed off with deionised water and collected. Some of the peptides released by the amphibian are cytotoxic not only to predators but to the amphibian itself. The skin secretion also contains enzymes which deactivate the toxic peptides within 5 to 30 minutes following secretion.[21] An equal volume of methanol must therefore be added to the deionised water and secretion mixture to denature these enzymes and stop them degrading the active peptides of interest.[29] The surface electrical stimulation method can be repeated on a monthly basis without harming the anurans. This eliminates the need to spend days searching for hundreds of frogs and ultimately killing them, and allows studies to be conducted with only several or even as little as one frog.



Figure 1.5. The surface electrical stimulation (SES) method of secretion collection and the crude material collected.

1.6 High Performance Liquid Chromatography

Once the secretions have been collected they are concentrated, and the various peptides separated and purified by high performance liquid chromatography (HPLC). Peptide separation requires a reverse phase column and a gradient elution where the concentration of organic solvent increases over time. The secretion is injected into the column and the peptides equilibrate between the stationary and mobile phases. The more hydrophobic the peptide, the stronger it adsorbs onto the stationary phase of the column. As the organic concentration in the mobile phase is increased it will reach a point where the least hydrophobic peptide will equilibrate in favour of the mobile phase and elute through the column to be collected.[30] As the organic concentration continues to rise more peptides will elute off the column with the most hydrophobic peptide coming off last. The eluting peptides are detected using an ultraviolet detector at λ = 214 nm which measures the $\pi \rightarrow \pi^*$ transition of the amide carbonyl group.

Acetonitrile is used as the organic modifier due to its low viscosity and ultraviolet transparency. The aqueous solvent contains trifluoroacetic acid (TFA) which maintains a low pH and minimises ionic interactions between the peptide and stationary phase.

Originally the B fragment ions were proposed to be acylium cations as shown in Figure 1.6.[31] More recent publications however suggest that B ions are predominantly the more stable oxazolones,[161-165] although additional cyclic structures have also been suggested.[166,167] The proposed mechanism for the formation of a cyclic oxazolone B ion is shown in Figure 1.6A.

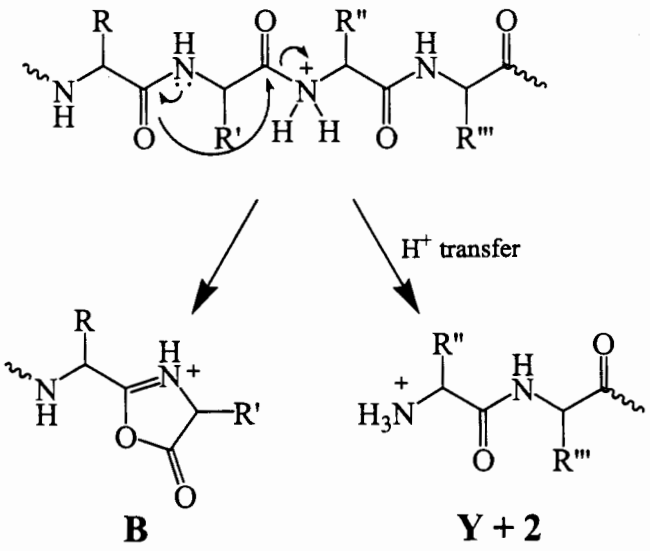


Figure 1.6A. Formation of oxazolone B and Y + 2 fragment ions of peptides in positive ion mass spectrometry.[161-165]

1.7 Peptide Sequencing

1.7.1 Mass Spectrometry

Mass spectrometry is a technique (requiring very little sample) used to determine both the molecular mass of a peptide, and more importantly, the sequence of the peptide. Collision induced dissociation (CID) in the positive ion mode is the most common method used to cleave the amide backbone bonds, to form B and Y + 2 fragmentation ions as shown in Figure 1.6.[31,32]

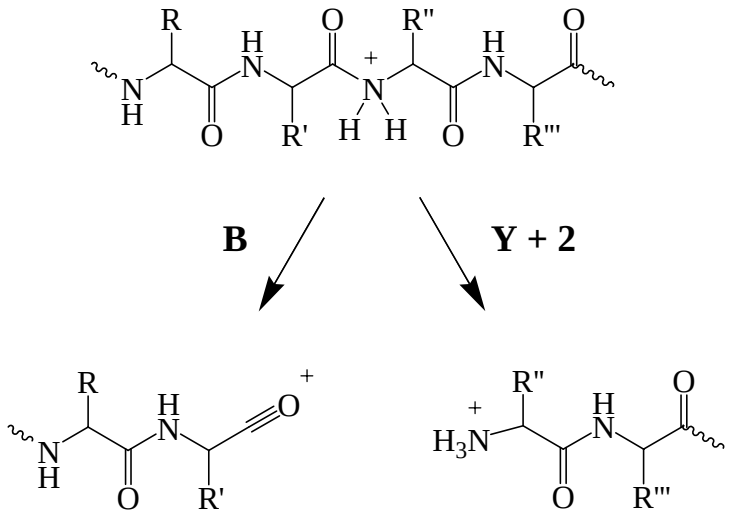


Figure 1.6. B and Y + 2 fragmentations of peptides in positive ion mass spectrometry.

The B fragmentation ions provide sequencing information from the C-terminal end of the peptide and the Y + 2 fragmentation ions from the N-terminal end.[33] One drawback of positive ion mass spectrometry has been the inability to distinguish between isomeric Leu and Ile (both *m/z* 113) and isobaric Lys and Gln (both *m/z* 128). A more detailed examination of positive and negative ion fragmentations is discussed in Chapter 2.

1.7.2 Lys-C Digestion

The inability of low resolution positive ion mass spectrometry to distinguish between isobaric Lys and Gln has necessitated the development of a new experiment to provide this sequencing information. Enzymatic digestion using a Lys-C endoprotease cleaves Lys containing peptides by hydrolysing the amide bond at the C-terminal side of a Lys residue (Figure 1.7).[34] Following digestion, the sample is submitted to mass spectrometric analysis to detect both the number of Lys residues present and their positions within the peptide.[11]

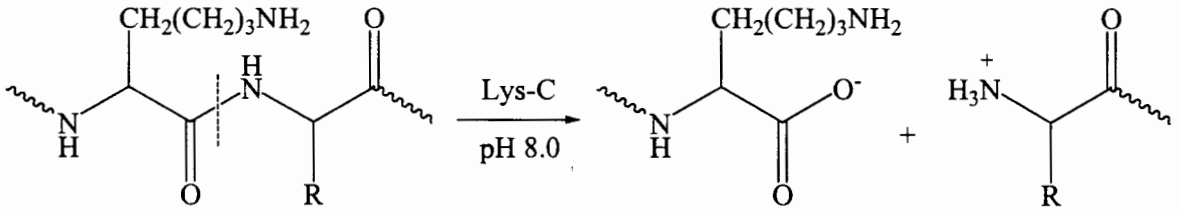


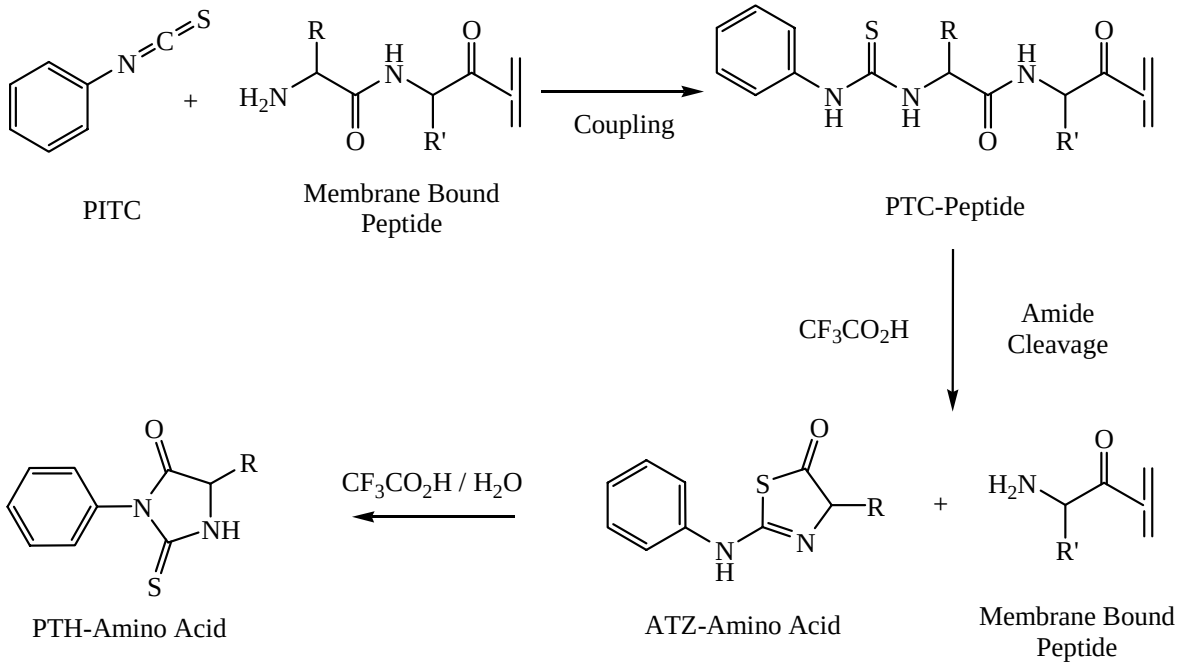
Figure 1.7. Lys-C digestion at the C-terminal Lys amide bond.

Enzymic digest is also used for large peptides and proteins to produce smaller peptides that are easier to work with. When two different endoproteases are used the fragments formed will overlap, allowing the entire sequence of the peptide or protein to be elucidated.[35]

An alternative method of differentiating between Lys and Gln involves the acetylation of Lys residues followed by mass spectrometric analysis. The Lys residue increases from 128 Da to 170 Da upon acetylation.[52]

1.7.3 Automated Edman Sequencing

Edman degradation is a technique used to determine the full sequence of a peptide.[36] An outline of the Edman degradation cycle is shown in Scheme 1.1. It involves binding the C-terminal end of the peptide to a membrane. This is followed by the addition of phenylisothiocyanate (PITC) to the free amine at the N-terminus of the peptide to form a phenylthiocarbamyl-peptide (PTC-peptide) complex. The first amino acid is cleaved at the amide bond under acidic conditions to afford the anilinothiazolinone (ATZ) derivative leaving the original peptide, minus the first N-terminal amino acid, bound to the membrane. The relatively unstable ATZ-amino acid rearranges to the more stable phenylthiohydantoin (PTH)-amino acid under aqueous acidic conditions and is then subjected to HPLC. The retention time of the PTH-amino acid complex identifies the amino acid residue. The cycle is repeated and can be automated to afford the entire sequence of the peptide from the N-terminal end. This technique is able to distinguish between isobaric Lys and Gln and also isomeric Leu and Ile. Automated Edman sequencing is most reliable for peptides less than 30 amino acid residues in length. Once the length increases beyond 30 residues there is a risk of the peptide washing away from the membrane because of the large number of cycles required.[31] This may be overcome by forming smaller fragments by enzymic digest, and then combining the sequence information for the individual fragments to provide the full sequence of the peptide or protein.



Scheme 1.1. The process by which the N-terminal amino acid residue is cleaved during automated Edman degradation.

There may also be difficulties in assigning post-translationally modified residues as they may have unknown retention times. Residues such as a N-terminal pyro-Glu prevent the addition of PITC to the N-terminal amino group and therefore no amino acids from that peptide can be assigned.[31] For this reason automated Edman sequencing is a complementary technique to mass spectrometry. A combination of both techniques generally provides the full sequence of a peptide.

1.7.4 C-Terminal End Group Determination

The C-termini of peptides often undergo post-translational modification from a carboxylic acid to a primary amide. Mass spectrometry can be used to distinguish between the two by converting the carboxylic acid or amide to a methyl ester and then re-measuring the mass of the peptide. Converting a carboxylic acid to a methyl ester increases the mass by 14 Da, whilst converting an amide to a methyl ester results in a 15 Da mass rise (Figure 1.8).[37,38]

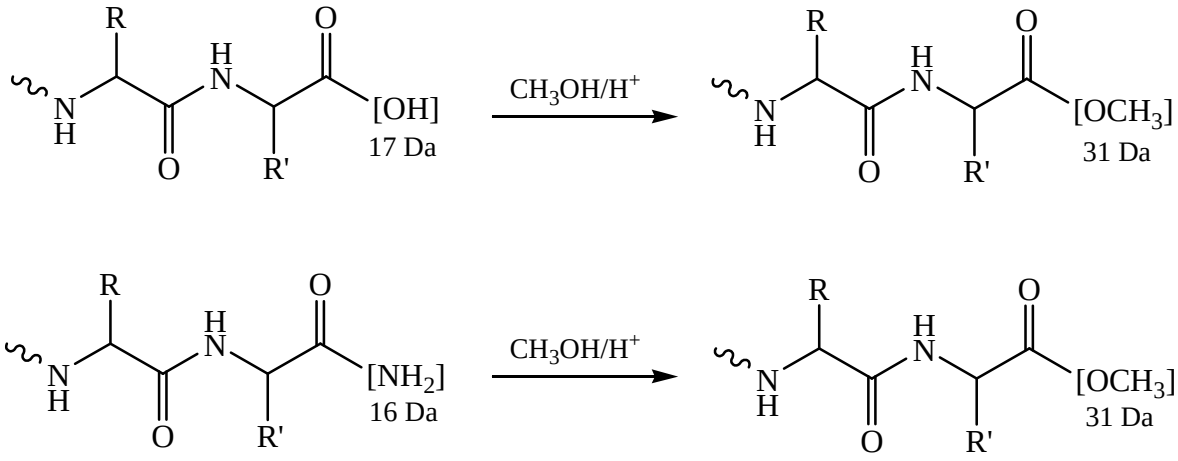


Figure 1.8. C-terminal end group determination by methylation and mass spectrometry.

Chapter 2:

Introduction to Positive and Negative Ion Mass Spectrometry

2.1 The Mass Spectrometers

2.1.1 Q-Tof 2 Mass Spectrometer

The original mass spectrometers consisted of a single magnetic sector followed by a detector.[39] These spectrometers were replaced with double focussing instruments which had both a magnetic and electric sector and provided better resolution.[40,41] One of the mass spectrometers used for this study is the Micromass Q-Tof 2 mass spectrometer. It consists of two sectors, the quadrupole and the time of flight tube, as shown in Figure 2.1. The ions are generated by the electrospray ionisation technique in a Z-Spray configured source which allows only ions to pass through the cone and enter the spectrometer. The ions are then accelerated by a potential difference. Adjusting the cone voltage can favour the selection of different mass to charge ratio (*m/z*) ions. The larger the voltage, the larger the *m/z* ions which are attracted in through the cone. Once the ion is accelerated through the potential difference it is focused into a beam before passing through the quadrupole analyser where the ions are differentiated by mass to charge ratio. In the MS mode, all the ions are allowed to pass through the quadrupole analyser to the pusher, which accelerates ions so that they all possess identical translational energy. The ions are collected at the pusher before they are accelerated so that they all leave from this point at the same time.

The ions are accelerated by a potential difference down the time of flight tube until they reach the reflector which decelerates the ions and reaccelerates them to the detector. The ions have the same translational energy and because translational energy is proportional to both mass and velocity, the heavier the ion the slower its velocity relative to a lighter ion. This means the heavier ion takes longer to reach the detector. The difference in time is transformed into a *m/z* value for that ion because the distance from the pusher to the detector is the same for all ions.

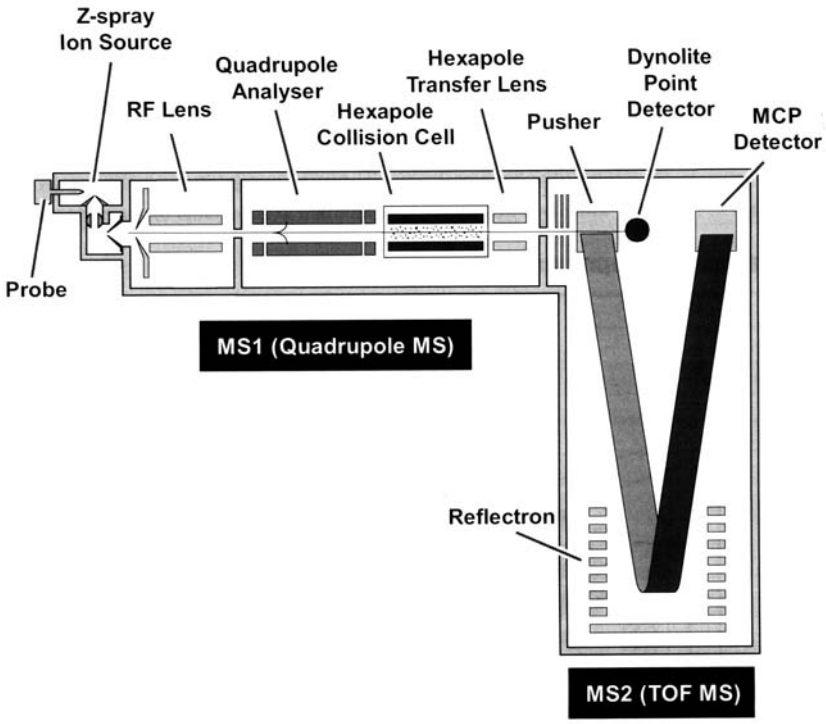


Figure 2.1. A schematic representation of the Micromass Q-Tof 2 mass spectrometer.

In the MS/MS mode, a particular ion is selected in the quadrupole analyser based on the *m/z* value of the ion, and allowed to pass into the hexapole collision cell. The ions are energised by a process known as collision induced dissociation (CID)[42] where the ions

undergo collisions with noble gases such as argon, and then fragment into daughter ions. The daughter ions enter the time of flight tube to produce MS/MS data of the selected ion.

It is possible to acquire "MS/MS/MS" data using the Q-Tof 2 instrument by increasing the energy in the ion source. This is essentially an MS/MS experiment on an energised source formed fragment ion.

2.1.2 LCQ Mass Spectrometer

The second mass spectrometer used in this study is the Finnigan LCQ Ion Trap mass spectrometer, and a schematic diagram of the instrument is shown is Figure 2.2. As with the Q-Tof 2, the ions are generated by the electrospray ionisation technique, but are then accelerated by a potential difference through an octapole and into the ion trap. The ion trap consists of three electrodes and a central ring electrode.[43] The induction of a radio frequency to the ring electrode establishes an electrical field gradient within the ion trap which is proportional to the distance from the centre of the ion trap. This allows ions with different *m/z* to each follow a stable trajectory within the ion trap for several seconds. When the amplitude of the radio frequency voltage is increased, the ions become unstable and are sequentially ejected in *m/z* sequence from the ion trap through an opening to a detector which records the mass spectrum.

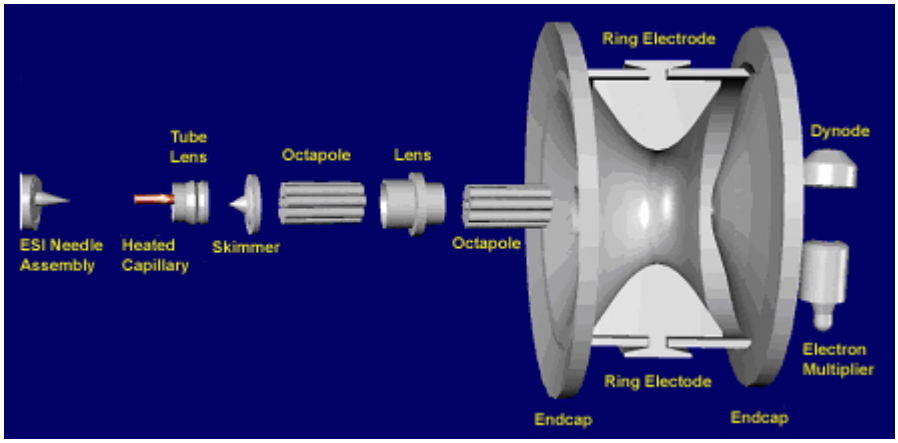


Figure 2.2. A schematic diagram of the Finnigan LCQ mass spectrometer.

MS/MS experiments can also be conducted using the ion trap mass spectrometer. This is achieved by increasing the internal energy of the system so that the ion of interest collides with helium gas present in the ion trap, and results in collision induced dissociation (CID) to form daughter fragmentation ions. The daughter ions are then sequentially ejected from the ion trap and are recorded by the detector.[44]

A benefit of ion trap mass spectrometry is the ability to conduct MS^n experiments. An MS/MS/MS experiment is where the daughter ion in a MS/MS experiment is isolated and collided with helium to produce a CID spectrum of its own daughter ions. These are grand-daughter ions of the original ion of interest. This process can be repeated routinely to MS^6 in the LCQ mass spectrometer.

2.2 Electrospray Ionisation

Ionisation techniques such as electron ionisation (EI) and chemical ionisation (CI) are not ideal for studying peptides.[45] Electrospray (ES) is the ideal method for ionising peptides as the method works well for large molecules.[46] It is a soft ionisation process which minimises the internal energy of the system, thus minimising fragmentation of the parent ion within the source.[47] Fragmentation may be achieved using the collision induced dissociation technique.

To effect electrospray ionisation (ESI), the peptide is dissolved in an aqueous organic solvent and the solution passed through a heated capillary with a potential difference of between 2-5 kV applied from the capillary needle to the electrode. Under near atmospheric conditions the solution exits the capillary as a fine mist of highly charged droplets. Depending upon the voltage applied, the droplets may emerge either positively or negatively charged. Inducing a charge onto the peptide can be achieved by applying a nebulising gas such as nitrogen which evaporates the solvent, reducing the size of the droplet to a point where the repulsive Coulombic forces at the surface of the droplet are greater than the cohesive forces holding the droplet together (Figure 2.3). This causes the droplet to break into many smaller droplets. This process of evaporation is repeated until all of the solvent is removed, liberating either a positively charged $(MH)^+$ or a negatively charged $(M-H)^-$ peptide parent ion. An electric field is used to transfer the ions to the mass analyser.

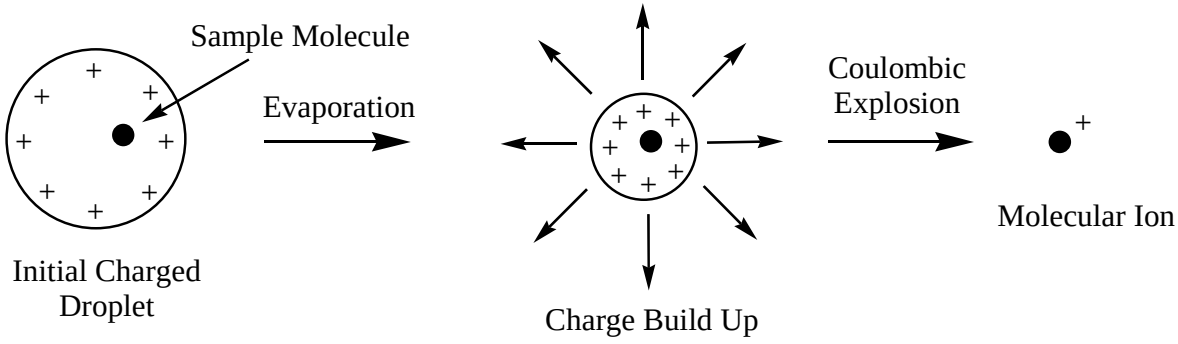


Figure 2.3. The process of electrospray ionisation (ESI) to form a $(MH)^+$ molecular ion.

Larger peptides with several protonation or deprotonation sites may be multiply charged, producing $(M+nH)^{n+}$ or $(M–nH)^{n-}$ ions where n is the number of protons added or removed respectively. This property of ESI is extremely valuable for studying large peptides and proteins because the mass scale is determined by the mass to charge ratio (*m/z*) of the species. Therefore a protein of mass 20000 Da would not be detected using a standard mass spectrometer, but if it produces an $(M+10H)^{10+}$ ion the *m/z* value will be 2001, well within the range of most mass spectrometers.

2.3 Positive Ion Mass Spectrometry

Positive ion mass spectrometry has been the primary tool used in determining the sequence of peptides. The reason for this has been the dominant cleavage of the amide bond linking adjacent amino acids together, allowing the sequence of the peptide to be determined.[48] The two fragmentations of the amide bond used for sequencing peptides are the B and Y + 2 cleavages.[31,32] The B fragmentations provide the sequence of the peptide from the C-terminal end, whilst the Y + 2 (the 2 referring to the two additional protons acquired by the C-terminal fragment) provide the sequence from the N-terminal end of the peptide.[33] An example of a positive ion spectrum is shown in Figure 2.4 for the peptide GLISSIGKALGGLIVDVIKPKS-OH **(1)**.[49] The B fragmentations are displayed above the spectrum and the Y + 2 fragmentations below the spectrum. A combination of B and Y + 2 fragmentations provide the entire sequence of the peptide with the exception of differentiation between isomeric Leu and Ile and isobaric Lys and Gln. This is achieved by utilising Edman degradation.

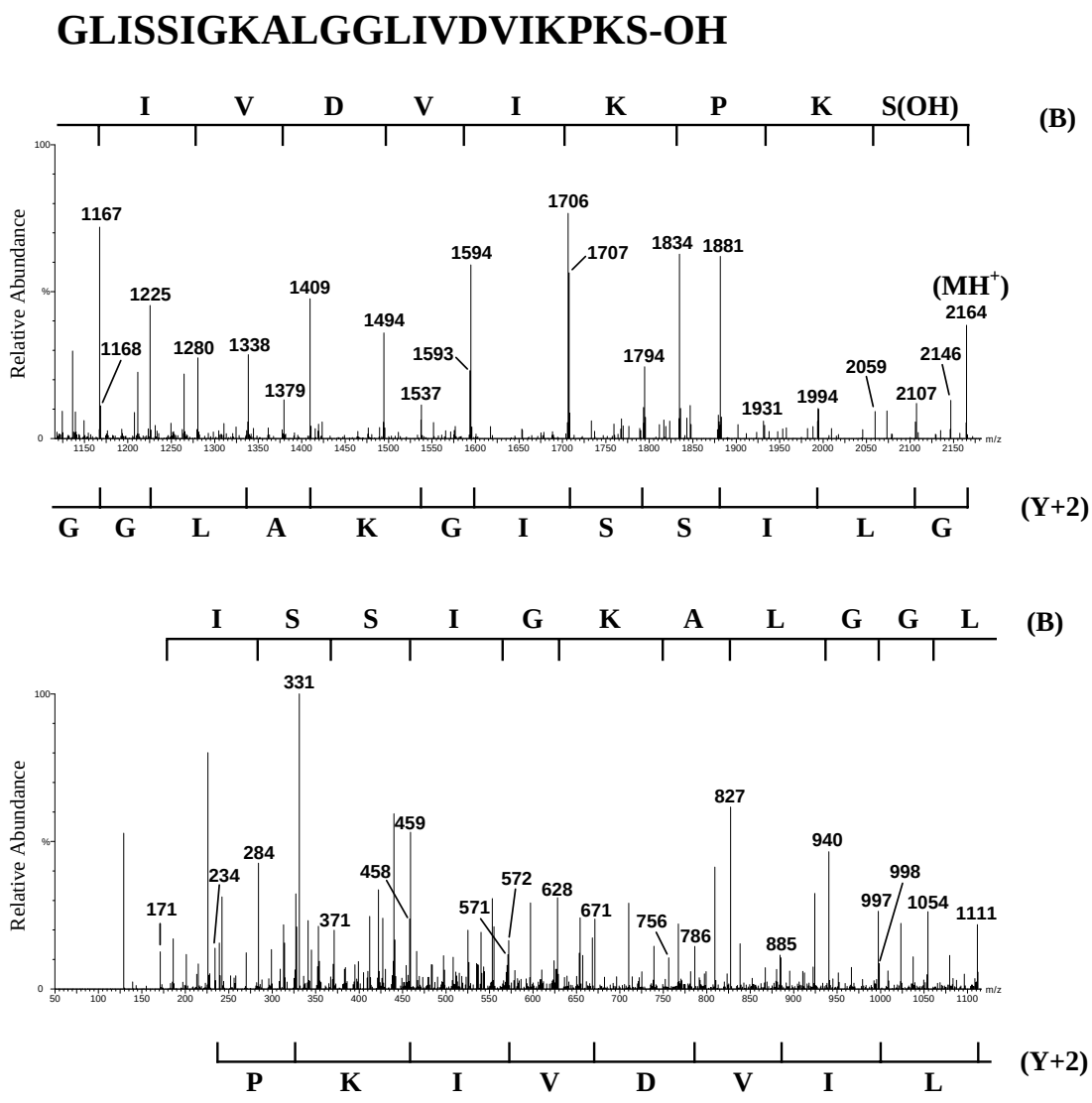


Figure 2.4. Collision-induced MS/MS data of **1**. Positive ion fragmentations of the $(MH)^+$ species, B cleavage ions are shown sequentially above the spectrum and Y + 2 cleavage ions are shown below the spectrum. Magnification range: *m/z* 1700-2150 (x4).[49]

In addition to the amide backbone cleavage there are two other less prevalent backbone cleavages involving the CH-CO and the NH-CH backbone bonds. (Figure 2.5).[31,50-53] The A and X fragmentation ions are afforded by the cleavage of the CH-CO backbone bond while cleavage of the NH-CH backbone bond forms the C + 2 and Z fragmentation ions. The D, V, W fragmentation ions are a result of one side chain and one

backbone cleavage.[51] As the A, C + 2, X, Z and D, V, W fragmentations are less prevalent than the B and Y + 2 cleavages, they are rarely used in the sequencing of peptides.

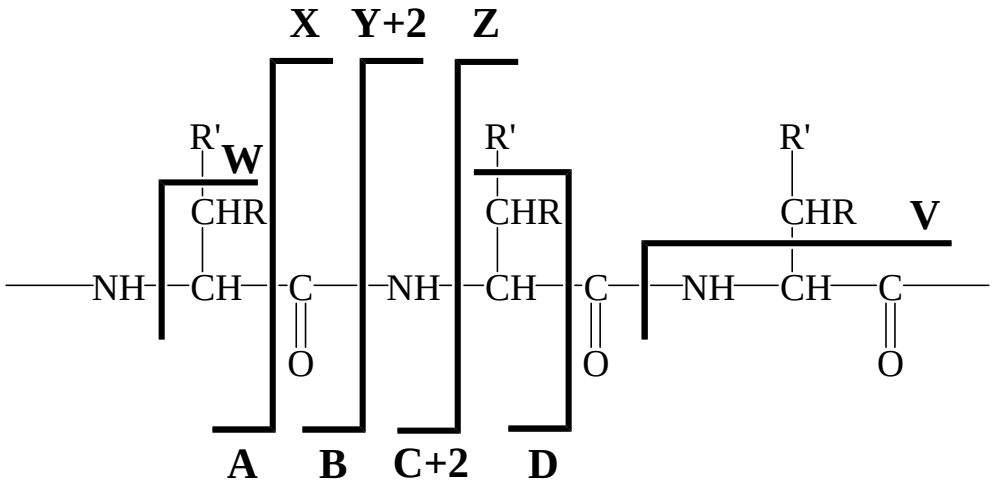


Figure 2.5. Positive ion fragmentations of peptides.[51]

Positive ion mass spectrometry does not always provide as much sequencing information as seen for **1** in Figure 2.4. For example, the positive ion MS/MS spectrum of a six residue peptide, pEPQVFV-NH_2 **(2)**[54] is shown in Figure 2.6. The B and Y + 2 fragmentations are shown schematically above and below the spectrum respectively with several A fragment ions shown on the spectrum. There are four B fragment ions which provide the sequence of the last four residues (except for differentiating between Lys and Gln). However, there are only two Y + 2 fragment ions formed and they do not distinguish the identity of the first two residues. The mass of 208 Da corresponds to three possible N-terminal sequences, pyro-Glu Pro, Ala His and His Ala. For the full sequence of a peptide to be obtained, the exploitation of Edman degradation is required. However, Edman degradation is incapable of sequencing peptides with an N-terminal pyro-Glu residue. The absence of Edman sequencing data for **2** would imply that the peptide has an N-terminal

pyro-Glu and that the second residue is therefore Pro. The differentiation between isobaric Lys and Gln is achieved by Lys-C enzymatic digest.

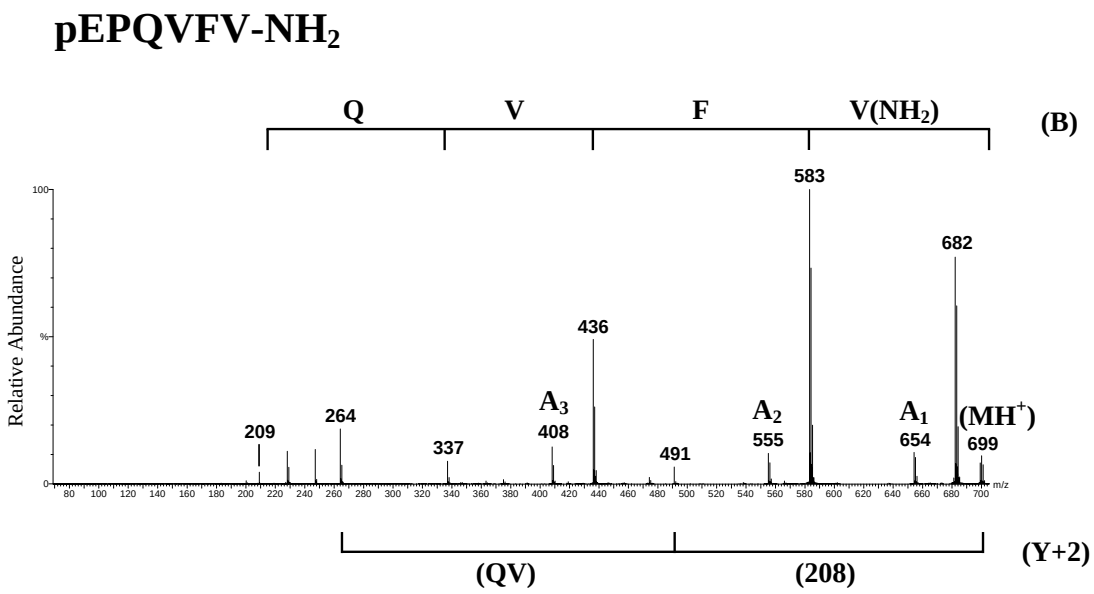


Figure 2.6. Collision-induced MS/MS data of **2**. Positive ion fragmentations of the $(MH)^+$ species, B and Y + 2 cleavage ions are shown above and below the spectrum respectively.[54]

2.4 Negative Ion Mass Spectrometry

The use of negative ion mass spectrometry in peptide sequencing has until recently been limited to the determination of the molecular weight of the peptide[55-57] and identifying post-translational modifications such as sulfate or phosphate groups. The main reason for this is the complex spectra produced in the negative ion mode. The complexity arrises from the large array of fragmentation ions produced, many of which do not correspond to backbone amide cleavages which are the most useful for peptide sequencing. Due to this, work in this area has been limited to only a small number of groups reporting

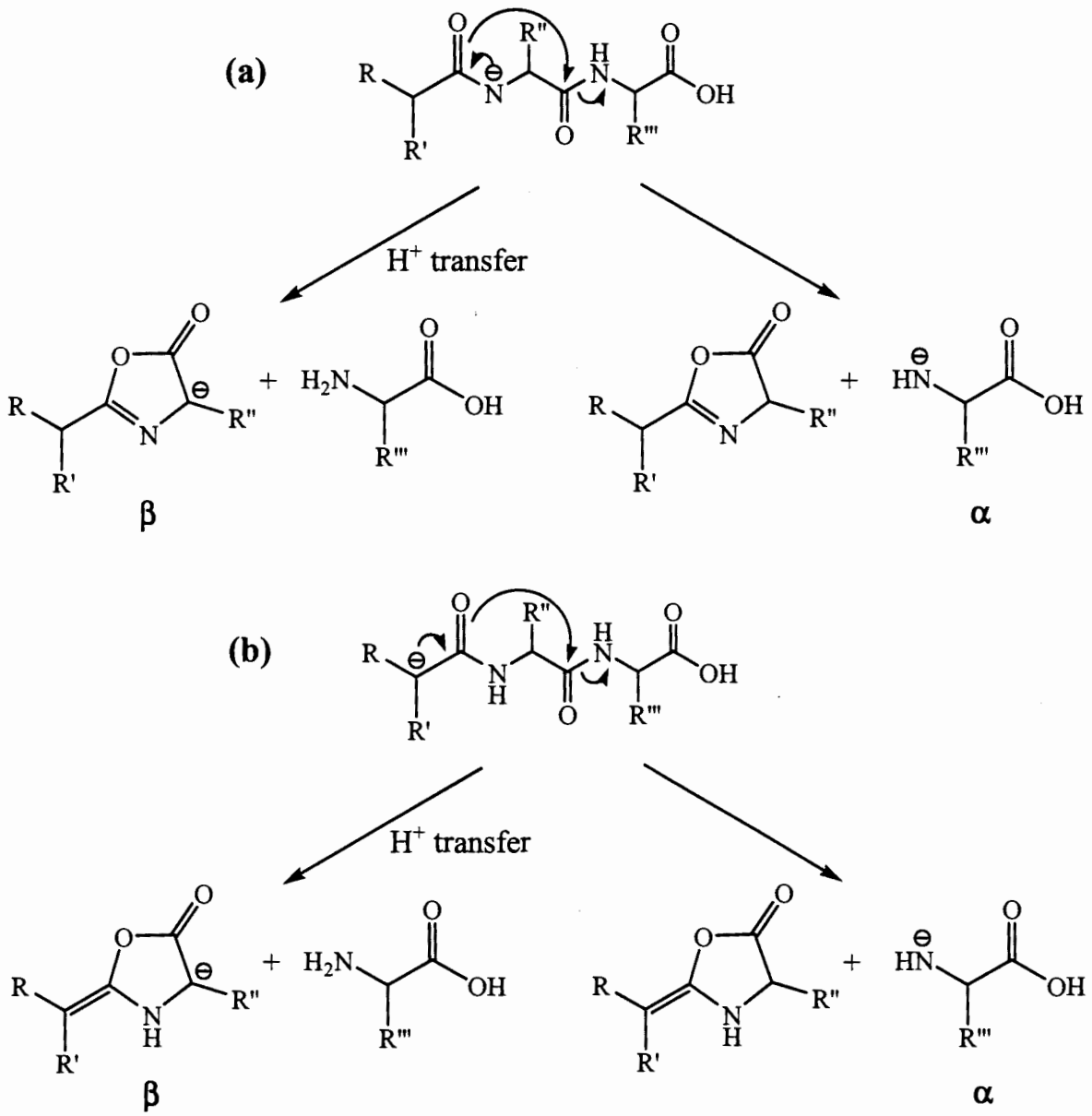
spectra of $(M-H)^-$ ions of peptides,[58,59] while several groups have been examining negative ion spectra in the hopes of explaining the formation of the magnitude of ions formed.[60-64]

The Bowie group has been involved in studying $(M-H)^-$ ions of small peptides for over a decade now. Initial work was carried out using Fast Atom Bombardment (FAB) spectroscopy on a two sector mass spectrometer (VG ZAB 2HF)[65,66] and more recently on the ESMS LCQ and Q-Tof 2.[54]

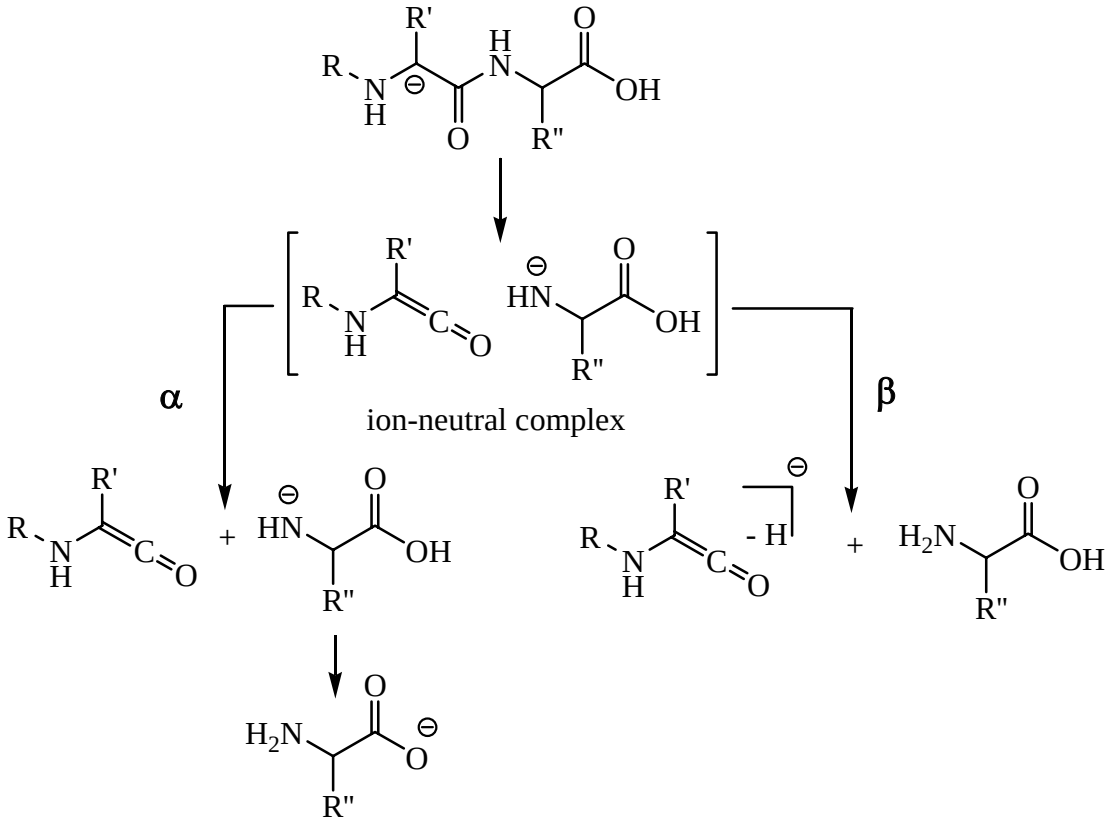
2.4.1 Amide Backbone Fragmentations

There are two backbone cleavage ions which involve fragmentation of the amide moieties and provide sequencing information equivalent to that provided by the B and Y + 2 cleavages in the corresponding positive ion spectra.[66,67] Although the information provided is analogous to that of the B and Y + 2 cleavages, the mechanisms by which the ions are formed in the positive and negative ion modes are completely different. These cleavages have therefore been named α and β backbone cleavages to distinguish them from the positive ion counterparts. The α cleavage ions provide the amino acid sequence from the N-terminal end of the peptide, while the β cleavage ions give the sequence from the C-terminal end of the peptide. The mechanism for α and β ion formation is shown in Scheme 2.1.

In addition to the mechanism for α and β fragment ion formation shown in Scheme 2.1, there are two other proposed mechanisms which result in a cyclic β fragment ion.[173,175,176,178] The first mechanism, shown in Scheme 2.1A (a), is initiated from an amide anion while the second mechanism, Scheme 2.1A (b), involves cyclisation from an enolate anion.



Scheme 2.1A. Formation of α and cyclic β fragment ions under negative ion mass spectrometry conditions from an initial (a) amide and (b) enolate anion.[173,175,176,178]



Scheme 2.1. The mechanism for the amide backbone cleavage of peptides in negative ion mass spectrometry to form α and β fragmentation ions.

When the peptide pEPQVFV-NH_2 **(2)** is re-examined using negative ion mass spectrometry, the sequence of the peptide with the exception of the last two residues is provided. The spectrum is shown in Figure 2.7 with the α cleavages displayed above the spectrum and the β cleavages below. The formation of the γ ion at *m/z* 472 confirms residue three as Gln. (Formation of γ and δ fragmentation ions will be discussed in full detail in sections 2.4.4 to 2.4.6). By combining the information provided by the positive and negative ion spectra, the entire sequence of the peptide can be determined.

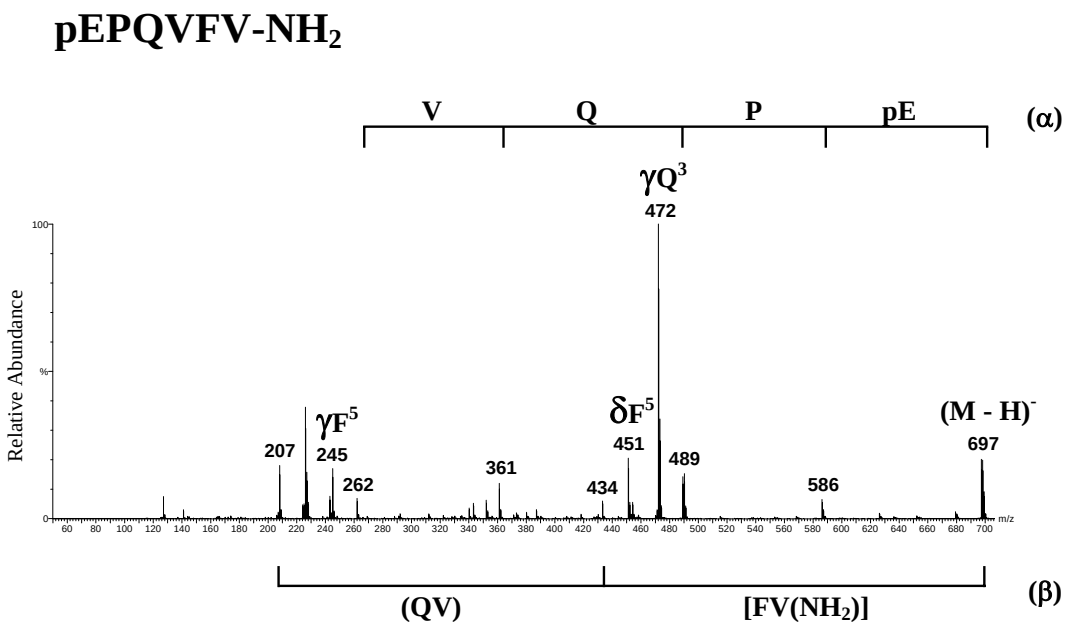


Figure 2.7. Collision-induced MS/MS data of **2**. Negative ion fragmentations of the $(M-H)^-$ ion, α and β cleavages are shown above and below the spectrum respectively. Magnification range: *m/z* 50-350 (x24).[54]

Close examination of many $(M-H)^-$ spectra including the one shown in Figure 2.7 reveal that the α cleavage ions are more abundant relative to the β cleavage ions. To explain this observation, theoretical calculations were undertaken to investigate the energetics of both the α and β cleavage processes. The reaction coordinate profile for the α and β cleavage processes are shown in Figure 2.8[68] as the energetics of other negative ion fragmentations will be compared to this.

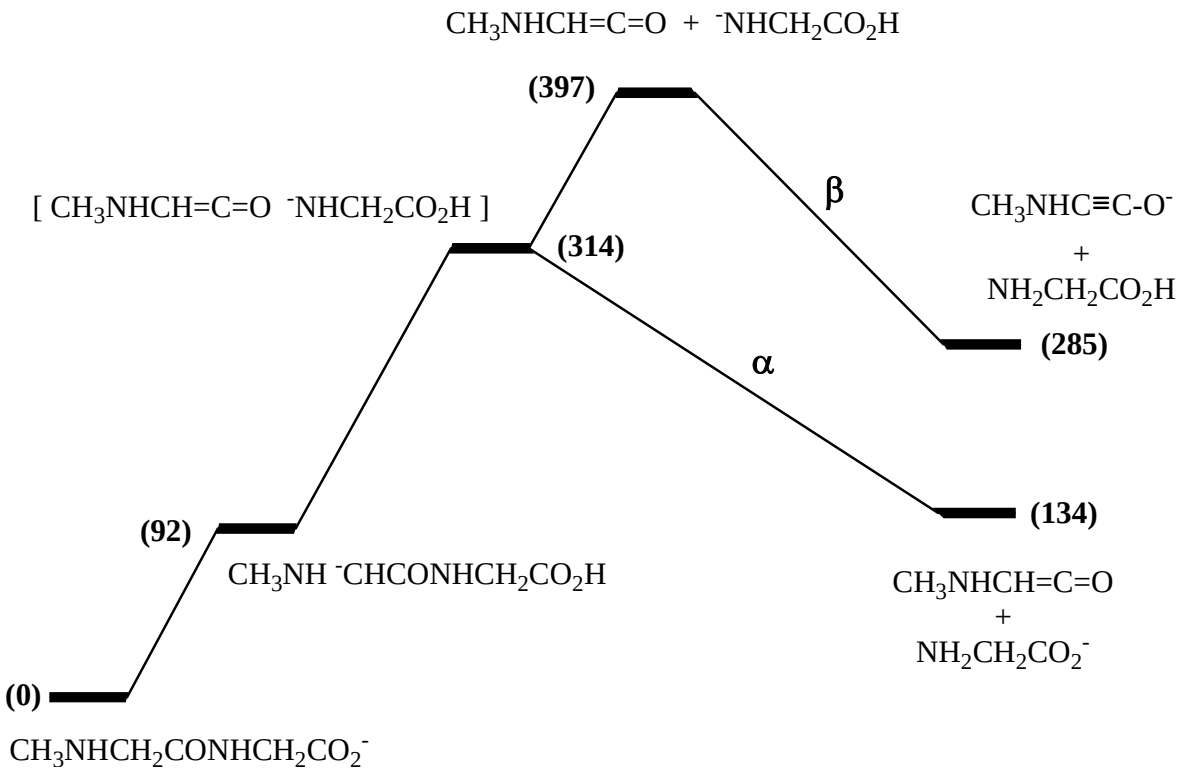


Figure 2.8. The α and β backbone cleavage fragmentations using the model system $CH_3NHCH_2CONHCH_2CO_2^-$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory.

The α cleavage process is endothermic by 134 kJ mol^{-1} and is thermodynamically more favourable than the β cleavage process by 151 kJ mol^{-1}. These results support the spectral observations that the α cleavage fragmentation ions are more numerous and abundant than the β cleavage fragmentations.

2.4.2 The $\beta^{|}$ Backbone Cleavage

In addition to the α and β amide backbone cleavages there is one more backbone amide cleavage. It provides sequencing information from the C-terminal end of the peptide but by a different mechanism to that of the β ion, and has therefore been termed the $\beta^{|}$ cleavage.[68] When considering the β cleavages observed in Figure 2.7, the first fragmentation ion from the C-terminal end is Phe Val (NH_2) and the second is Gln Val Phe Val (NH_2). However, in the negative ion mass spectrum of GLFAVIKKVAAVIKKL-NH_2 **(3)**[54] shown in Figure 2.9, there are no fragment ions corresponding to the loss of Leu (NH_2), Lys Leu (NH_2), Lys Lys Leu (NH_2) etc. There are however fragmentation ions which correspond to the C-terminal loss of Leu followed by Lys Leu, then Lys Lys Leu and so on. For this type of fragmentation to occur the NH from the C-terminal amide anion must be transferred to a portion of the peptide which is N-terminal in relation to the position of the amide backbone cleavage. ($\beta^{|}$ cleavages of peptides can occur from both C-terminal amides and carboxylic acids).[69]

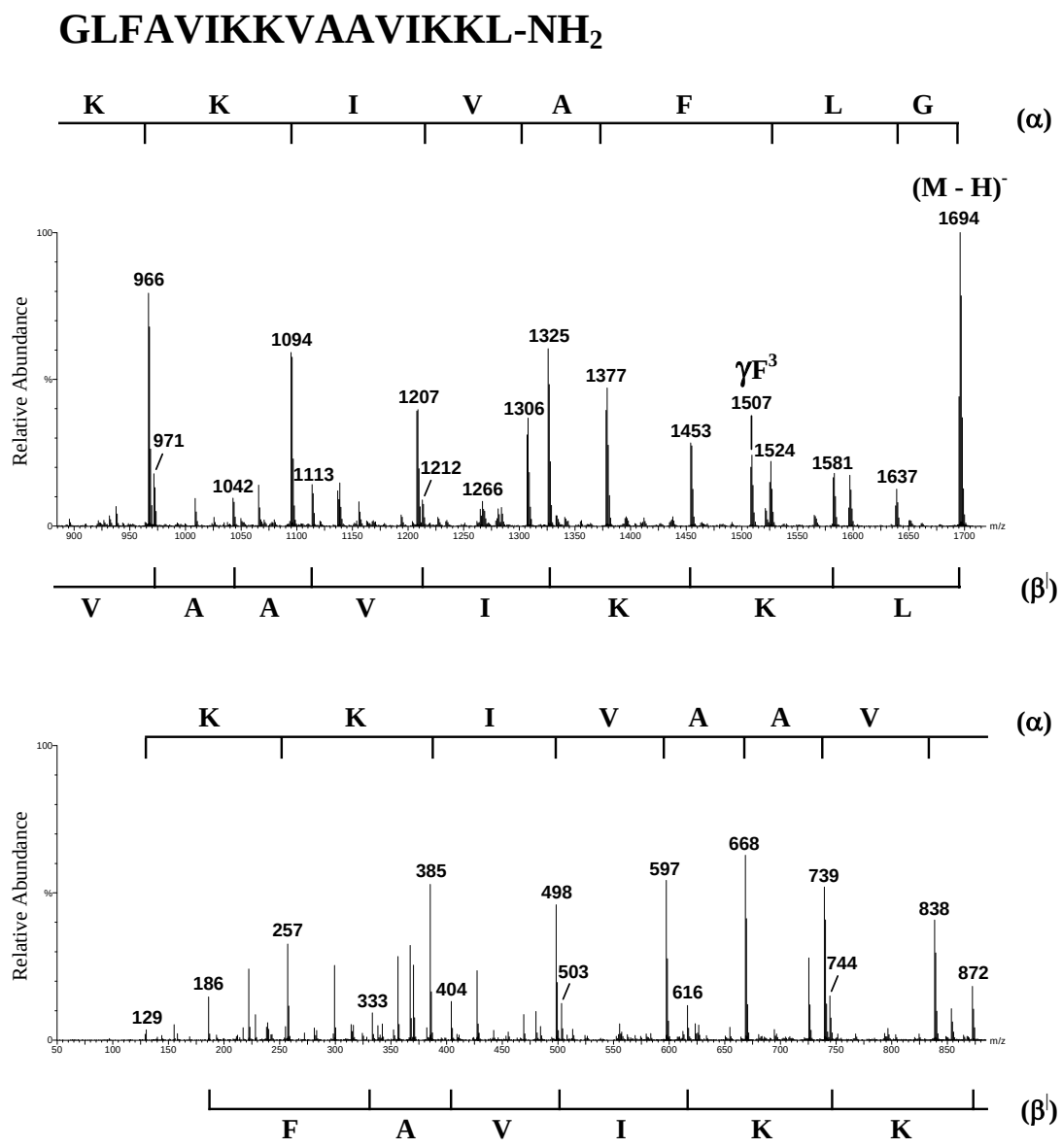
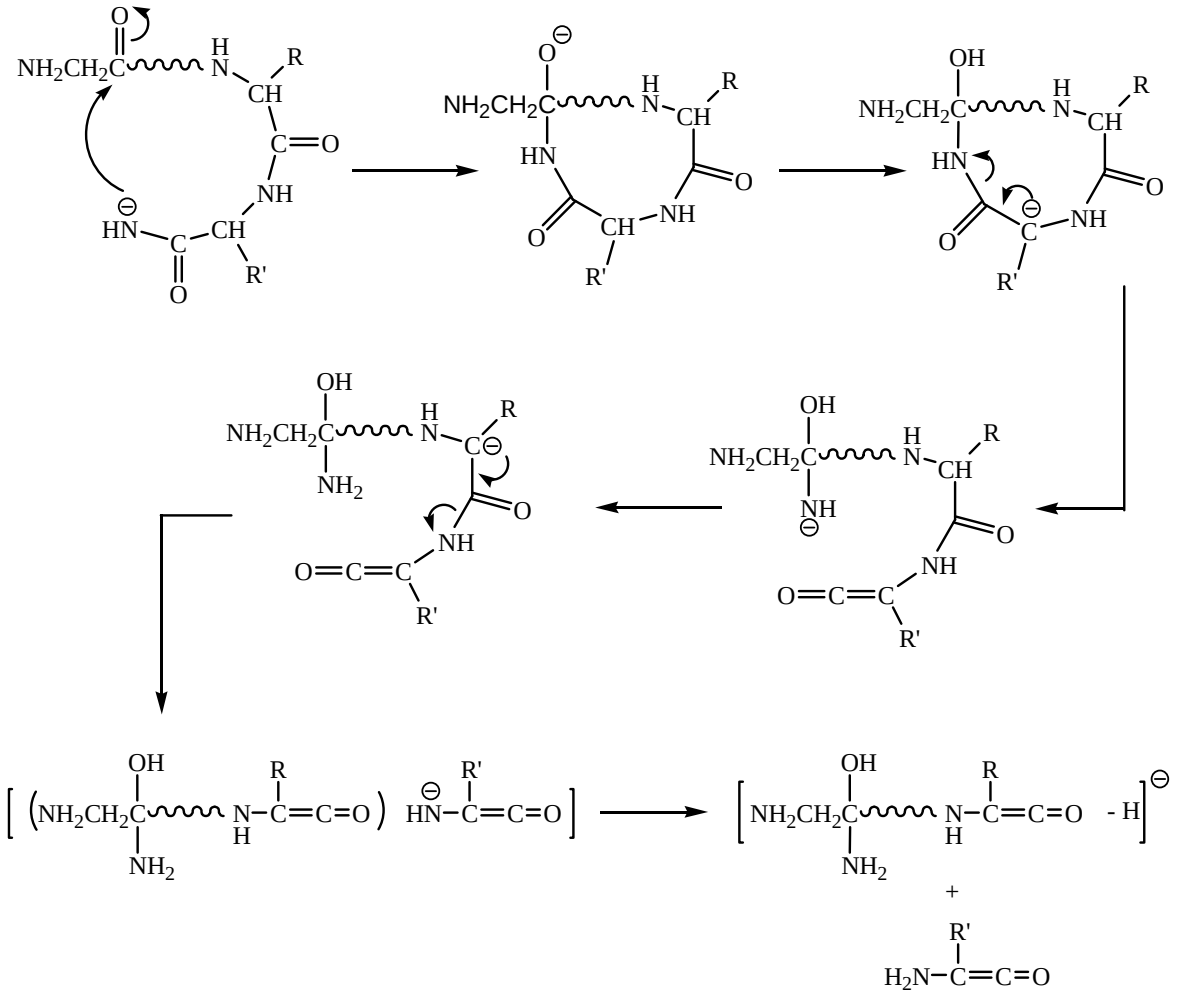


Figure 2.9. Collision-induced MS/MS data of **3**. Negative ion fragmentations of the $(M\text{-}H)^-$ ion, α cleavages are shown above the spectrum with $\beta^|$ cleavages displayed below the spectrum. Magnification ranges: *m/z* 50-450 (x2), *m/z* 950-1650 (x2).[54]

This type of fragmentation is achieved by a combination of two methods. The first method involves the C-terminal amide (or carboxylic acid) cyclising to each residue at the carbonyl group. This is then followed by an enolate induced amide cleavage at the residue adjacent on the C-terminal side of the cyclised intermediate as shown in Scheme 2.2. The second method is mechanistically very similar. The only difference is that the C-terminal

amide migrates to the N-terminal end of the peptide and is then followed by enolate initiated amide backbone cleavages at each of the amide bonds.[68]



Scheme 2.2. The mechanism of the $\beta^{|}$ cleavage process.

Molecular modelling studies were conducted to confirm the mechanism shown in Scheme 2.2, and to determine if the $\beta^{|}$ backbone cleavage is energetically competitive with the α and β backbone cleavages.[68] The results of the modelling study reveal that the cyclisation of the C-terminal amide and formation of the enolate ion is exothermic by as much as 125 kJ mol^{-1}. From the enolate anion, the mechanism of backbone amide cleavage is similar to the β fragmentation process shown in Figure 2.8. This means that

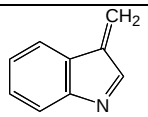
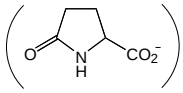
the overall $\beta^{|}$ cleavage mechanism is endothermic by only 68 kJ mol^{-1} and is energetically favourable in comparison with both α and β backbone cleavages.

The results suggest that the $\beta^{|}$ process is more facile than both the α and β cleavages yet spectra with $\beta^{|}$ cleavages often also contain α cleavages. This could be a result of the theoretical calculations failing to take into account the entropy of the cyclisation of the C-terminal amide.

2.4.3 Characteristic Side Chain Fragmentations

Many amino acids undergo characteristic side chain cleavages in the negative ion mode, a feature not generally observed in the positive ion mode.[70] These fragmentations often confirm the presence of a particular amino acid in the peptide but do not define its relative position within the peptide. A complete list of side chain cleavages is outlined in Table 2.1.

Table 2.1. Characteristic side chain fragmentations.

Amino acid	Loss (or formation)	Mass
Ala	$Me^{\bullet}$	15
Val	$Pr^{\bullet}$	43
Leu (Ile)	$Bu^{\bullet}$	57
Phe	$(PhCH_2^-)$	91
Tyr	$(p\text{-}HOC_6H_4CH_2^-)$ $O{=}C_6H_4{=}CH_2$	107 106
Trp		129
Ser	CH_2O	30
Thr	MeCHO	44
Cys	H_2S	34
Met	MeSH	48
Asp	H_2O	18
Asn	NH_3	17
Glu	H_2O 	18 128
Gln	NH_3	17
Arg	HN=C=NH	42

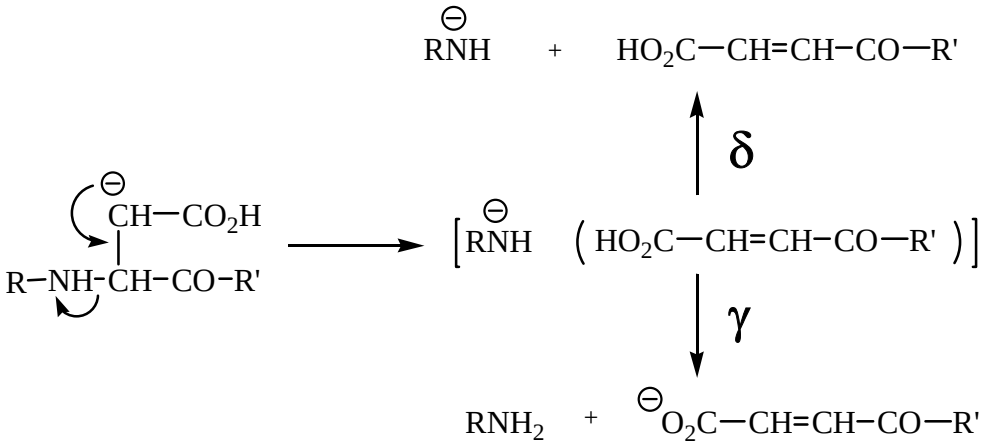
The first type of side chain cleavage involves the complete loss of the side chain as an alkyl radical and is most pronounced for residues such as Ala, Val and Leu/Ile losing $Me^{\bullet}$, $Pro^{\bullet}$ and $Bu^{\bullet}$ respectively.[71-73] However, many of these side chain fragmentations only occur for small peptides up to four residues in length, and become less pronounced as the size of the peptide increases.

There are some amino acids however which provide pronounced side chain cleavages irrespective of the size of the peptide. These include Ser ($-CH_2O$), Thr ($-CH_3CHO$), Asp and Glu ($-H_2O$) and Asn and Gln ($-NH_3$). In fact, as the size of the peptide increases in length these side chain cleavages often begin to suppress the α, β and $\beta^{|}$ backbone cleavage ions. In other cases the α, β and $\beta^{|}$ cleavage ions are observed not from the $(M-H)^-$ parent ion but in the case of Ser for example, from the $[(M-H)^- -CH_2O]^-$ fragment ion.

2.4.4 Side Chain Induced Backbone Cleavages

In addition to the α, β and $\beta^{|}$ fragmentation ions formed by the cleavage of the backbone amide bond, there is a second type of backbone cleavage which is initiated by the side chain of certain amino acids. This occurs for amino acids that can be deprotonated on the side chain carbon adjacent to the backbone to form a resonance stabilised anion and was first observed for Asp, Asn, Phe, Tyr, His and Trp.[54] As the size of the peptide increases in length the presence of the Phe, Tyr and His induced backbone cleavages often decrease, whereas for Asp and Asn, these fragmentation ions remain intense.

In comparison to the α, β and $\beta^{|}$ fragmentations which involve the cleavage of the amide backbone bond, side chain induced backbone cleavages involve cleavage of the NH-CH bond to form the ion-neutral complex shown in Scheme 2.3.[74] Dissociation of the ion-neutral complex results in the formation of RNH^- known as the δ ion, whilst proton transfer from the C-terminal portion of the peptide followed by dissociation results in the loss of RNH_2 forming a γ cleavage ion.



Scheme 2.3. The mechanism for the formation of the γ and δ fragmentation ions from Asp.

The spectrum shown in Figure 2.10 is of an Asp and Asn containing peptide, GFADIFGKVANLIKS-NH_2 **(4)**.[75] The α and β fragment ions are shown above and below the spectrum respectively with the side chain induced backbone cleavages identified on the spectrum. The peak at *m/z* 1284 is the Asp γ ion which then undergoes the additional loss of water to form the base peak in the spectrum. This is followed by the side chain loss of CH_2O from the Ser residue. The γ Asn ion is at *m/z* 554 with the additional loss of ammonia forming the ion at *m/z* 537. The δ fragmentation ions from the Asp and Asn residues can also be seen in the spectrum at *m/z* 291 and 1021 respectively. A reduction in the number of α and β ions indicates that these cleavages are being suppressed by the side chain induced backbone cleavages. This suggests that these fragmentations are energetically competitive with the α, β and $\beta^{|}$ processes. (Nomenclature: formation of a γ ion from an Asp residue at position 4 will be abbreviated to γAsp4 within the text and labelled γD^4 in figures).

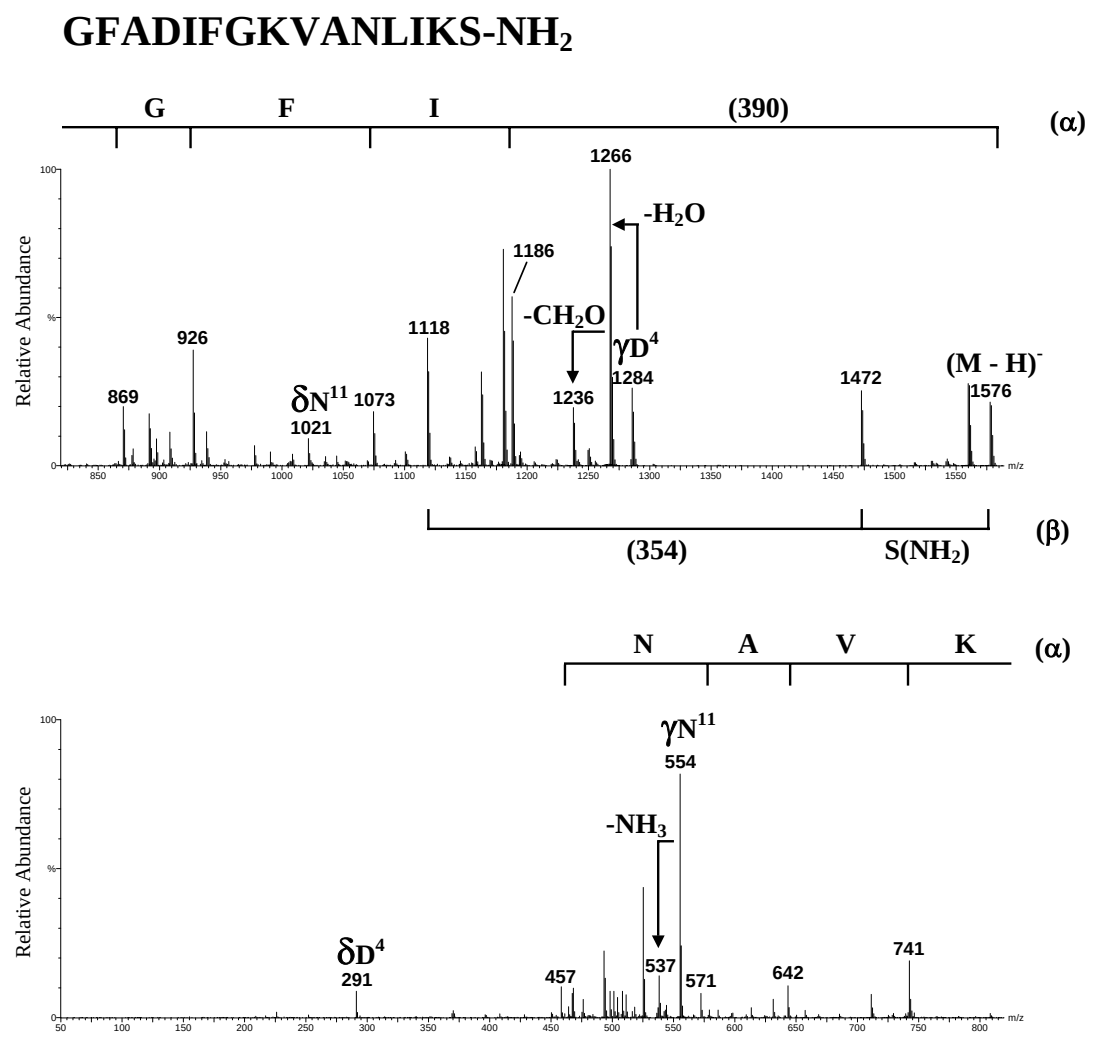
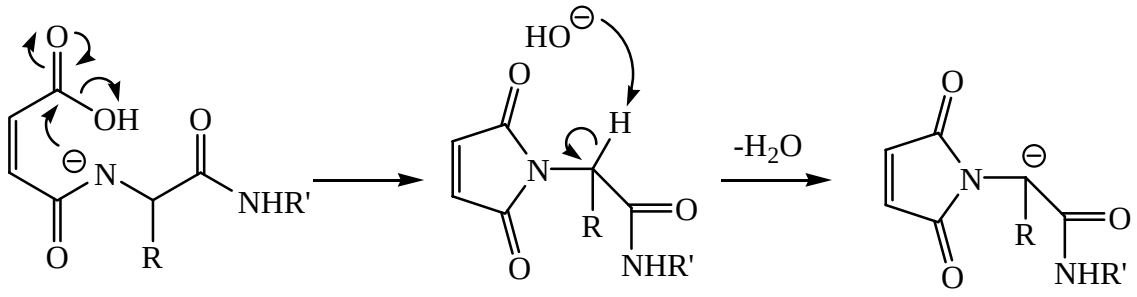


Figure 2.10. Collision-induced MS/MS data of **4**. Negative ion fragmentations of the $(M-H)^-$ ion, α and β cleavages are shown above and below the spectrum respectively with γ and δ fragmentation ions displayed on the spectrum. Magnification ranges: *m/z* 450-750 (x6), *m/z* 850-1200 (x12).[75]

Theoretical calculations have been conducted on the mechanism by which Asp initiates backbone cleavages.[76] The formation of the γ ion is exothermic by 41 kJ mol^{-1} from the enolate anion with the δ ion formation endothermic by 310 kJ mol^{-1}. The results confirm that the γ cleavage is a more favourable process than the δ cleavage, which is

consistent with spectral data where the γ ions are more abundant and intense than the δ ions.

The mechanism for the loss of water from the Asp γ ion to form the base peak in the spectrum is shown in Scheme 2.4. In the case of Asn, the γ ion eliminates NH_3. This allows for easy differentiation of Asp and Asn residues.



Scheme 2.4. The loss of water from the Asp γ fragmentation ion.

2.4.5 γ and δ Cleavages of Glu and Gln

For the amino acids Glu and Gln, the side chain enolate anion is one carbon too many to initiate γ and δ backbone cleavages by the mechanism shown previously, and it was assumed that these amino acids would not initiate backbone cleavages.[54] However the negative ion spectrum of a Glu containing peptide NEEEKVKWEFPDVP-NH_2 **(5)**[77] shown in Figure 2.11 reveals fragmentation ions corresponding to the γ cleavage initiated by each of the four Glu residues.

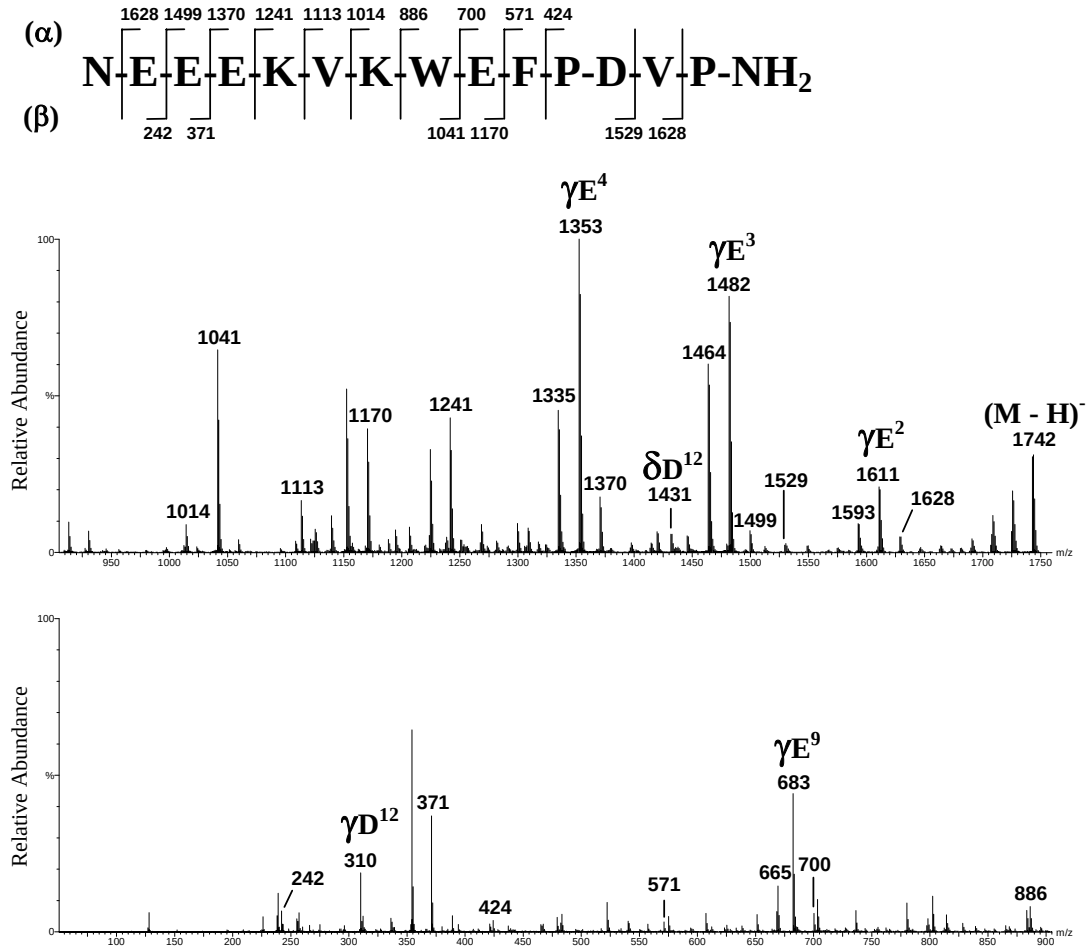
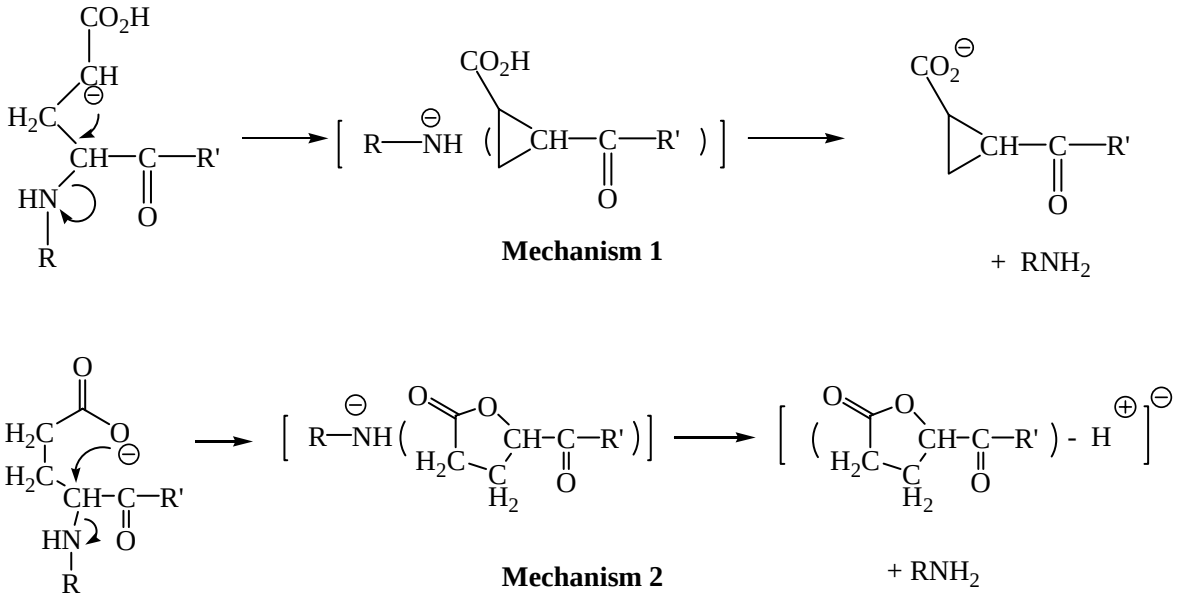


Figure 2.11. Collision-induced MS/MS data of **5**. Negative ion fragmentations of the $(M-H)^-$ ion, α and β cleavages are shown on the formula above the spectrum with γ fragmentation ions displayed on the spectrum. Magnification ranges: *m/z* 200-1025 (x3), *m/z* 1100-1340 (x4), *m/z* 1360-1750 (x2).[54]

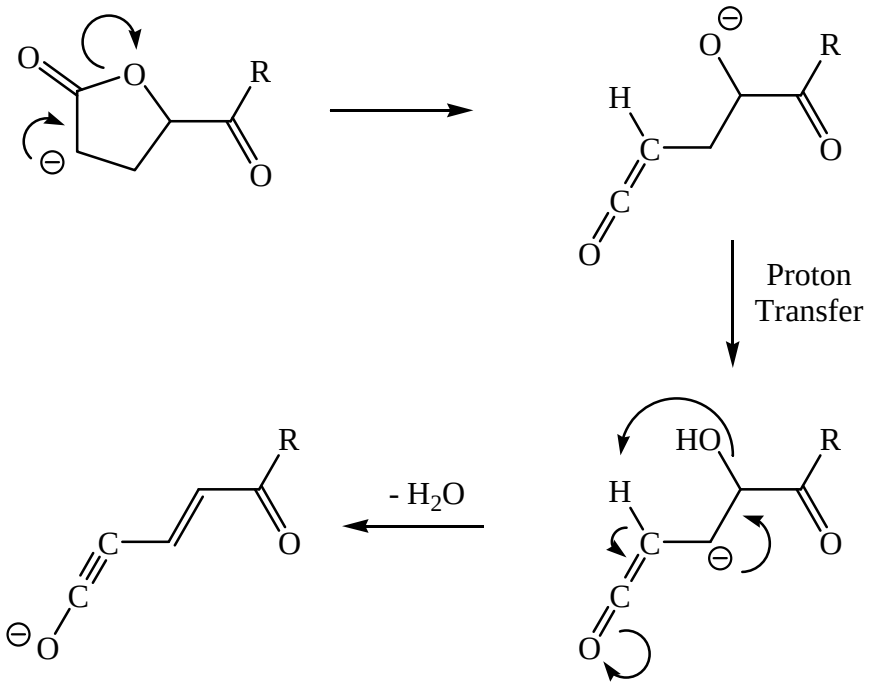
There are several possible mechanisms for the formation of a Glu γ ion. The side chain enolate anion may cyclise to form a three membered ring intermediate followed by the elimination of RNH^- (the δ ion), with proton transfer effecting formation of the γ ion (Mechanism 1, Scheme 2.5). Alternatively, nucleophilic attack by the side chain carboxylate anion could form a five membered ring intermediate which may decompose to form the γ and δ ions (Mechanism 2, Scheme 2.5).



Scheme 2.5. Two possible mechanisms for the γ cleavage process by Glu. Mechanism 1: The three membered ring intermediate initiated from the enolate anion. Mechanism 2: The five membered ring intermediate initiated by the carboxylate anion.

Theoretical calculations were conducted to confirm which of the two processes was more likely to form the γ ions observed in the spectrum.[76] The formation of the γ ion by the three membered ring intermediate is endothermic by 337 kJ mol^{-1}. The five membered ring process however is endothermic by only 100 kJ mol^{-1}.

As was the case for Asp, a loss of water follows the formation of the Glu γ ion. To rationalise the presence of the γ-H_2O peak it was proposed that ring opening of the γ ion takes place as shown in Scheme 2.6.[76]



Scheme 2.6. The loss of water from the Glu γ fragmentation ion.

In the case of Gln, the γ ion loses NH_3. This observation is extremely valuable for peptide sequencing as the amino acids Lys and Gln are isobaric and it is not possible to distinguish between the two using positive ion mass spectrometry (unless the instrument has high resolution capability). However, in the negative ion mode the δ, γ and γ-NH_3 fragmentation ions from Gln clearly distinguish this residue from Lys which cannot initiate backbone cleavage. This means that automated Edman degradation or Lys-C enzymic digestion followed by mass spectrometric identification is no longer the only way of differentiating between Lys and Gln residues.[78]

2.4.6 Ser and Thr Side Chain and Backbone Cleavages

As mentioned previously, the residues Ser and Thr undergo specific side chain fragmentations resulting in the loss of CH_2O and CH_3CHO respectively. As a result, the Ser and Thr residues are converted to Gly. The loss of CH_2O from Ser is so facile that the $[(M-H)^- -CH_2O]^-$ fragmentation ion is often the base peak in the spectrum. Additionally, α, β and $\beta^|$ cleavages often originate from the $[(M-H)^- -CH_2O]^-$ ion rather than the parent $(M-H)^-$ ion.[54] The spectrum of a Ser containing peptide GLFAVIKKVASVIKGL-NH_2 **(6)**[54] is shown in Figure 2.12. The $[(M-H)^- -CH_2O]^-$ fragmentation ion is the base peak in the spectrum with all α and $\beta^|$ cleavages originating from this ion. The formation of the γF^3 ion at *m/z* 1422 also originates from the $[(M-H)^- -CH_2O]^-$ ion rather than the parent ion. The loss of CH_3CHO from Thr residues is not as facile as the loss of CH_2O from Ser. The $[(M-H)^- -CH_3CHO]^-$ ion is rarely the base peak of the spectrum, and the α, β and $\beta^|$ cleavages generally originate from the $(M-H)^-$ ion.[54]

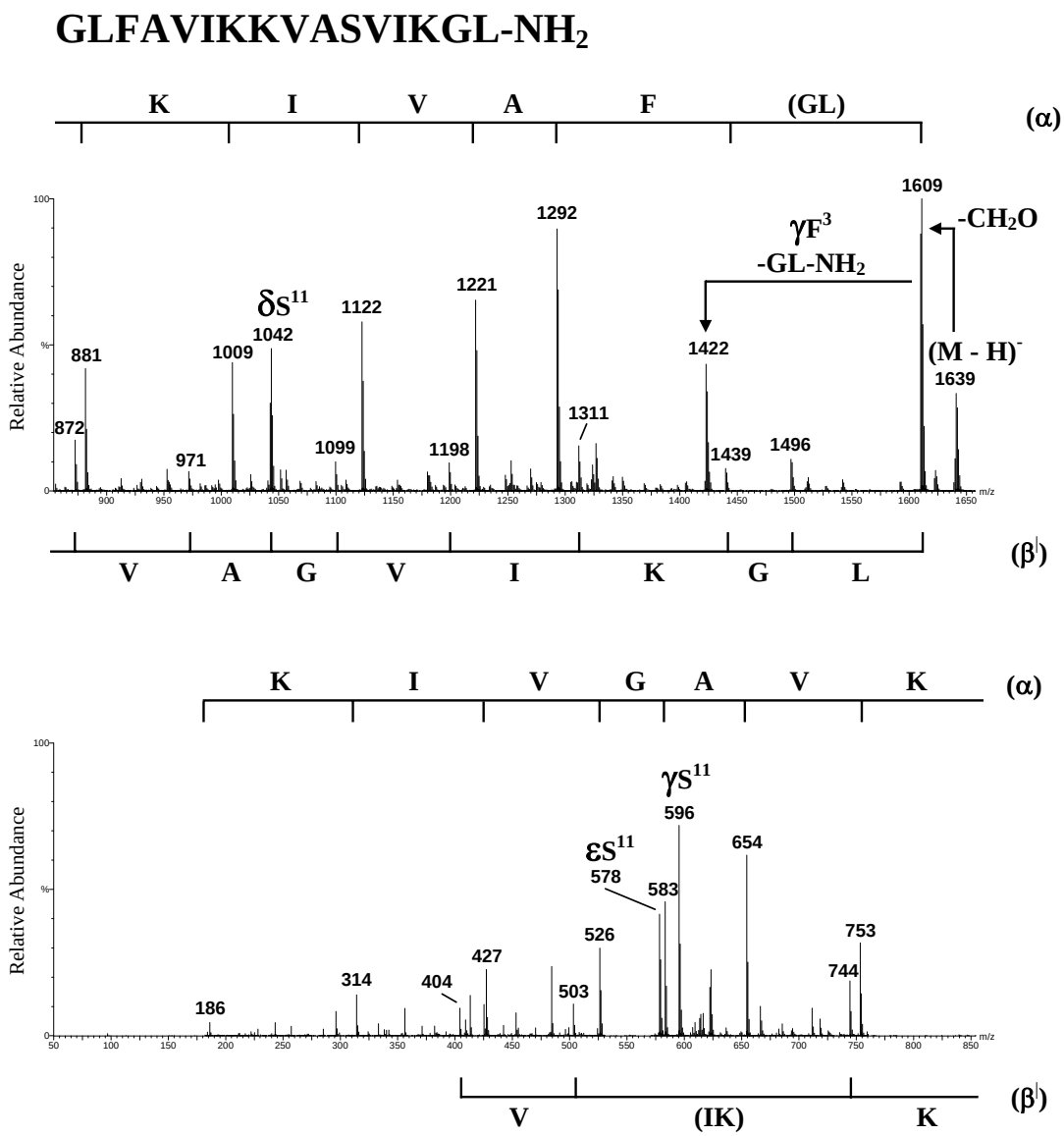
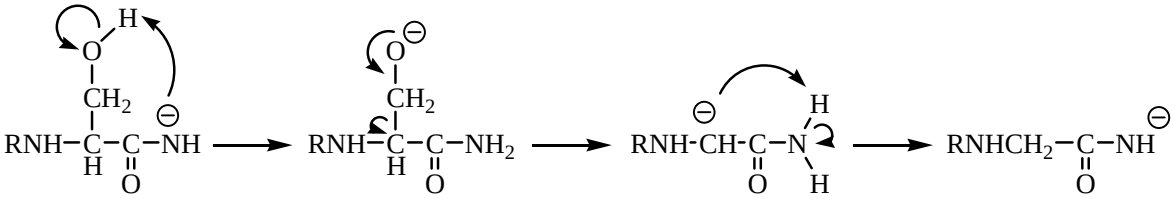


Figure 2.12. Collision-induced MS/MS data of **6**. Negative ion fragmentations of the $(M-H)^-$ ion. α and $\beta^|$ cleavages are shown above and below the spectrum respectively. Magnification ranges: *m/z* 150-550 (x54), *m/z* 550-1430 (x24).[54]

The mechanism for the loss of CH_2O from Ser containing peptides requires the transfer of the proton from the side chain hydroxyl group to form an alkoxide anion. The loss of CH_2O results in the formation of an enolate anion which can deprotonate an amide to form the resonance stabilised amide anion (Scheme 2.7).[78]



Scheme 2.7. The side chain loss of CH_2O from the Ser residue.

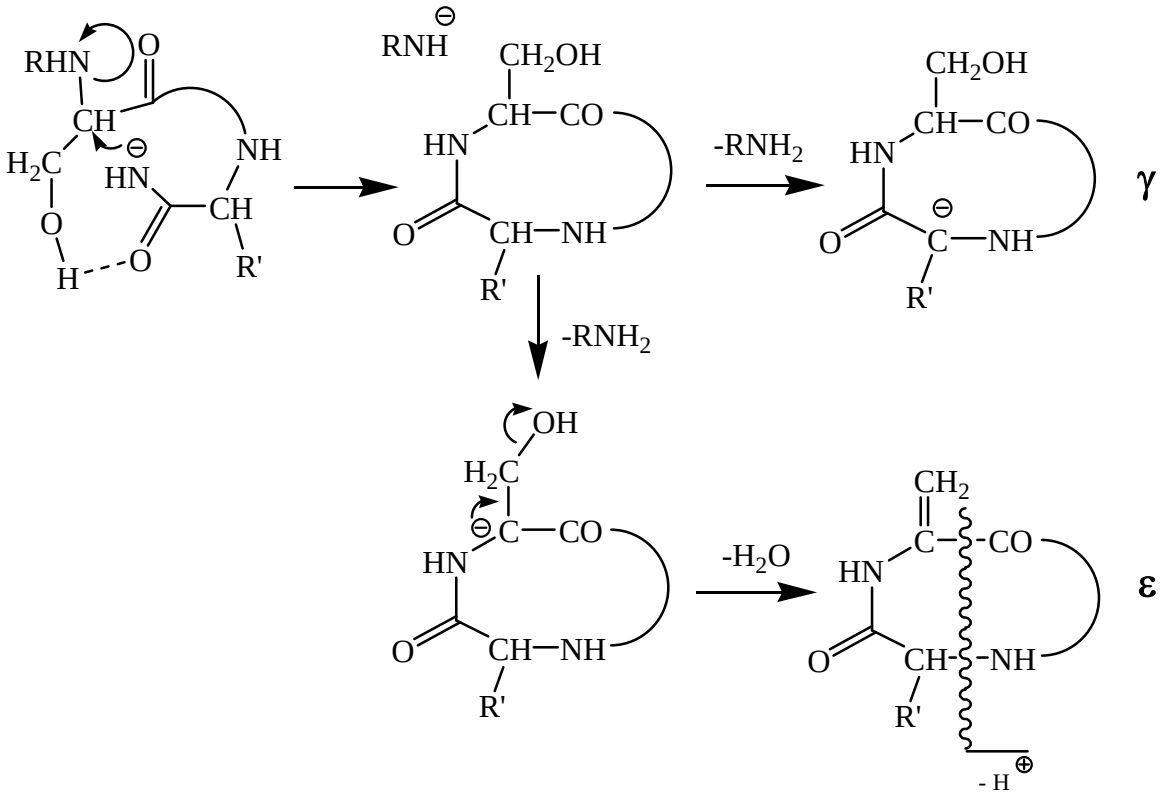
Theoretical calculations were conducted to understand why the $[(M-H)^- - CH_2O]^-$ ion is the base peak of the spectra of Ser containing peptides and why α, β and β$^|$ cleavages originate from the $[(M-H)^- - CH_2O]^-$ ion rather than the $(M-H)^-$ ion.[78] The loss of CH_2O from Ser converts the Ser residue into a Gly residue and requires only 149 kJ mol^{-1} with the final product endothermic by only 57 kJ mol^{-1}.[78]

When the energetics of the CH_2O loss from Ser are compared with those of α, β and β$^|$ cleavages, it is clearly evident that the Ser side chain loss is both kinetically and thermodynamically more favourable. (The α cleavage is endothermic by 134 kJ mol^{-1} with a transition state barrier of 314 kJ mol^{-1}.[68] The β cleavage process is endothermic by 285 kJ mol^{-1} with a barrier of 379 kJ mol^{-1}, and the β$^|$ backbone cleavage is endothermic by 68 kJ mol^{-1} with a barrier of 180 kJ mol^{-1}).[68] This explains why the $[(M-H)^- - CH_2O]^-$ ion is the base peak of the spectrum for Ser containing peptides, and why α, β and β$^|$ cleavages originate from the $[(M-H)^- - CH_2O]^-$ ion rather than the $(M-H)^-$ ion.

An examination of the negative ion spectrum of **6** shown in Figure 2.12 reveals a peak at *m/z* 596 which corresponds to a γ fragmentation ion resulting from the presence of the Ser residue in the peptide. The abundance of *m/z* 596 in the spectrum indicates that it is formed by a facile process which must compete favourably with loss of CH_2O. There is

also a large peak at *m/z* 578, 18 Da lower than the γ ion at *m/z* 596. The difference of 18 Da between the two peaks indicates that the γ ion may also lose water to generate the lower mass ion. MS/MS experiments of the source formed ion at *m/z* 596 however did not fragment to give *m/z* 578, meaning that the two peaks originate from either a common intermediate or two independent processes.

The mechanism for the formation of the γ fragmentation ion and the γ-H_2O ion is shown in Scheme 2.8.[78] It involves the cyclisation of the C-terminal carboxylic acid or amide anion to form a hydrogen bond between the carbonyl oxygen of the terminal group with the hydrogen of the Ser side chain. Cyclisation would then result in an ion-neutral complex where RNH^- could deprotonate the cyclic intermediate to form the γ ion. Alternatively, if the cyclic intermediate is deprotonated at the α-carbon of the Ser residue, an additional loss of water will result in the formation of the lower mass peak in the spectrum. This process has been called the ε backbone cleavage.



Scheme 2.8. The mechanism for the Ser induced formation of the γ and ε ions.

2.5 Theoretical Calculations[79]

The use of theoretical chemistry has expanded rapidly over the past decade, largely due to advances in the computational speed of modern super-computers. It is often used in concert with experimental data, to give insight into mechanisms and energetics of reaction pathways.

There are two main areas of computational chemistry, molecular mechanics and electronic structure theory. Both perform similar types of calculations such as determination of energies or vibrations of molecules, however, the approach used is significantly different. Molecular mechanics use the laws of classical physics to predict

the structures and properties of molecules. The calculations are based upon the interaction among nuclei, without explicitly considering the electrons in the system. This approach is limited and cannot be accurately used for chemical problems where electronic effects dominate, such as those involving the breakage or formation of bonds.

Electronic structure methods use the laws of quantum mechanics as the basis for computations. Quantum theory states that the energy and other related properties of a molecule can be determined by solving the Schrödinger equation ($H\psi = E\psi$). E is the energy of the system, H is the Hamiltonian operator and ψ, the wave function. Apart from the smallest systems, solving the exact solution to the Schrödinger equation is not possible. Instead, various mathematical approximations of the Schrödinger equation are utilised.

There are two major classes of electronic structure methods. The semi-empirical methods use parameters derived from experimental data to simplify the calculations. The *ab initio* methods on the other hand use no experimental parameters in their calculations. Their computations are based solely on the laws of quantum mechanics and several physical constants such as the speed of light, masses and charges of nuclei and Planck's constant. Density functional theory methods are a new class in electronic structure methods. Although they are similar to *ab initio* methods in many ways, they do not try to solve the wave function, but rather, the electron probability density.

A basis set is a mathematical description of the molecular orbitals within a system. It can be interpreted as restricting each electron to a particular region of space. Larger basis sets impose fewer constraints on electrons and result in more accurate calculations.

Most peptides are too large to study accurately using theoretical calculations, even at the lower levels of theory. For this reason, smaller molecules which contain the functional groups and regions of the peptide involved in the reaction are used to represent the entire peptide. Even these molecules are often too large to study at the highest levels of theory available. A compromise must therefore be made between the accuracy of the calculation and the available computational resources.

Most of the reaction systems investigated in this thesis are performed at the HF/6-31G(d)//AM1 level of theory. The structures of minima and transition states are optimised at the semi-empirical AM1 level of theory.[80] The AM1 level has been used previously for geometry optimisations of minima and transition states and demonstrated good correlation with structures obtained using *ab initio* methods.[81] The energies of the optimised structures are calculated using the HF/6-31G(d) *ab initio* method.

Chapter 3:

Negative Ion Cleavages of the Cys Residue

3.1 Introduction

Negative ion cleavages initiated by the Cys residue have not been studied in detail. A major factor for this has been the absence of the Cys residue in host defence peptides of Australian anurans reported so far.[54] Only the negative ion spectra of $(M-H)^-$ parent ions of small peptides such as CG-OH,[82] GC-OH,[82] and GAVC-OH[83] have been investigated. The collision induced mass spectra of the $(M-H)^-$ parent anions of the CG-OH dipeptide forms a base peak corresponding to the loss of H_2S to form a $[(M-H)^- -H_2S]^-$ fragmentation ion. The corresponding spectra of GC-OH and GAVC-OH reveal fragmentations which form both the $[(M-H)^- -H_2S]^-$ and $[(M-H)^- -(H_2S+CO_2)]^-$ ions.

Recent work on *Crinia* species however has revealed major peptide components which contain the Cys residue.[84] This has led to an investigation into Cys initiated cleavages in the negative ion mode. The negative ion spectrum of the Cys containing peptide GLFKVLGCVAKHLLPHVVPVIAEKL-NH_2 **(7)**[84] is displayed in Figure 3.1, and reveals a facile loss of H_2S to form a major peak in the spectrum. All of the α and β cleavages originate exclusively from the $[(M-H)^- -H_2S]^-$ ion rather than the $(M-H)^-$ parent ion. The fragmentation ion at *m/z* 2088 corresponds to the Ile21/Glu23 cyclisation and cleavage, a cleavage ion formed as a result of the α-helical structure of the peptide.[54] (This type of negative ion fragmentation will be discussed in more detail in Chapter 5).

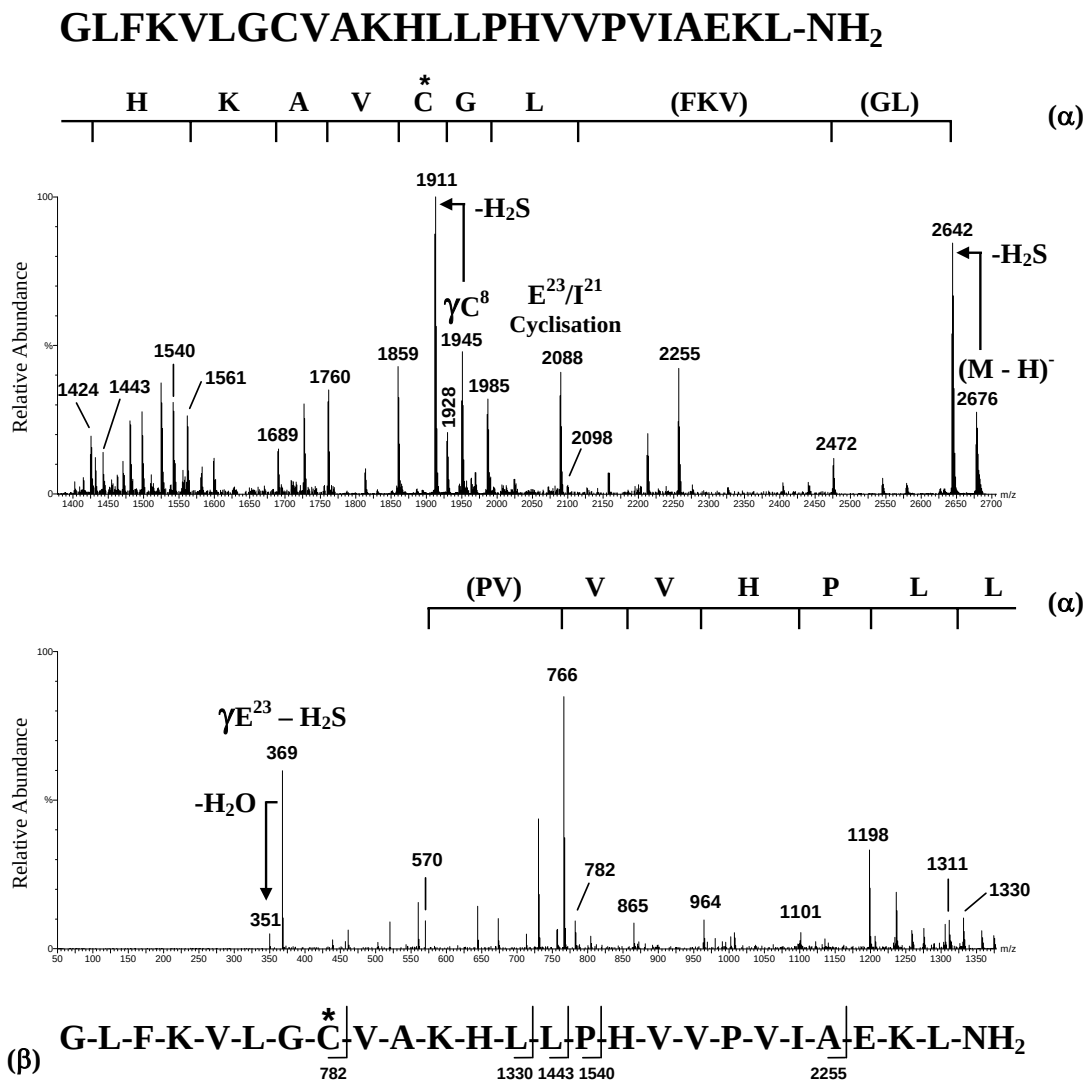


Figure 3.1. Collision-induced MS/MS data for the $(M-H)^-$ parent anion of **7**. α Cleavages are shown above the spectrum with β cleavages displayed on the formula below the spectrum. C* corresponds to a Cys residue following the loss of H_2S and has the structure $NHC(=CH_2)CO$. Magnification ranges: *m/z* 300-1350 (x12), *m/z* 1400-1900 (x16), *m/z* 1940-1980 (x6), *m/z* 1980-2500 (x10).

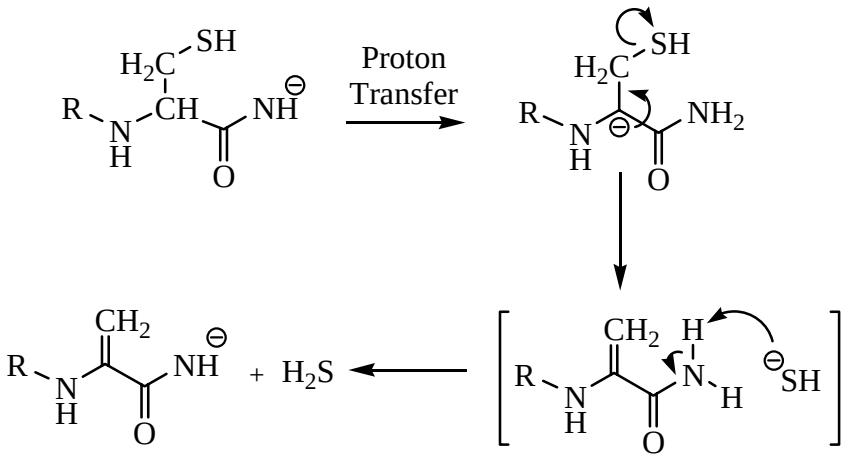
As described previously, the negative ion spectra of Ser containing peptides show facile loss of CH_2O. Since the only difference between the Ser residue and a Cys residue is the substitution of the side chain oxygen with a sulfur group, it would be expected that Cys should exhibit the facile loss of CH_2S. This is not evident from the spectrum in Figure 3.1,

although the $[(M-H)^- -CH_2S]^-$ fragmentation ion observed in some spectra of Cys containing peptides suggests that the side chain loss of CH_2S does occur, but it is a minor process.

3.2 Results and Discussion

3.2.1 The Side Chain Loss of H_2S From Cys

The proposed mechanism for the side chain loss of H_2S from Cys is depicted in Scheme 3.1. It involves transfer of an α proton to the C-terminal $CONH^-$ (or CO_2^-) group to form an enolate anion. This is followed by the loss of an SH^- ion which deprotonates the amide (or carboxylic acid) to form H_2S.



Scheme 3.1. The proposed mechanism for the side chain loss of H_2S from Cys.

This mechanism has been investigated by molecular modelling at the HF/6-31G(d)//AM1 level of theory for a simplified model system. The model system selected is $CH_3NHCH(CH_2SH)CONH^-$ where the N-terminal portion of the peptide has been replaced with CH_3 and the C-terminal portion with H (Figure 3.2). Once the enolate anion has formed, an additional 5 kJ mol^{-1} is required to reach the transition state at 84 kJ mol^{-1}. A stable ion-neutral complex is formed in which HS^- deprotonates the amide. The overall process is endothermic by 64 kJ mol^{-1}. The loss of H_2S is energetically more favourable

than the α, β and β$^{|}$ cleavages which explains why the $[(M-H)^- - H_2S]^-$ ion is often the base peak in the spectrum, and why α, β and β$^{|}$ cleavages originate from the $[(M-H)^- - H_2S]^-$ ion rather than the $(M-H)^-$ parent ion.

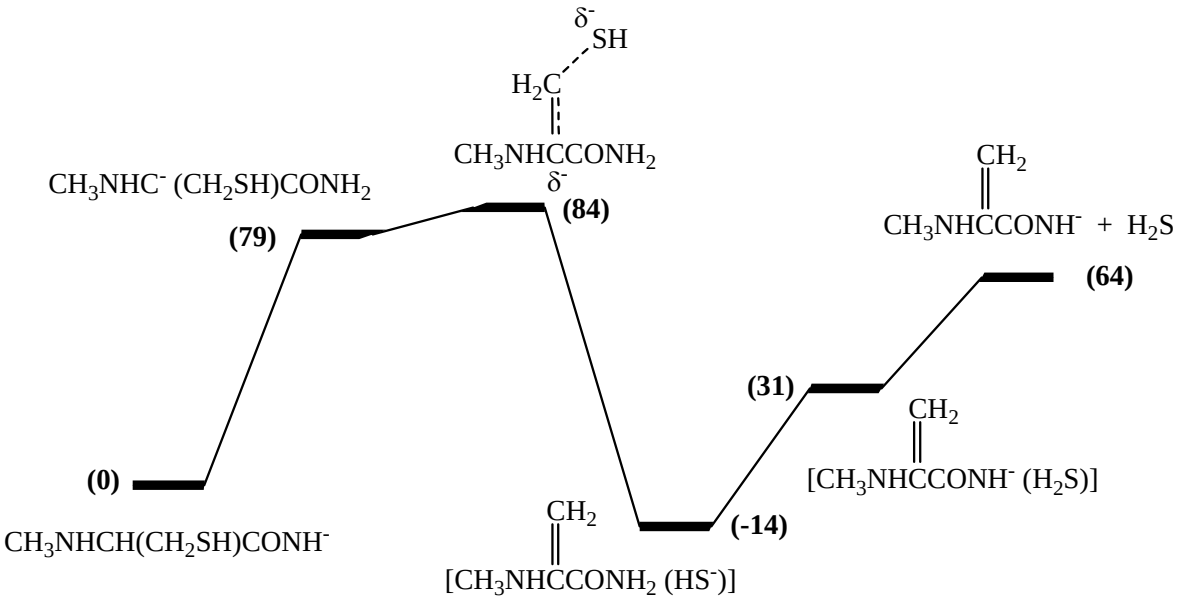


Figure 3.2. The side chain loss of H_2S from a Cys residue in the model system $CH_3NHCH(CH_2SH)CONH^-$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory. Transition states were confirmed by IRC calculations.

The same model system and level of theory was used to calculate the energetics involved in the side chain loss of CH_2S, so that the two processes could be directly compared (Figure 3.3). The hydrogen abstraction from the sulfur onto the amide is exothermic by 63 kJ mol^{-1} but the transition state for the loss of CH_2S is endothermic by 181 kJ mol^{-1}. The overall barrier is 210 kJ mol^{-1} for the model system used. These results reveal that the side chain loss of H_2S from Cys is both kinetically and thermodynamically more favourable than the loss of CH_2S. This is in agreement with the observed experimental results.

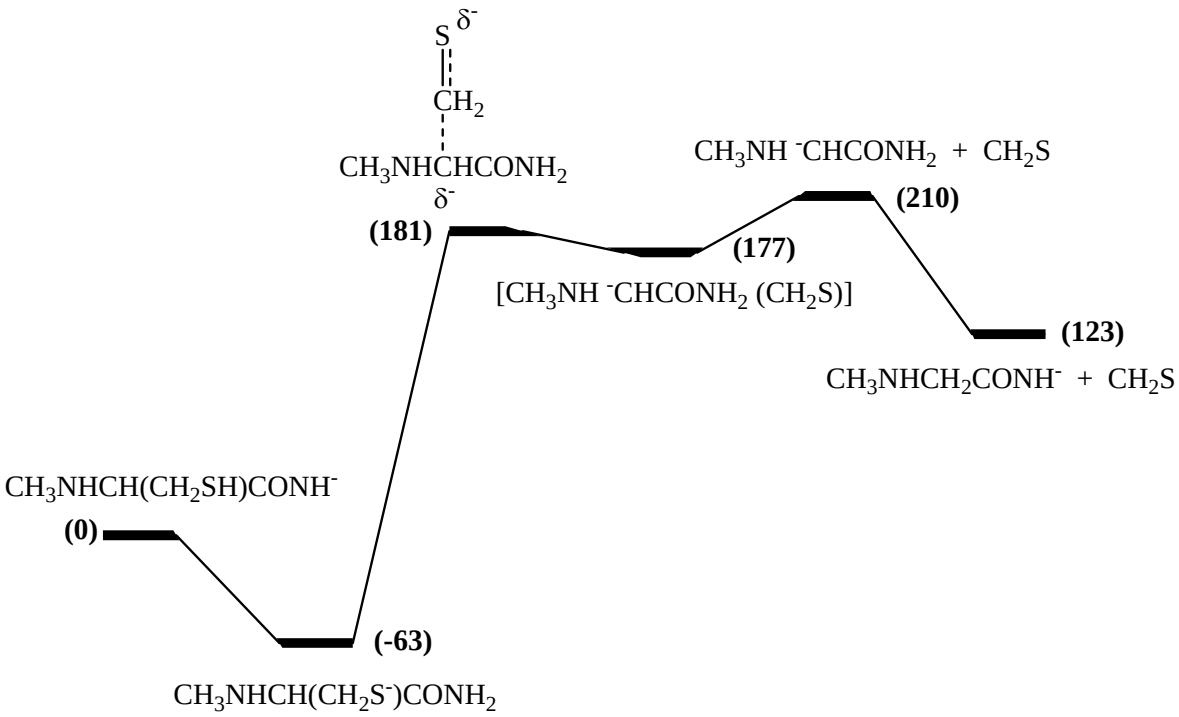


Figure 3.3. The side chain loss of CH_2S from a Cys residue in the model system $CH_3NHCH(CH_2SH)CONH^-$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory. Transition states were confirmed by IRC calculations.

3.2.2 Ser Side Chain Loss Re-Investigated

With the loss of H_2S from the Cys residue both kinetically and thermodynamically more favourable than the loss of CH_2S, the side chain losses from Ser have been re-investigated. The side chain loss of CH_2O has already been reported[78] but a more detailed look at the energetics for this process is shown in Figure 3.4.

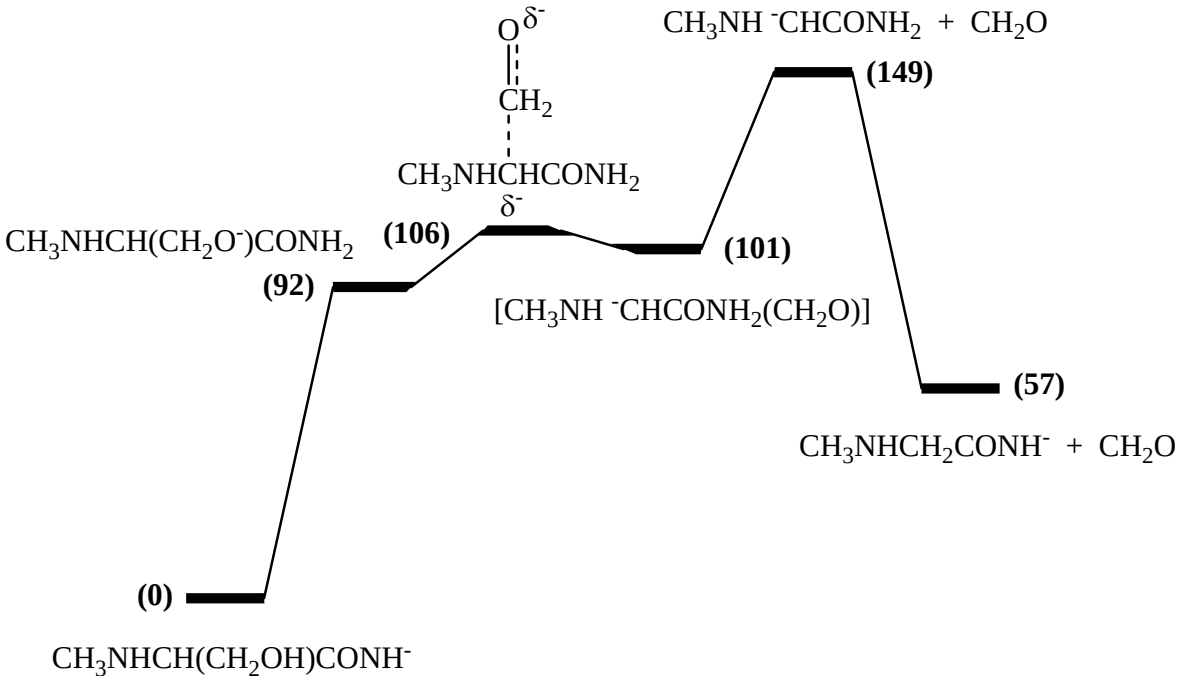


Figure 3.4. The Ser side chain loss of CH_2O using the model system $CH_3NHCH(CH_2OH)CONH^-$. Relative energies in kJ mol^{-1}. Geometries and energies at the AM1 level of theory.[78]

Molecular modelling was conducted for the side chain loss of H_2O from Ser to determine if the theoretical results coincide with the experimental results; i.e. where loss of CH_2O is more facile than loss of H_2O. The calculations were performed using the same model system and level of theory used previously for the loss of CH_2O[78] so that the two processes can be directly compared. The results are shown in Figure 3.5.

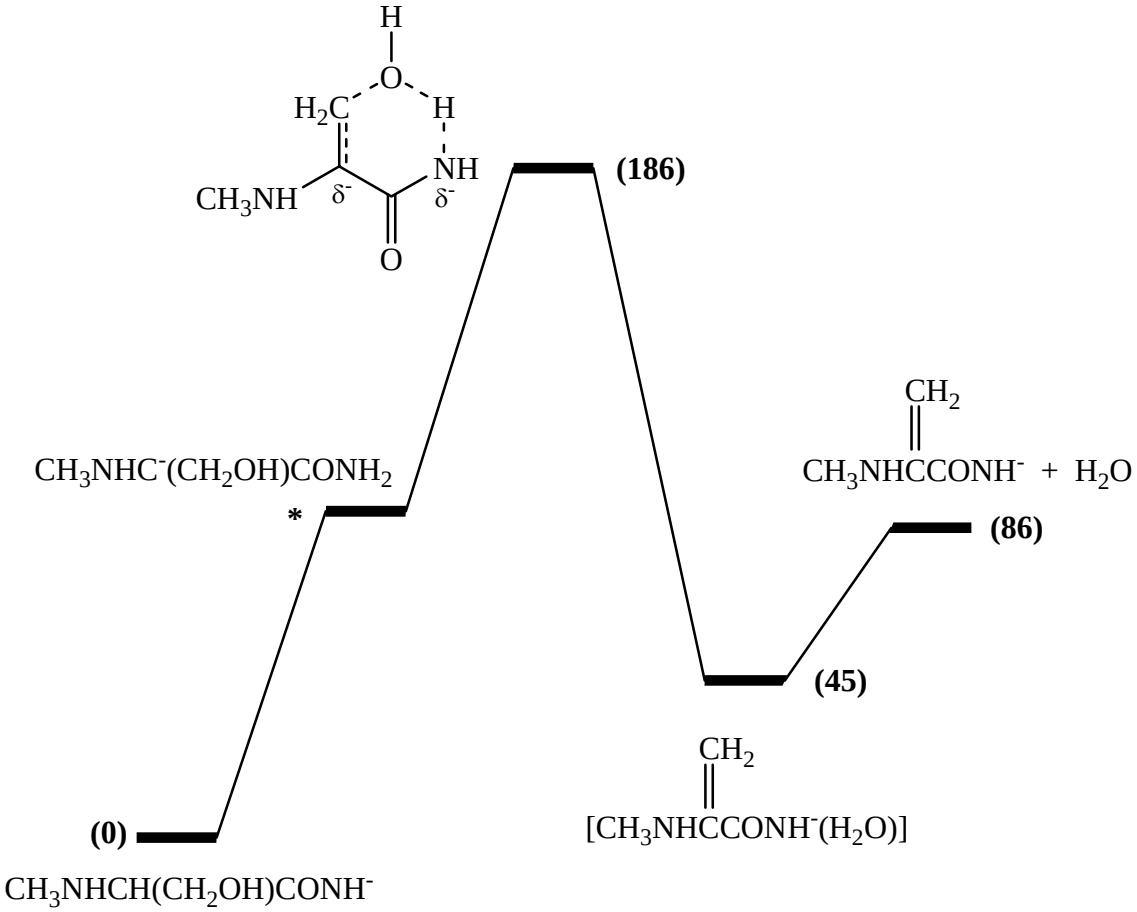


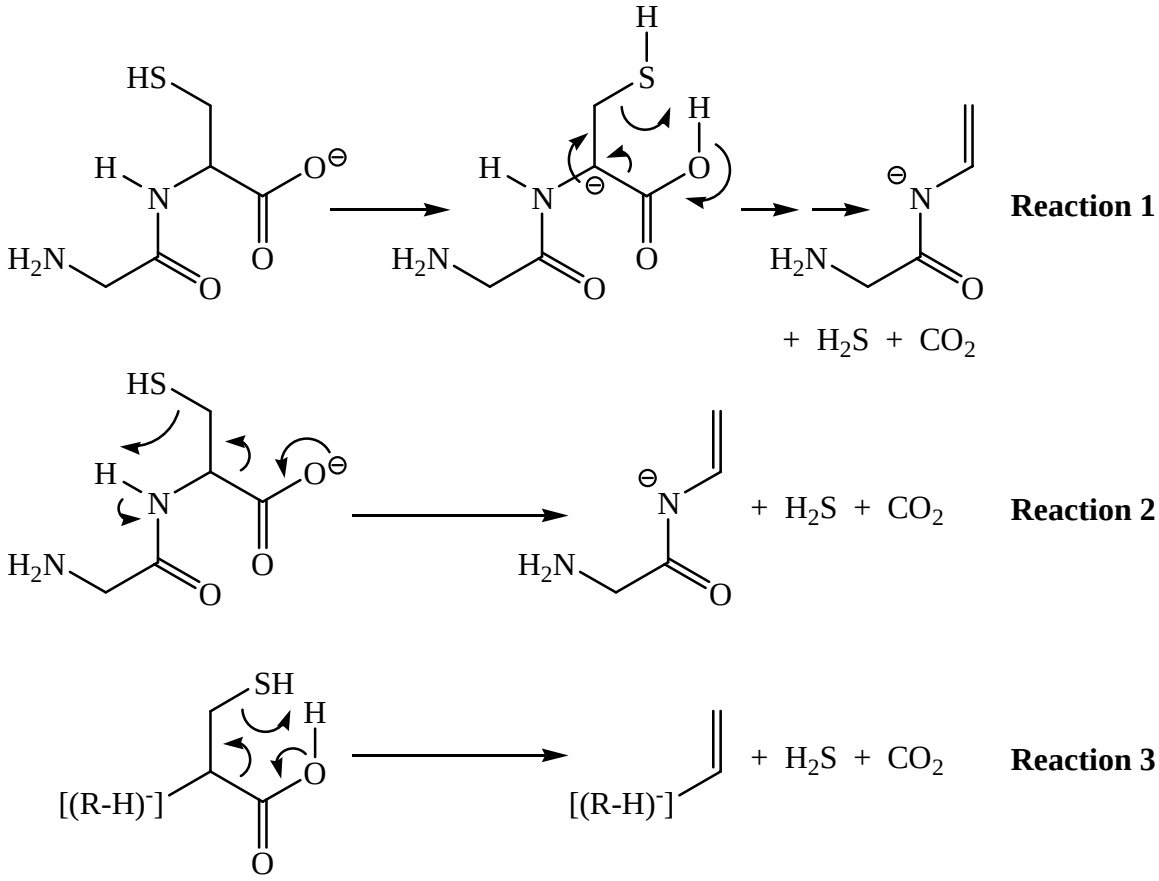
Figure 3.5. The side chain loss of H_2O from a Ser residue in the model system $CH_3NHCH(CH_2OH)CONH^-$. Relative energies in kJ mol^{-1}. Geometries and energies at the AM1 level of theory. Transition states were confirmed by IRC calculations. *****There were difficulties with this calculation not apparent with any of the other calculations conducted thus far. The energy difference between the carboxylate and enolate anions are -11 kJ mol^{-1} at the AM1 level of theory but +99 kJ mol^{-1} at the HF/6-31G(d)//AM1 level of theory. Calculations were run at higher levels of theory to solve this problem. The differences in energy were +103 and +92 kJ mol^{-1} at the HF/6-31G(d)// HF/6-31G(d) and B3LYP/6-31G(d)// B3LYP/6-31G(d) levels of theory respectively. These results compare well with the HF/6-31G(d)//AM1 level of theory.

The loss of H_2O from Ser occurs as shown in Figure 3.5. No lower energy pathway was observed. The cleavage of the carbon-oxygen bond requires 186 kJ mol^{-1} with the overall process being endothermic by 86 kJ mol^{-1}. When this is compared to the side chain

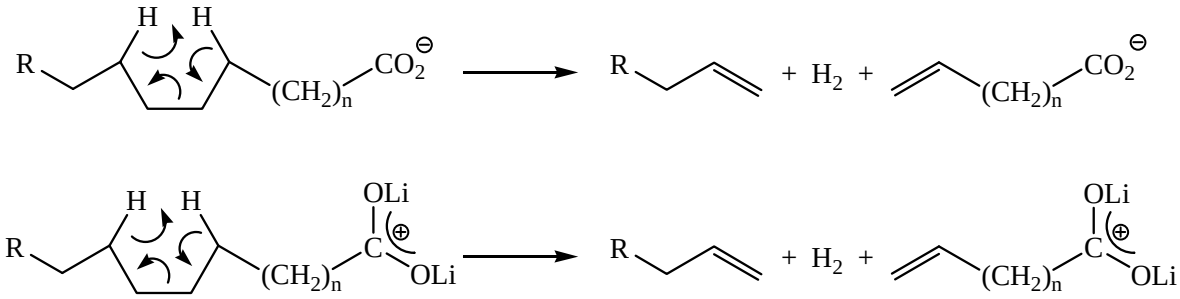
loss of CH_2O, the loss of H_2O is shown to be both kinetically and thermodynamically unfavourable. Theory and experiment are therefore in accord for this system.

3.2.3 Fragmentation of C-Terminal Cys

When a Cys residue occupies the C-terminal position of a peptide, the major fragmentation involves simultaneous loss of both H_2S and CO_2 to form the $[(M-H)^- -(H_2S+CO_2)]^-$ fragment ion.[82,83] Three possible mechanisms are proposed for the fragmentation. The first, Reaction 1 of Scheme 3.2, is where the enolate anion initiates either a concerted or stepwise loss of H_2S and CO_2. The second process (Reaction 2, Scheme 3.2) is again a charge-induced process, but this time initiated from the carboxylate anion. The process may be concerted as shown in Scheme 3.2 or may be stepwise. An alternative process may involve a charge-remote reaction (Reaction 3, Scheme 3.2). A charge-remote process is a reaction which occurs remote from and uninfluenced by the charged centre in the molecule.[85-87] Such reactions have been extensively studied by Gross *et al*.[86,87] The work originally focussed on cleavages of fatty acids under both negative[88] and positive[89] ion conditions. Scheme 3.3 illustrates the charge-remote mechanism from the $(M-H)^-$ and $(M+2Li-H)^+$ parent ions of a fatty acid. The nature of the charge is not important,[90] however, charge-remote reactions normally only occur when charge-directed cleavages are energetically unfavourable.[85]



Scheme 3.2. Reaction 1: The charge-induced cleavage of C-terminal Cys, initiated by the enolate anion. Reaction 2: The charge-induced cleavage of C-terminal Cys, initiated by the carboxylate anion. Reaction 3: The charge-remote cleavage of C-terminal Cys.



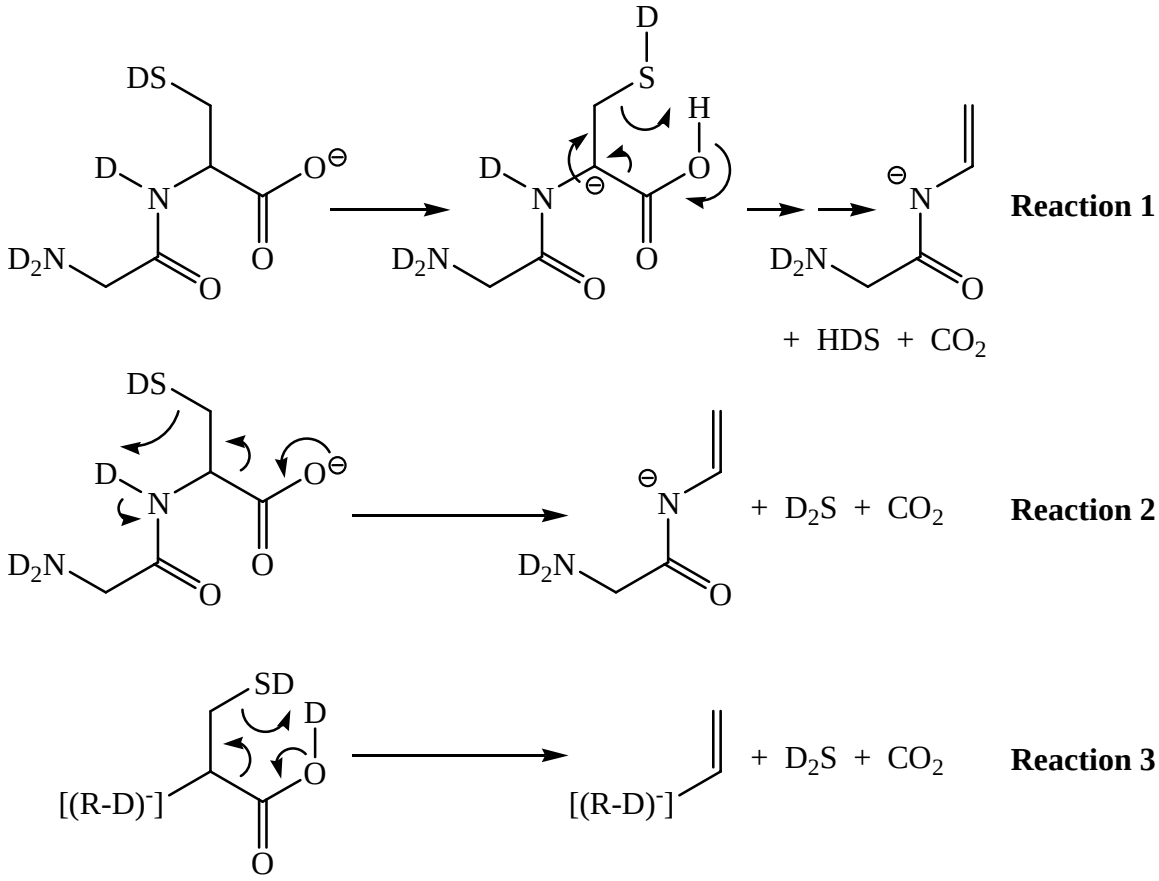
Scheme 3.3. Charge-remote mechanism of a fatty acid under negative[88] and positive[89] ion conditions.

In an attempt to exclude or confirm one or more mechanisms displayed in Scheme 3.2, the GC-OH peptide was deuterium labelled and re-submitted to mass spectrometric analysis. The MS/MS data for the GC-OH peptide and deuterium analogue are displayed in Table 3.1.

Table 3.1. Negative ion MS/MS of GC-OH and deuterium-labelled analogue [*m/z* (loss) relative abundance (%)].

$[NH_2CH_2CONHCH(CH_2SH)CO_2H - H]^-$	177 $[M-H]^-$, 143 (H_2S) 54, 131 (CH_2S) 14, 99 (H_2S+CO_2) 100.
$[ND_2CH_2CONDCH(CH_2SD)CO_2D - D]^-$	181 $[M-D]^-$, 146 (HDS) 45, 135 (CH_2S) 11, 102 ($HDS+CO_2$) 6, 101 (D_2S+CO_2) 100.

The results displayed in Table 3.1 reveal that there are two alternative pathways for the loss H_2S and CO_2 from a Cys C-terminal peptide. Examination of the MS/MS data for the deuterium labelled analogue indicates that a minor pathway leads to the loss of HDS and CO_2 while a major mechanistic pathway produces $[(M-D)^- - (D_2S+CO_2)]^-$ as the base peak in the spectrum. The three proposed mechanisms in Scheme 3.2 are re-considered in deuterium labelled analogue form, as shown in Scheme 3.4.



Scheme 3.4. Deuterium labelled analogues of the three proposed mechanisms for fragmentation of a C-terminal Cys residue.

The Reaction 1 pathway of Scheme 3.4 produces the loss of HDS and CO_2. This is consistent with the minor process as indicated in Table 3.1. For this enolate directed reaction to lose D_2S and CO_2, it would require at least two H^+ and/or D^+ transfers prior to or during the elimination of D_2S and CO_2, an unlikely event. Both Reaction 2 (carboxylate anion induced fragmentation – concerted or stepwise) and Reaction 3 (charge-remote fragmentation) of Scheme 3.4 produce the loss of D_2S and CO_2. The loss of D_2S and CO_2 is consistent with the major fragmentation pathway. This means that both are still plausible mechanisms by which a C-terminal Cys residue loses H_2S and CO_2 in the negative ion mode. Theoretical studies were considered on a model system to shed light on which pathway was most likely.

The charge-directed mechanism was investigated at the HF/6-31G(d)//AM1 level of theory on the model system $HCONHCH(CH_2SH)CO_2^-$. The N-terminal portion of the molecule has been replaced with H. The results of the theoretical studies are shown in Figure 3.6. The charge-induced cleavage occurs in a stepwise manner originating from the carboxylate anion. The overall process is endothermic by 138 kJ mol^{-1} with a maximum barrier of 281 kJ mol^{-1}.

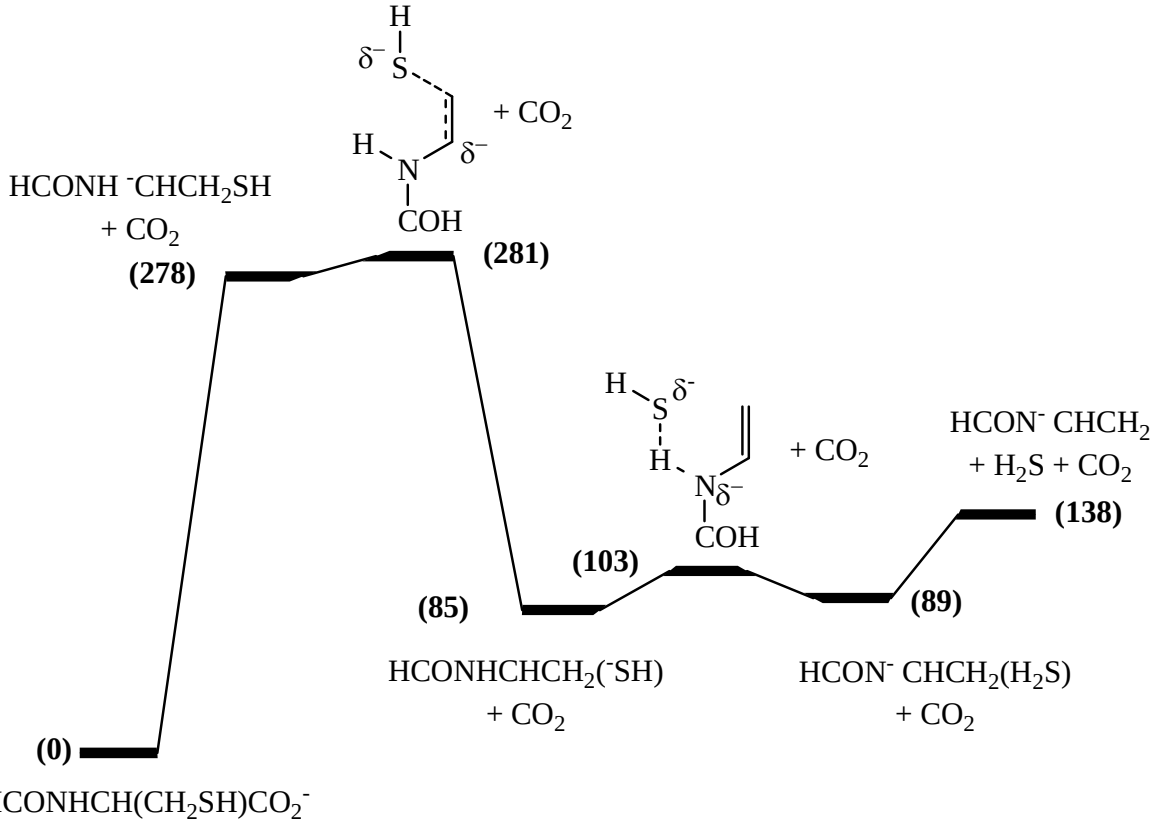


Figure 3.6. The charge-induced loss of H_2S and CO_2 for a C-terminal Cys residue in the model system $HCONHCH(CH_2SH)CO_2^-$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory. Transition states were confirmed by IRC calculations.

In comparison, the charge-remote process occurs via a concerted process and is shown in Figure 3.7. The barrier to the transition state is higher than that of the charge induced reaction by 52 kJ mol^{-1}. However, the charge-remote process is the more thermodynamically favourable of the two (-33 kJ mol^{-1} compared with +138 kJ mol^{-1} for the charge-induced process).

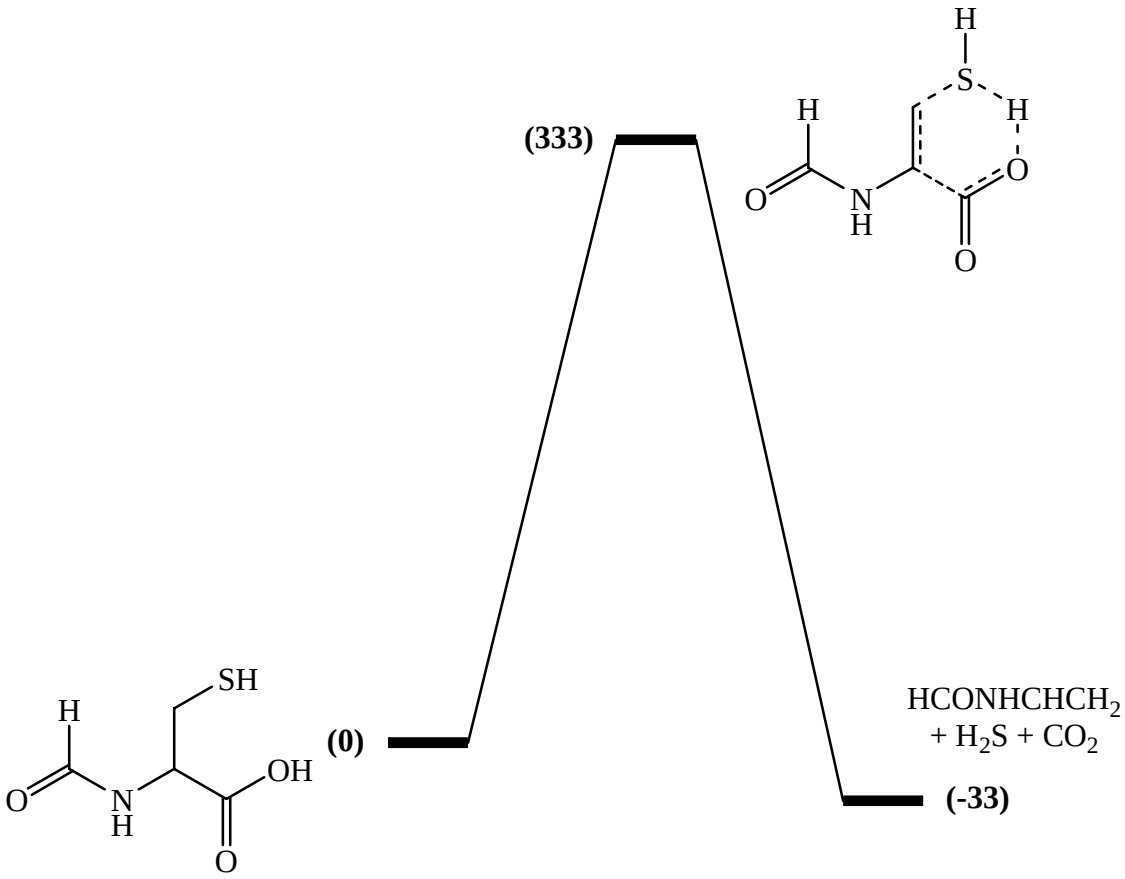


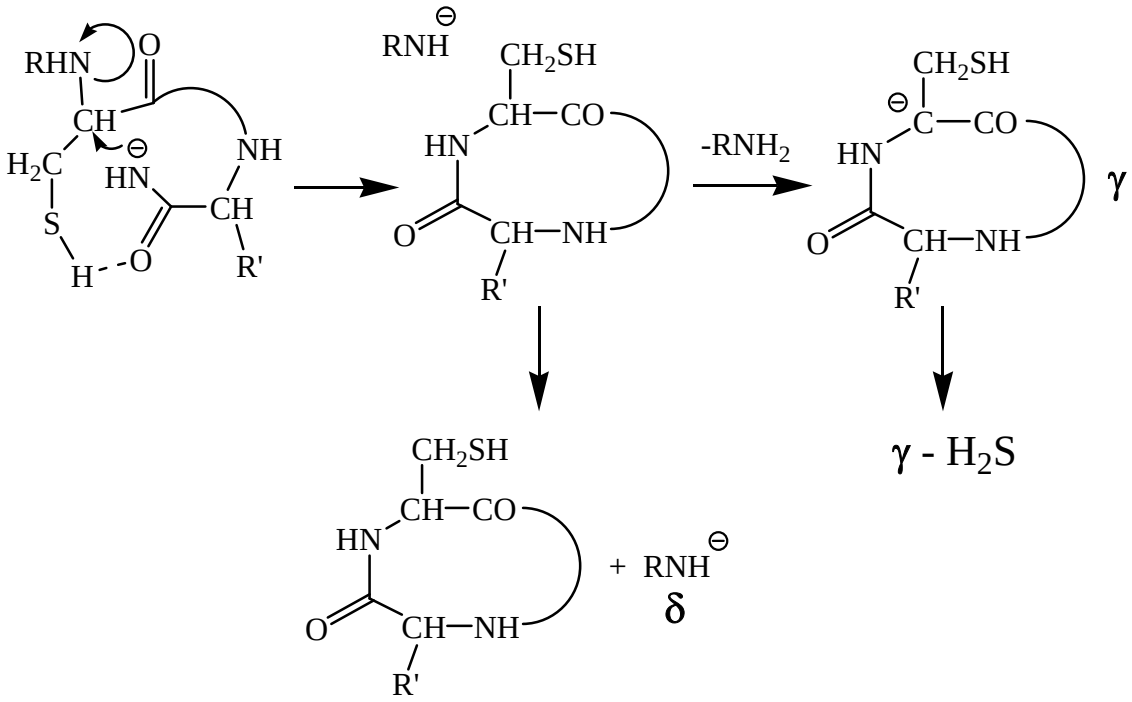
Figure 3.7. The charge-remote loss of H_2S and CO_2 for a C-terminal Cys residue in the model system $HCONHCH(CH_2SH)CO_2H$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory. Transition state confirmed by IRC calculations.

Comparing the energetics of the two processes reveals that the charge-induced process is kinetically favourable while the charge-remote process is thermodynamically

favourable. The results so far obtained do not allow for the unequivocal determination for which of the two processes (or both) actually takes place.

3.2.4 γ and δ Cleavages of Cys

The spectrum shown in Figure 3.1 of the Cys containing peptide **7** reveals a fragmentation ion at *m/z* 1945 which corresponds to the γ backbone cleavage initiated by the Cys residue. The γ fragmentation of Ser has already been examined and the Cys γ fragmentation is likely to occur in a similar manner. Scheme 3.5 illustrates the proposed mechanism for the Cys γ cleavage.



Scheme 3.5. The proposed mechanism for the Cys induced formation of the δ, γ and γ-H_2S ions.

The fragmentation involves the C-terminal amide oxygen forming a hydrogen bond with the SH side group of Cys. This places the amide nitrogen anion in close proximity to the electrophilic α-carbon of the Cys residue. The amide nucleophile can initiate nucleophilic attack onto the electrophile to form an ion-neutral complex. Dissociation of the ion results in a δ fragmentation ion, with proton transfer followed by dissociation of the ion-neutral complex providing the γ fragment ion. The peak at *m/z* 1911 in the spectrum of Figure 3.1 is 34 Da less than the Cys γ fragmentation ion and is a result of the loss of H_2S from the γ ion.

Theoretical calculations have been conducted for these fragmentations using the HF/6-31G(d)//AM1 level of theory. The selected model system for this investigation involves $CH_3CONHCH_2(CH_2SH)CHO$ and $CH_3NHCH_2CONH^-$. They represent the Cys residue and the C-terminal amide respectively. The formation of the δ and γ anions are illustrated in Figure 3.8. The key to this process is the formation of the hydrogen bond between the oxygen of the C-terminal amide with the SH group of the Cys side chain. This intermediate (exothermic by 60 kJ mol^{-1}) locks the C-terminal amide nitrogen into a position where it can initiate attack onto the electrophilic Cys α-carbon to form an ion-neutral complex which is exothermic by 84 kJ mol^{-1}. The transition state barrier for this S_N2 process is only 132 kJ mol^{-1}. Dissociation of the ion-neutral complex furnishes the δ ion, a process endothermic by 37 kJ mol^{-1}. Alternatively, proton transfer followed by dissociation forms the γ ion in a reaction exothermic by 40 kJ mol^{-1}. The theoretical results agree with the experimental observations; i.e. γ backbone cleavages are more common than the corresponding δ ions for Cys.

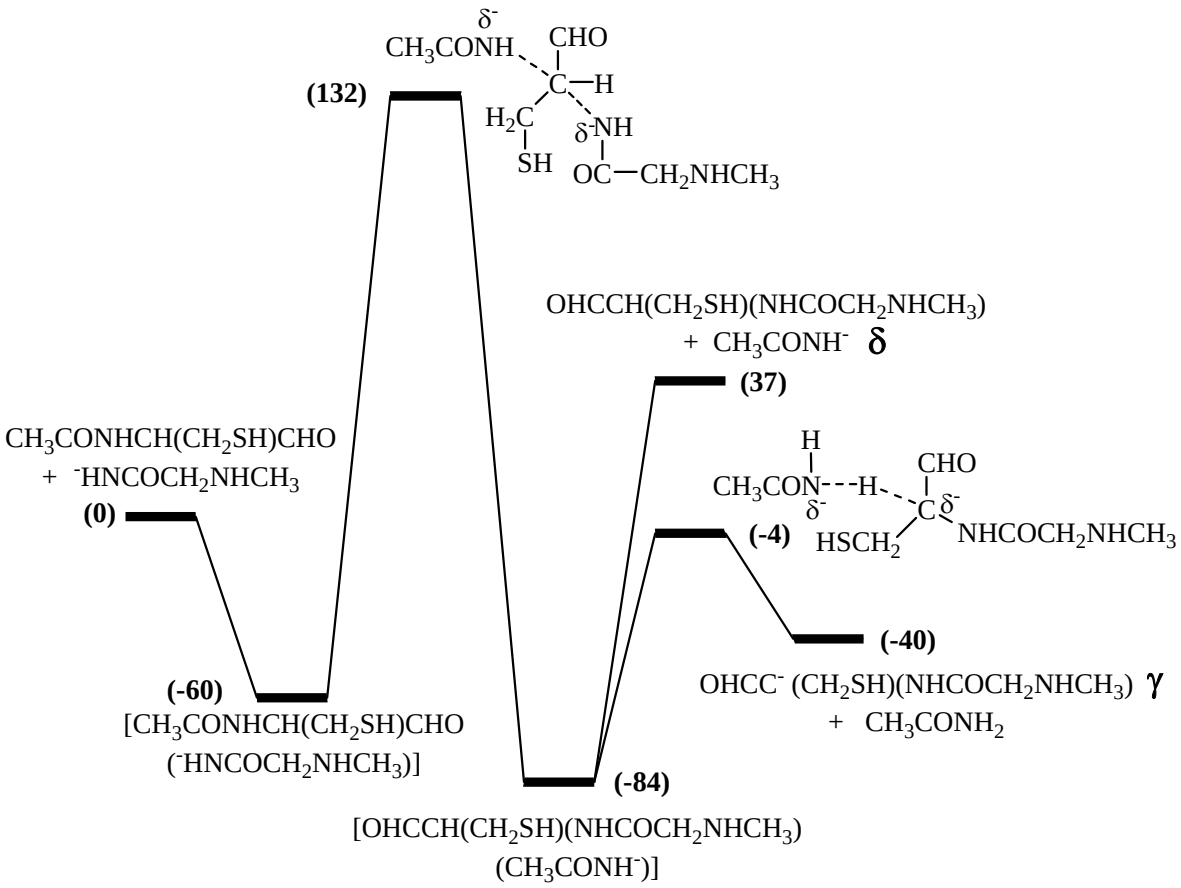


Figure 3.8. The formation of the Cys γ and δ fragmentation ions using the model system $CH_3CONHCH_2(CH_2SH)CHO$ and $CH_3NHCH_2CONH^-$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory. Transition states confirmed by IRC calculations.

Following formation of the γ ion, an additional 68 kJ mol^{-1} will overcome the barrier to the Cys side chain cleavage of the SH group to form the ion-neutral complex (-77 kJ mol^{-1}) shown in Figure 3.9. From this point the loss of H_2S may occur by one of two processes. The first is deprotonation to form the enolate anion at 74 kJ mol^{-1} or the formation of the amide anion which is the more favourable of the two processes, by 112 kJ mol^{-1}.

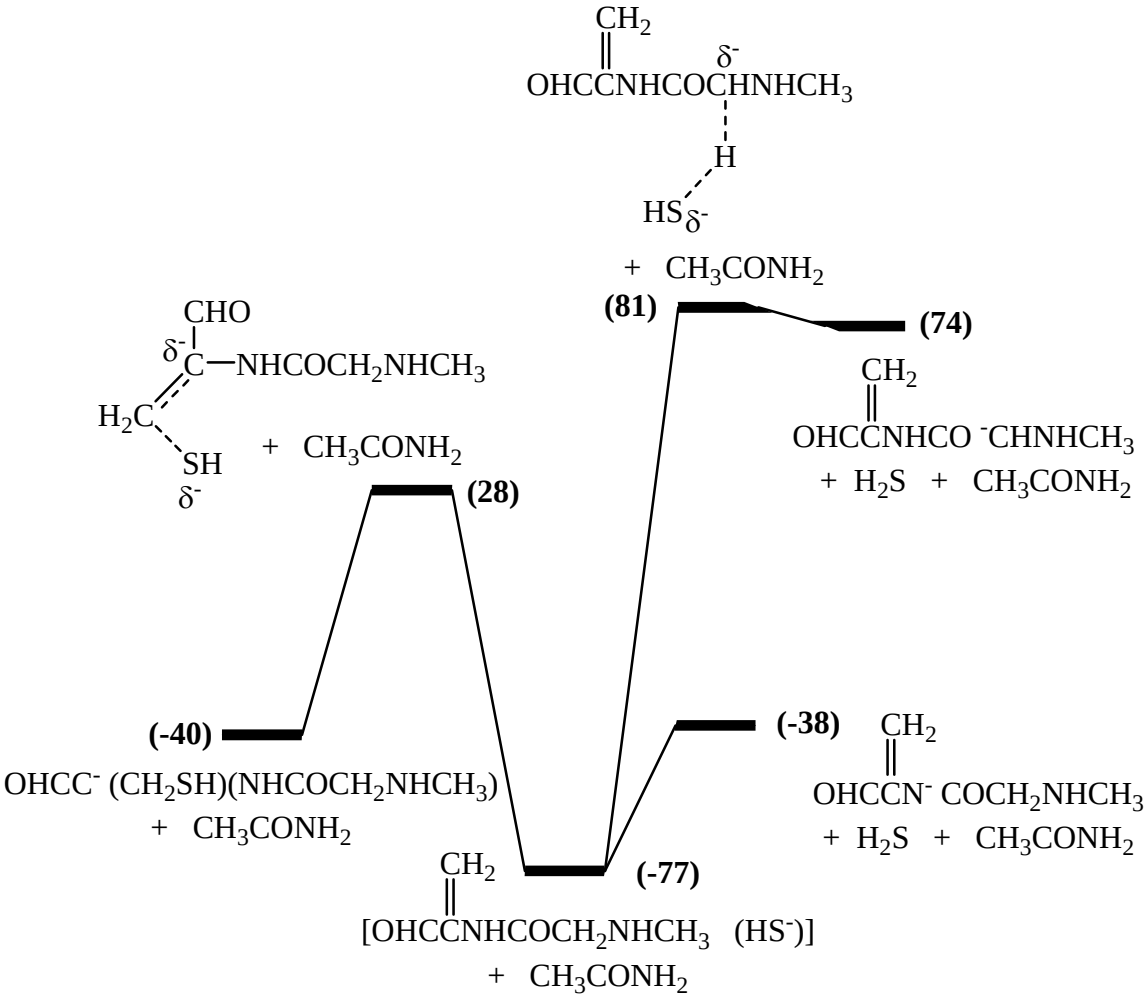


Figure 3.9. The additional loss of H_2S from the Cys γ fragmentation ion using the model system $CH_3CONHCH_2(CH_2SH)CHO$ and $CH_3NHCH_2CONH^-$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory. Transition states confirmed by IRC calculations.

The negative ion mass spectral data of several Cys containing peptides are summarised in Table 3.2.

Table 3.2. Negative ion fragmentation data for Cys containing peptides. C* corresponds to a Cys residue following the loss of H_2S and has the structure $NHC(=CH_2)CO$.

GLFAVIKKVACVIGGL-NH_2

$(M-H)^-$ *m/z* 1584.

$[(M-H)^- -H_2S]^-$ *m/z* 1550.

α ions from $[(M-H)^- -H_2S]^-$ *m/z* 1380, 1233, 1162, 1063, 950, 822, 694, 595, 524, 455, 356, 243, 186 [(GL)FAVIKKVAC* VIG].

$\beta^|$ ions from $[(M-H)^- -H_2S]^-$ *m/z* 1437, 1380, 1111, 1042, 971, 872, 744, 616 [KKVAC* (VIG)GL].

γC^{11} *m/z* 541, γC^{11}-H_2S *m/z* 507, γF^3 from $[(M-H)^- -H_2S]^-$ *m/z* 1363.

Sequence: (GL)FAVIKKVACVIGGL-NH_2.

GLFDVIKKVACVIGGL-NH_2

$(M-H)^-$ *m/z* 1628.

$[(M-H)^- -H_2S]^-$ *m/z* 1594.

α ions from $[(M-H)^- -H_2S]^-$ *m/z* 1277, 1162, 1063, 950, 822, 694, 595, 524, 455, 356, 243, 186 [(GLF)DVIKKVAC* VIG].

γC^{11} *m/z* 541, γC^{11}-H_2S *m/z* 507, δC^{11} *m/z* 1086, γD^4 *m/z* 1294, γD^4-H_2O *m/z* 1276, γD^4-(H_2O+H_2S) *m/z* 1242, δD^4 *m/z* 333.

Sequence: (GLF)DVIKKVACVIG(GL)-NH_2.

GLLGLLGCVVCHVVPAIVGHF-NH_2

$(M-H)^-$ *m/z* 2099.

$[(M-H)^- -2H_2S]^-$ *m/z* 2031.

α ions from $[(M-H)^- -2H_2S]^-$ *m/z* 1861, 1748, 1691, 1578, 1465, 1408, 1339, 1240, 1141, 1072, 935, 737, 357 [(GL)LGLLGC* VVC* H(VV)(PAIV)-(GHF)-NH_2].

$\beta^|$ ions from $[(M-H)^- -2H_2S]^-$ *m/z* 1747, 1310, 906, 708, 639, 582, 469, 356, 299, 186 [(GLL)GLLGC*(VV)(C* HVV)(PAIVG)(HF)-NH_2].

γC^8 *m/z* 1459, γC^8-H_2S *m/z* 1425, δC^8 *m/z* 639, γC^{11} *m/z* 1158, γC^{11}-H_2S *m/z* 1124.

Sequence: (GL)LGLLGCVVCH(VV)(PAIV)G(HF)-NH_2.

Table 3.2 (Continued).

GLLCVLGCVAKHVLPHVVPVIAEHL-NH_2

$(M-H)^-$ *m/z* 2612.

$[(M-H)^- -2H_2S]^-$ *m/z* 2544.

α ions from $[(M-H)^- -2H_2S]^-$ *m/z* 2261, 2192, 2093, 1980, 1923, 1854, 1755, 1684, 1556, 1419, 1207, 973, 874, 775, 678, 579 [(GLL)C* VLGC* VAKH(VL)(PH)VVPV(IAEHL)-NH_2].

$\beta^|$ ions from $[(M-H)^- -2H_2S]^-$ *m/z* 1981, 1882, 1785, 1686, 1587 [(1640)VVPV(IAEHL)-NH_2].

γC^4 *m/z* 2312, γC^4-H_2S *m/z* 2278, γC^8 *m/z* 1940, γC^8-H_2S *m/z* 1906, γE^{23} *m/z* 378.

Sequence: (GLL)CVLGCVAKH(VL)(PH)VVPV(IA)E(HL)-NH_2.

3.3 Summary and Conclusions

1. Negative ion MS/MS of the $(M-H)^-$ parent ions of Cys containing peptides result in the side chain loss of H_2S to form the $[(M-H)^- -H_2S]^-$ daughter ion. This process is extremely facile, requiring just 84 kJ mol^{-1} to overcome the highest transition state barrier. The overall fragmentation is endothermic by just 64 kJ mol^{-1} at the HF/6-31G(d)//AM1 level of theory on the model system used. The loss of H_2S is kinetically and thermodynamically more favourable than the α and β fragmentation pathways. This explains why α and β fragmentations of Cys containing peptides originate from the $[(M-H)^- -H_2S]^-$ ion rather than the $(M-H)^-$ parent ion.

2. The side chain loss of CH_2S from a Cys residue is a minor process. The loss of H_2S is both kinetically (by 126 kJ mol^{-1}) and thermodynamically (by 59 kJ mol^{-1}) more favourable than loss of CH_2S.

3. When a Cys residue is in the C-terminal position of a peptide, the negative ion MS/MS spectrum is dominated by the formation of an $[(M-H)^- -(H_2S+CO_2)]^-$ fragment anion. Following theoretical calculations on a model system, two mechanisms are plausible for this fragmentation. The kinetically favoured process is a stepwise, carboxylate anion, charge-induced process. The highest transition state barrier is 281 kJ mol^{-1} with the overall process endothermic by 138 kJ mol^{-1}. The concerted charge-remote mechanism is thermodynamically favoured. The transition state barrier is 333 kJ mol^{-1} with the loss of H_2S and CO_2 exothermic by 33 kJ mol^{-1} at the HF/6-31G(d)//AM1 level of theory.

4. The side chain loss of H_2S from the $(M-H)^-$ parent ion of a peptide confirms the presence of the Cys residue. The position of the Cys residue is determined by the formation of the γ ion. This charge induced process is initiated by the formation of a hydrogen bond between the Cys side chain SH and the carbonyl of a C-terminal amide or carboxylic acid moiety. The formation of the hydrogen bond is exothermic by 60 kJ mol^{-1} and juxtaposes the nucleophilic C-terminal amide nitrogen (or carboxylate anion in the case of a C-terminal carboxylic acid) in a position where it can initiate S_N2 attack on the electrophilic α-carbon of the Cys residue. The highest transition state barrier for this process is only 132 kJ mol^{-1} with the formation of the γ ion exothermic by 40 kJ mol^{-1}. (Formation of the δ ion is endothermic by 37 kJ mol^{-1}). Formation of an enolate anion on the Cys α-carbon following γ ion formation can result in an additional loss of H_2S from the γ ion.

3.4 Experimental

3.4.1 Peptide Synthesis

All peptides used in this study were synthesised by Mimotopes (Clayton, Victoria, Australia) using L-amino acids via the *N*-α-Fmoc method as reported.[91]

3.4.2 Mass Spectrometry

Electrospray MS/MS data were determined using a Micromass Q-Tof 2 orthogonal acceleration time of flight mass spectrometer with a mass range to 10 000 Da. The Q-Tof 2 is fitted with an electrospray source in an orthogonal configuration with a Z-Spray interface. Samples were dissolved in acetonitrile/water (1:1) and infused into the electrospray source at a flow rate of 5 μL min^{-1}. Conditions were as follows: capillary voltage, 3.10 kV; source temperature, 80 ºC; desolvation temperature, 150 ºC; cone voltage, 50 V. MS/MS data were acquired with the argon collision gas energy set between 40-80 eV to give optimal fragmentations.

3.4.3 Theoretical Calculations

Geometry optimisations were carried out using the semi-empirical AM1[80] method with the MAC SPARTAN program.[92] Stationary points were characterised as either minima (no imaginary frequencies) or transition states (one imaginary frequency) by calculation of frequencies using analytical gradient procedures. The AM1 level of

calculations has been used previously for geometry optimisation and transition structure searches and has demonstrated good correlation with structures obtained using *ab initio* methods.[81] Transition states were confirmed using Intrinsic Reaction Coordinate (IRC) calculations. Energies of transition states and reactive intermediates have been calculated at the HF/6-31G(d)//AM1 level of theory using the GAUSSIAN 98 suite of programs[93] on the Alpha Server SC [Australian Partnership for Advanced Computing (APAC) National Facility, Canberra, Australia].

Chapter 4:

Negative Ion Cleavages of Intramolecular Disulfide Bridged Peptides

4.1 Introduction

One of the most common post translational modifications in peptides and proteins is the formation of a disulfide bridge between two Cys residues.[94-99] The two Cys residues are covalently linked by the sulfur atom of each residue. Disulfide bridges are important in establishing and maintaining secondary and tertiary structure, essential for biological activity.[94-99] Many biologically active peptides such as insulin contain one or more disulfide bridges.[98]

Over 200 peptides have been identified from around 30 species of Australian amphibians studied thus far.[6] The diversity among the sequences of these peptides is immense. They vary in size from just four residues long in the case of tryptophyllins (isolated from *Litoria electrica*[77] and *Litoria rubella*[100]) up to 30 residues for rotheins and uperins (isolated from *Litoria rothii*[101] and *Uperoleia inundata*[34] respectively). Many have post-translational modifications to the C-terminus forming C-terminal amides. Others have post translational pyro-Glu N-terminal residues. This is a result of cyclisation of N-terminal Glu and more commonly Gln, eliminating H_2O and NH_3 respectively in the cyclisation process. On the other hand, the caerulein peptides (found in several species of Australian amphibians)[2,11,49,101-107] have a post translational sulfated Tyr.

Only recently have studies of the Australian Common Froglet *Crinia signifera* identified an Australian amphibian peptide containing a disulfide bridge.[75] This peptide is named signiferin 1 and has the sequence RLCIPYIIPC-OH **(8)** where the underline between the two Cys residues identifies the disulfide link. This linkage between the Cys residues forces the peptide to adopt a horse shoe shaped secondary structure (Figure 4.1).

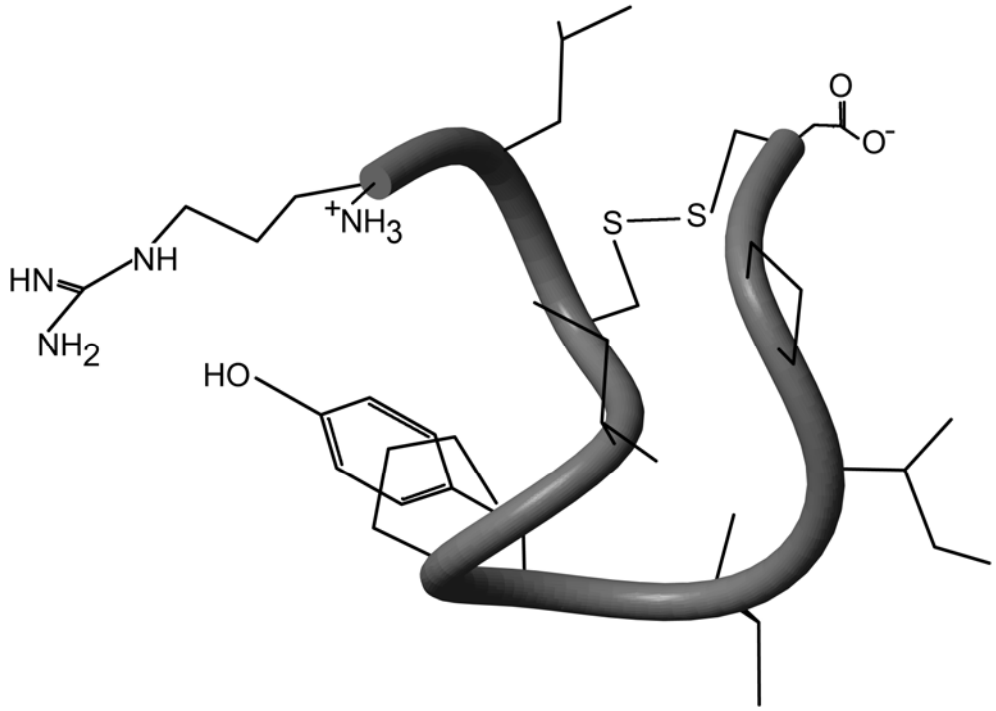


Figure 4.1. The 3-dimensional structure of signiferin 1, RLCIPYIIPC-OH **(8)** as determined by 2D NMR spectroscopy.[108]

Full sequencing of intramolecular disulfide linked peptides using positive ion mass spectrometry is difficult. Fragmentations are normally limited to residues outside the disulfide bridge.[51,109-115] Cleavage of an amide bond within the disulfide link does not form a fragment ion because the N-terminal portion of the peptide remains bound to the C-terminal end via the disulfide bridge. For this reason, sequencing information is rarely provided for residues within the disulfide link.[51]

The positive ion MS/MS spectrum of signiferin 1 **(8)** is shown in Figure 4.2. The spectrum is complex. The expected Y + 2 fragmentations for the loss of the first two residues (outside the disulfide bridge) are not evident. Instead, it appears that the amide bond between the Cys3 and Ile4 residues is cleaved with the N-terminal portion of the peptide remaining bound to the C-terminal portion through the disulfide bridge. This cleavage is followed by a second amide bond cleavage after the Ile or Pro or Tyr etc. to give the sequence IPYII. Examination of the B fragmentations reveal that the amide bond between the Ile8 and Pro9 residues is cleaved first. The peptide C-terminal end remains bound to the N-terminal end and additional amide cleavages provide the sequence as given above the spectrum in Figure 4.2.

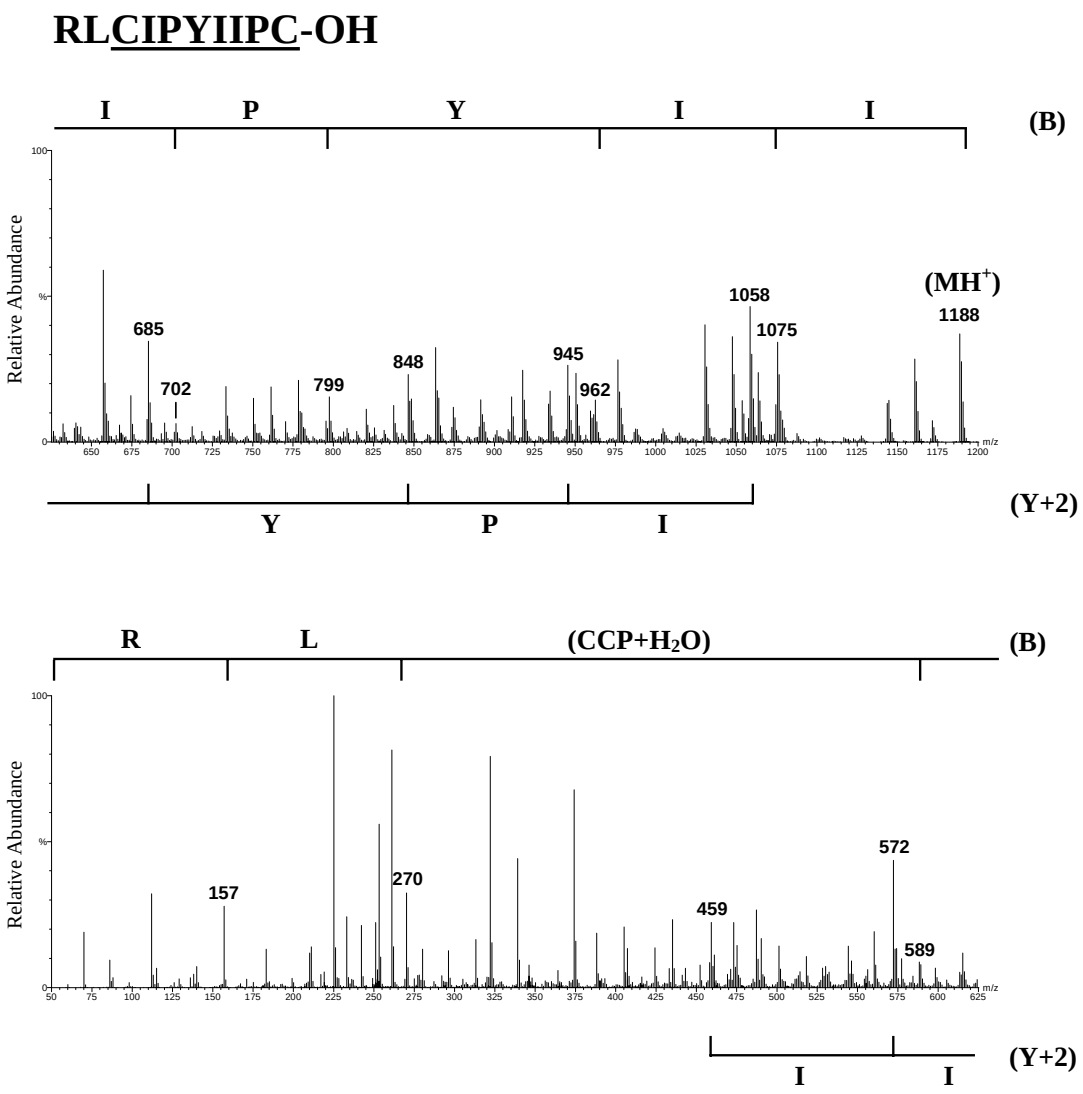
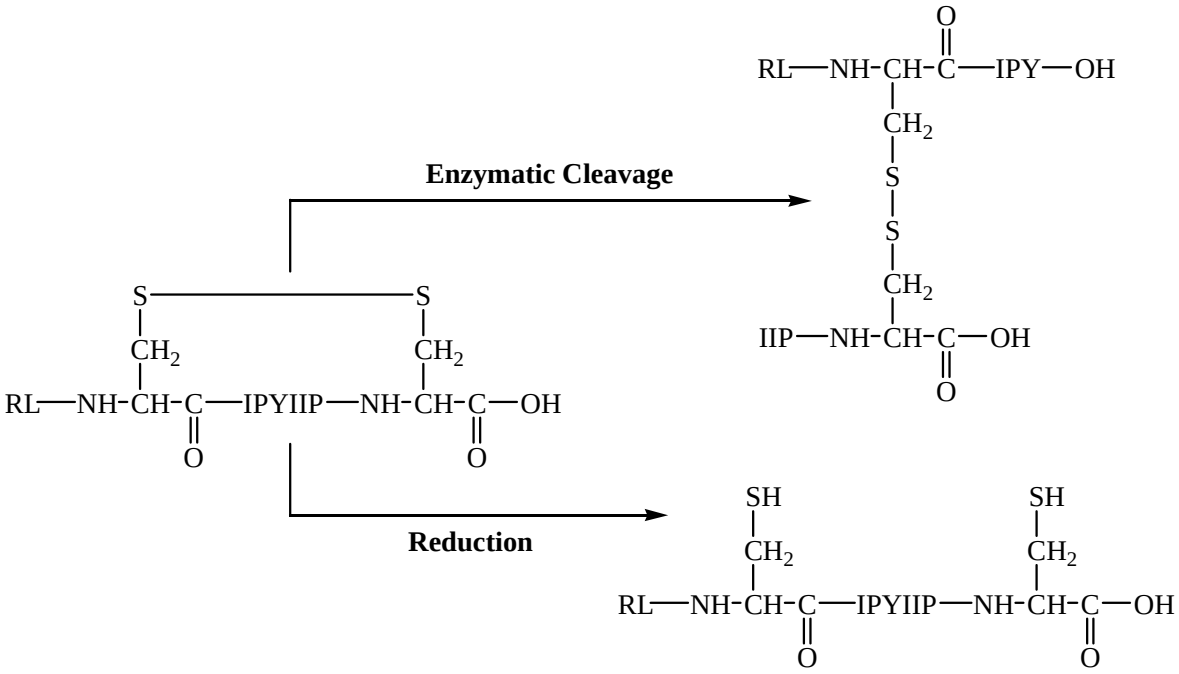


Figure 4.2. Collision-induced MS/MS data of **8**. Positive ion fragmentations of the $(MH)^+$ species, B and Y + 2 cleavage ions are shown above and below the spectrum respectively. Magnification ranges: *m/z* 450-1050 (x2), *m/z* 1050-1100 (x4).

The search for improved methods of sequencing disulfide linked peptides has taken place for over 20 years.[116] During this time, two main methods have proved most successful. They are summarised in Scheme 4.1 using RLCIPYIIPC-OH **(8)** as an example.



Scheme 4.1. The two main approaches used to assign sequences of intramolecular disulfide bridged peptides.

4.1.1 Intramolecular Disulfide Bridged Peptide Sequencing by Enzymatic Cleavage Followed by Positive Ion Mass Spectrometry

The first method involves enzymatic cleavage of the peptide within the intramolecular disulfide bridge to form an intermolecular disulfide peptide.[97,117,118] In the case of RLCIPYIIPC-OH, α-chymotrypsin would be an ideal endoprotease to use. (α-chymotrypsin cleaves amide bonds C-terminal to amino acids with aromatic side chains such as Tyr).[119,120] By forming an intermolecular disulfide linked peptide, only one amide bond cleavage is required to form a fragment ion. (Two cleavages are required to produce fragment ions of residues within an intramolecular disulfide bridged peptide).

One downside to this approach is that an inter-disulfide linked peptide contains two N-terminal ends and two C-terminal ends. This means that the MS/MS spectra of such peptides will contain two sets of overlapping B fragment ions and two sets of overlayed Y + 2 ions. This increases the complexity of sequence determination of underivatised peptides.

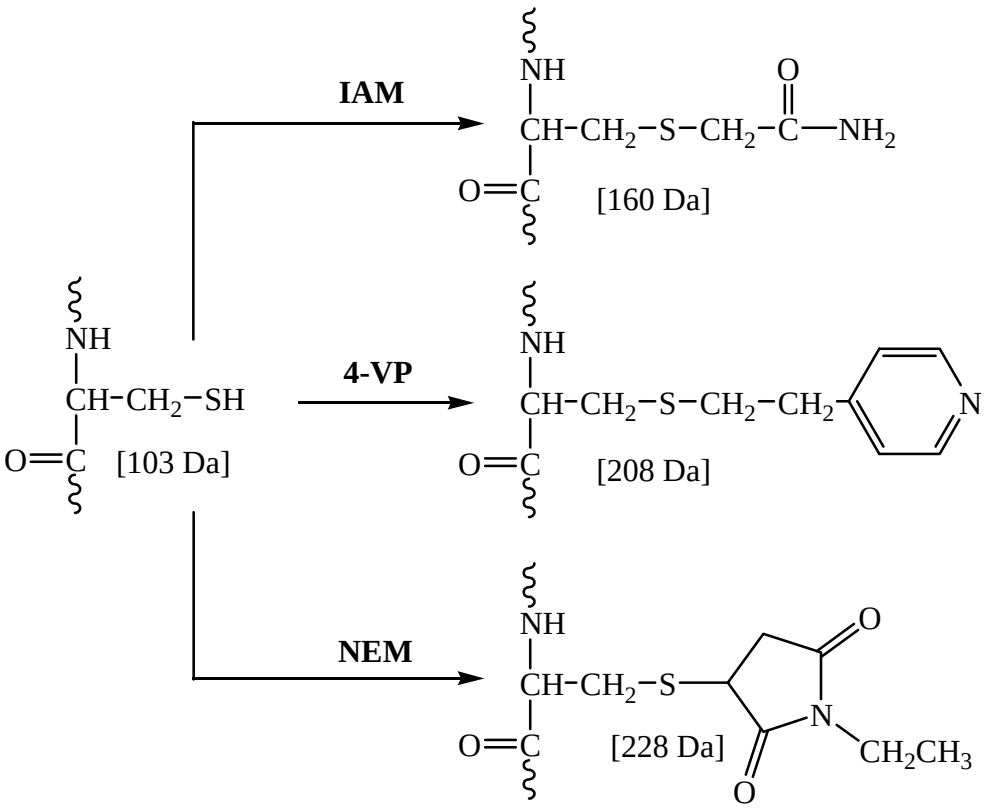
In the example shown in Scheme 4.1, α-chymotrypsin can be used to cleave at the C-terminal amide of Tyr, an aromatic residue.[119,120] When no aromatic containing amino acids are present within the disulfide link, alternative enzymes such as Lys-C or trypsin etc. have to be utilised. The main drawback of sequencing intra-disulfide linked peptides by first converting to their inter-counterparts, is that not every intra-disulfide peptide will necessarily be cleaved within the disulfide link by known endoproteases.

4.1.2 Intramolecular Disulfide Bridged Peptide Sequencing by Disulfide Reduction/Alkylation Followed by Positive Ion Mass Spectrometry

The second method of Scheme 4.1 is an alternative approach for sequencing intramolecular disulfide bridged peptides. It involves the reduction of the disulfide to form two individual Cys residues. The most common reducing agents used are dithiothreitol (DTT)[121-124] and tris(2-carboxyethyl)phosphine (TCEP).[123-125] To achieve complete reduction of the disulfide it is often required to optimise the reduction conditions for each peptide.[111]

Alkylation of reduced peptides is essential to prevent the reformation of the disulfide link during isolation and analysis of the peptide.[126] Examples of common

alkylating agents include iodoacetamide (IAM),[127-130] iodoacetic acid (IAA),[130-132] 4-vinylpyridine (4-VP),[133-135] fluorescein-5-maleimide (FM)[136] and *N*-ethylmaleimide (NEM).[136-138] Alternatively, the Cys residue can be protected by cyanylation with 1-cyano-4-(dimethylamino)pyridinium (CDAP) tetrafluoroborate.[138,139] The Cys alkylated analogues and corresponding masses of IAM, 4-VP and NEM are illustrated in Scheme 4.2.



Scheme 4.2. The IAM, 4-VP and NEM alkylated analogues of Cys, and their corresponding masses.[140,141]

The use of some alkylating agents can cause additional complications for sequencing due to alkylation of residues other than Cys. IAM and IAA have been reported to occasionally alkylate His, Lys and Met.[142-144] 4-VP is known to alkylate amines, particularly Lys and the N-terminus of peptides.[144]

There are several less common methods of cleaving the S-S bond of disulfide peptides. One approach involves the scission of the disulfide bond and oxidation of the free Cys residues with performic acid.[145,146] A similar result can also be achieved using sodium sulfite.[147] The use of synchrotron X-rays to cleave and oxidise the disulfide to two sulfonic acid analogues has been reported recently.[148] However, whichever method is utilised, the assignment and sequencing of disulfide linked peptides remains a time consuming process.[149]

4.2 Results and Discussion

Two types of intramolecular disulfide bridged peptides are considered for negative ion mass spectrometric analysis. The first type is where the disulfide occupies the central portion of the peptide (i.e. there are one or more amino acid residues on either side of the disulfide bridge). The second example is where one of the Cys residues involved in the disulfide link is in the C-terminal position.

4.2.1 Loss of H_2S_2 from Intramolecular Disulfide Bridged Peptides

The first disulfide linked peptide to be considered is RLCIPYIIPCG-OH **(9)**, a synthetic modification of signiferin 1 **(8)** where a Gly residue has been placed in the C-terminal position. The positive ion MS/MS spectrum of **9** is shown in Figure 4.3. The Y + 2 fragment ions provide the sequence of the first two residues (the two residues N-terminal of the disulfide link). The B fragmentations shown schematically above the spectrum indicate the presence of the C-terminal Gly residue but then continue to provide the sequence of several residues within the disulfide bridge. It appears that the amide bond after Pro9 is cleaved with the C-terminal portion of the peptide remaining bound to the N-terminal through the disulfide link. This is followed by a second amide cleavage after Ile7 and Tyr6 to give the sequence as shown on the spectrum.

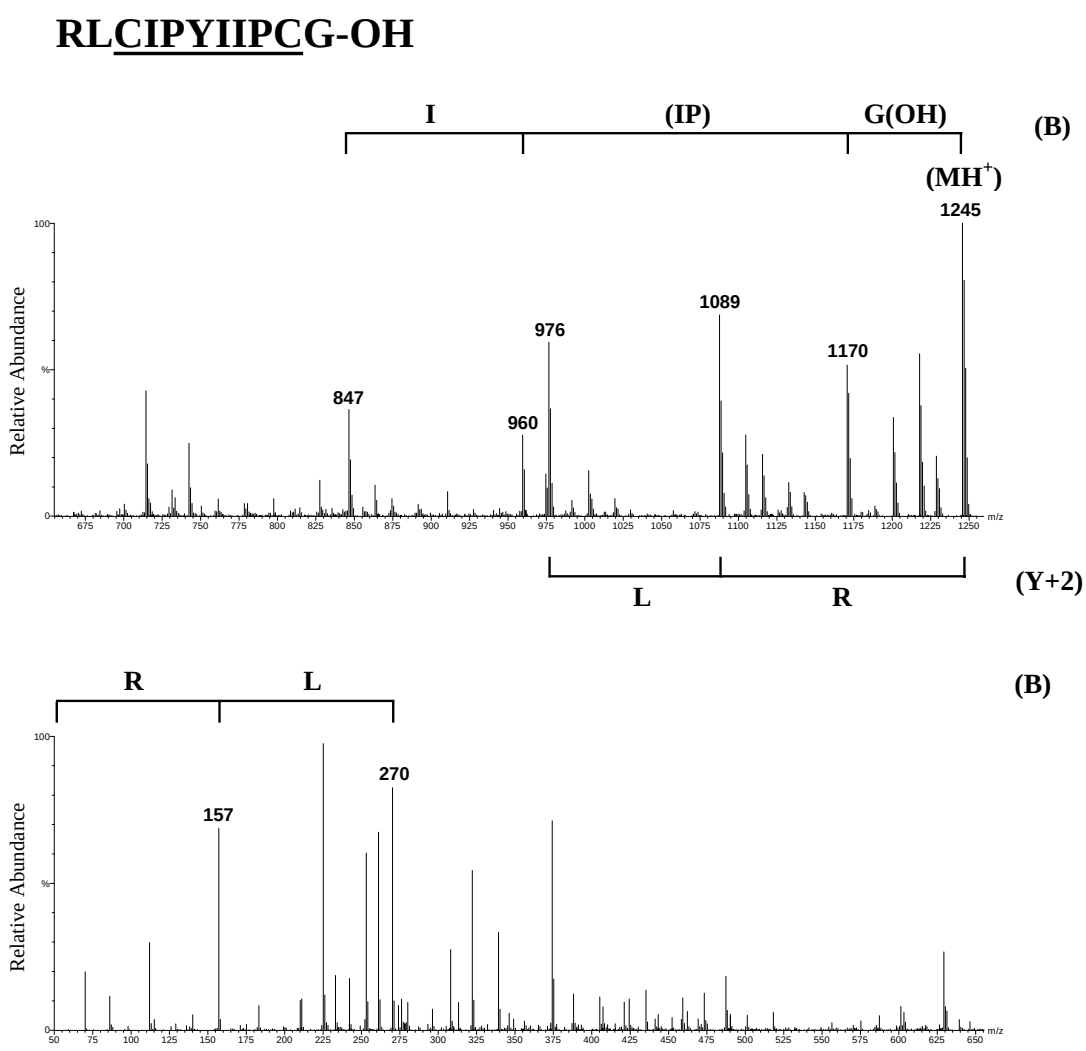


Figure 4.3. Collision-induced MS/MS data of **9**. Positive ion fragmentations of the $(MH)^+$ species, B cleavage ions are shown sequentially above the spectrum and Y + 2 cleavage ions are shown below the spectrum.

The negative ion mass spectrum of **9** is shown in Figure 4.4. The base peak in the spectrum is *m/z* 1177 and corresponds to the loss of H_2S_2 from the parent ion.

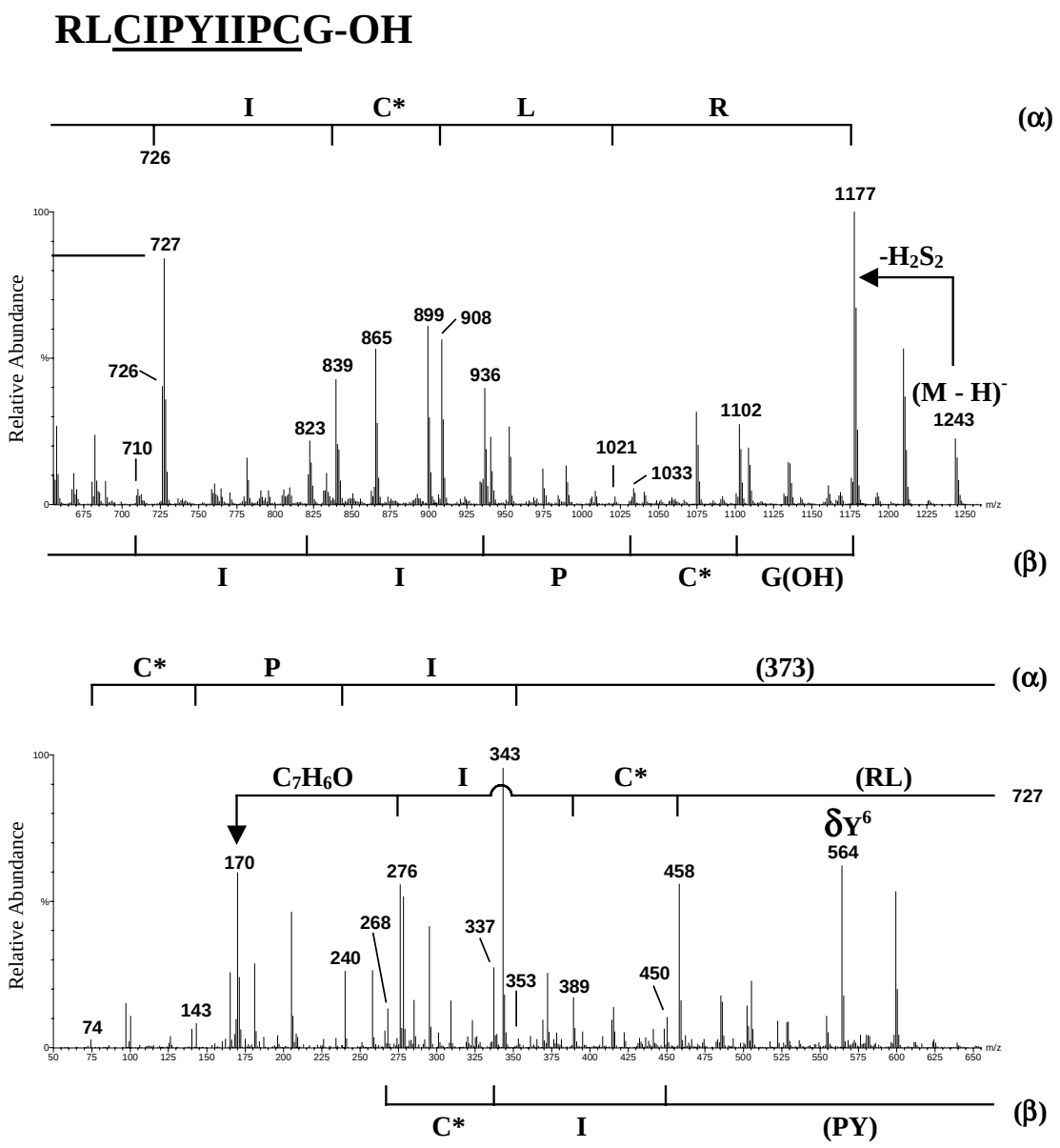
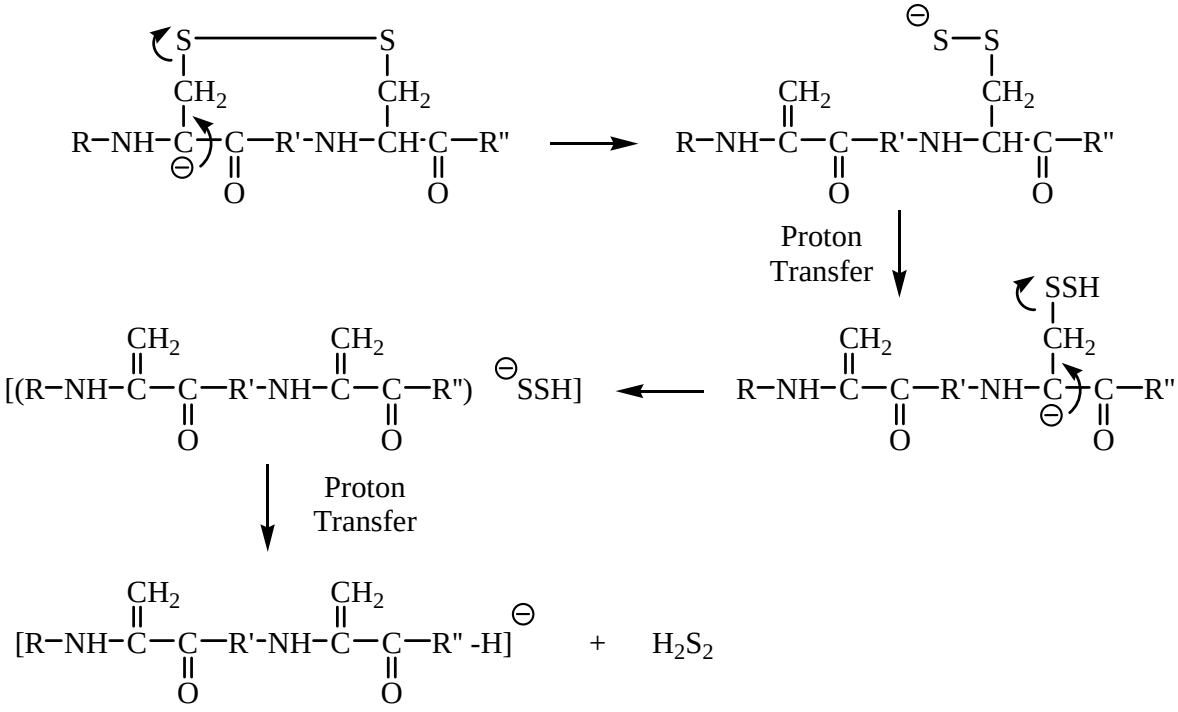


Figure 4.4. Collision-induced MS/MS data for the $(M-H)^-$ parent anion of **9**. α Cleavages are shown above the spectrum and β cleavages below the spectrum. C* = $NHC(=CH_2)CO$. Magnification ranges: *m/z* 50-600 (x4), *m/z* 650-850 (x2), *m/z* 900-1050 (x4).

The loss of H_2S_2 must involve cleavage of the disulfide bridge. A proposed charge induced mechanism for this process is shown in Scheme 4.3. The mechanism involves the enolate anion on the N-terminal Cys residue inducing the cleavage of the carbon-sulfur bond. This is followed by proton abstraction at the α-carbon of the C-terminal Cys to form

a second enolate. The cleavage of the second C-S bond forms HSS^- which deprotonates the peptides to form H_2S_2. The two Cys residues acquire the structure $NHC(=CH_2)CO$ which is abbreviated to C*.



Scheme 4.3. A charge-induced mechanism for the loss of H_2S_2 from an intramolecular disulfide bridged peptide.

Re-examination of the spectrum in Figure 4.4 reveals that the α and β backbone fragmentation ions do not originate from the parent ion, but rather the $[(M-H)^- -H_2S_2]^-$ fragment ion. This implies that the loss of H_2S_2 is a facile process. It also means that the disulfide is cleaved first, exposing the residues originally within the disulfide link to amide backbone cleavages. This allows for these residues (originally within the disulfide link) to be sequenced. Conveniently, the mass of C* (69 Da) differs from the 20 common amino acids allowing the position of this residue to be confirmed within the peptide by examination of the α and β fragmentations.

Theoretical calculations of the proposed mechanism were examined on the model system HCO-<u>CGC</u>-O^- where the N-terminal portion of the peptide was replaced the HCO and the C-terminal with O^-. The amino acids within the disulfide were replaced with one Gly residue to simplify the calculations. The reaction coordinate profile for the charge-induced loss of H_2S_2 is displayed in Figure 4.5.

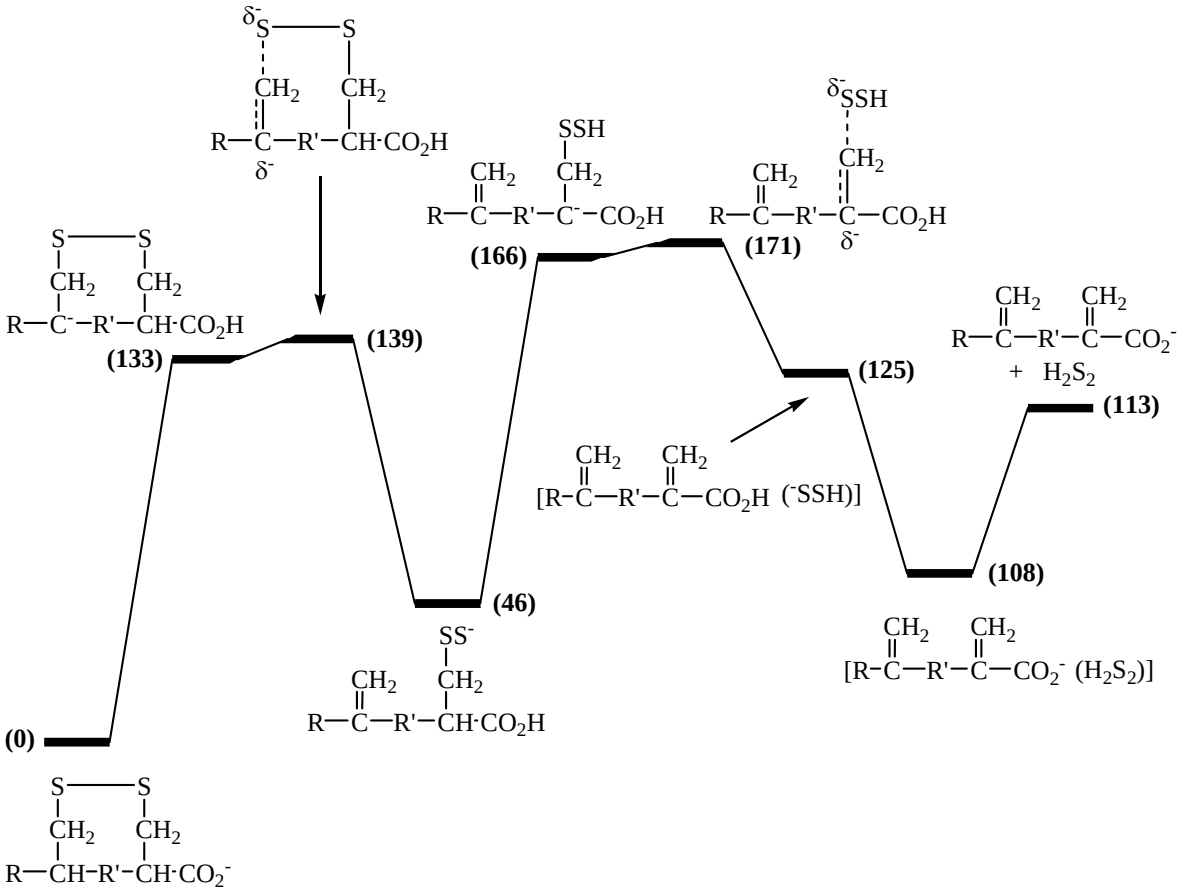


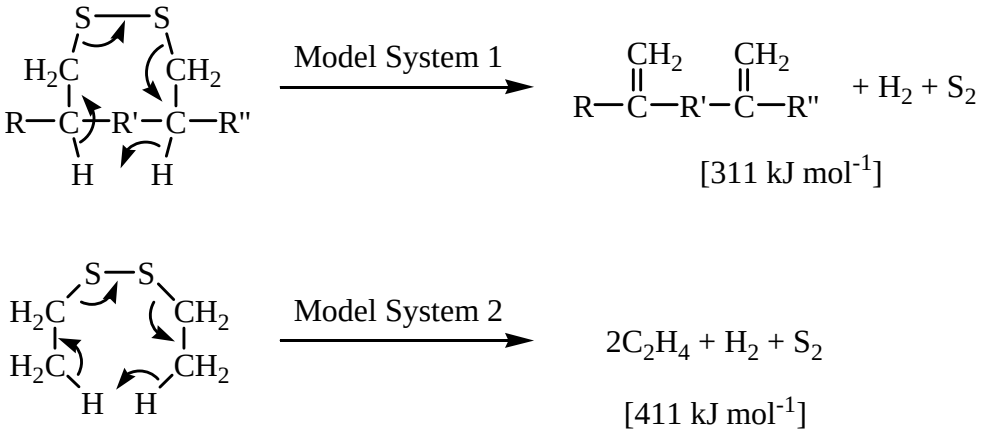
Figure 4.5. Loss of H_2S_2 from an intramolecular disulfide peptide on the model system HCO-<u>CGC</u>-O^-. R = HCONH, R' = $CONHCH_2CONH$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory. Transition states were confirmed by IRC calculations.

The enolate anion is one of many species which can be directly formed in the source of the spectrometer. Alternatively, proton transfer to the carboxylate anion to form the enolate is endothermic by 133 kJ mol^{-1} at the HF/6-31G(d)//AM1 level of theory. Once the enolate anion is formed, only 6 kJ mol^{-1} is required to cleave the carbon-sulfur bond and open the peptide from its cyclic form. This is followed by a second proton transfer to produce the second enolate. Once formed, only 5 kJ mol^{-1} is require to cleave the second C-S bond with the overall process endothermic by 113 kJ mol^{-1} (or exothermic by 20 kJ mol^{-1} from the first enolate ion). The highest transition state barrier for the loss of H_2S_2 is 171 kJ mol^{-1} (or 38 kJ mol^{-1} from the enolate anion), substantially lower than the α and β fragmentation processes (Figure 2.8). This explains why the α and β fragmentation ions originate from the $[(M-H)^- -H_2S_2]^-$ fragment ion rather than the parent ion.

The positive ion MS/MS spectrum of **9** (Figure 4.3) did not provide the full sequence of this peptide. Under negative ion conditions, the presence of a disulfide bridge in an underivatised peptide is confirmed by the major loss of H_2S_2 from the parent ion. This opens the peptide from the cyclic form and allows amide backbone cleavages to sequence the peptide from one end to the other. In the case of **9** (Figure 4.4), the α and β fragment ions provide the entire sequence of the peptide except the relative positions of Pro5 and Tyr6 and differentiation between isomeric Leu and Ile. However, formation of the Tyr δ ion at *m/z* 564 from the $[(M-H)^- -H_2S_2]^-$ ion confirms the position of this residue, and therefore the position of the Pro residue.

A charge-remote mechanism for the loss of the elements of H_2S_2 (Scheme 4.4) was also considered. Theoretical calculations were conducted using the same model system

and level of theory as the one used for the charge-induced mechanism. A simpler model system ($CH_3CH_2SSCH_2CH_3$) was also considered.

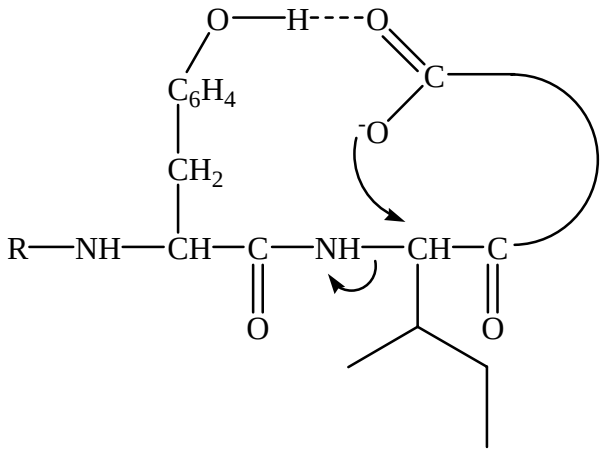


Scheme 4.4. The charge-remote energetics for loss of the elements of H_2S_2 from two disulfide model systems. Model System 1: R = HCONH, R' = $CONHCH_2CONH$, R'' = CO_2H. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory.

The charge-remote process is strongly endothermic for both model systems when compared to the charge-induced mechanism. No transition states were found for either model system. It is unlikely therefore that a disulfide bridged peptide will lose the elements of H_2S_2 under negative ion conditions via a charge-remote process.

The negative ion mass spectrum of **9** (Figure 4.4) produces a pronounce peak at *m/z* 727. "MS/MS/MS" experiments reveal that this ion originates from the $[(M-H)^- - H_2S_2]^-$ fragment ion and corresponds to the structure RLC*IPY-NH^- [where C* has the structure $NHC(=CH_2)CO$]. The *m/z* 727 ion undergoes further amide backbone cleavages to produce α fragmentation ions. The loss of C_7H_6O is the characteristic side chain loss from a Tyr residue[54] (see Table 2.1).

The formation of the ion at *m/z* 727 may occur via a process similar to the side chain induced backbone cleavages from the residues Ser, Thr and Cys. Following the loss of H_2S_2 from the parent ion, the C-terminal carboxylate group hydrogen bonds with the OH side chain of Tyr (Scheme 4.5). This places the carboxylate anion in a position where it can induce an S_N2 attack on the electrophilic α-carbon of Ile. The NH-CH bond of Ile is cleaved forming the fragment ion RLC*IPY-NH^- at *m/z* 727.



Scheme 4.5. The proposed mechanism for the formation of the *m/z* 727 fragment ion in Figure 4.4.

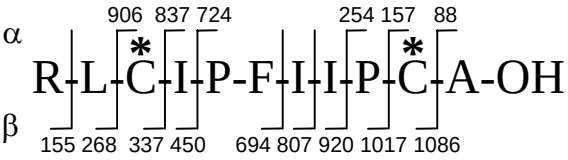
Mass spectrometric analysis of a second signiferin 1 analogue was also investigated. The modification includes the replacement of Tyr6 with Phe6 and the insertion of Ala in the C-terminal position. The MS/MS spectral data of this peptide are summarised in Table 4.1. The loss of H_2S_2 from the parent ion forms the base peak in the spectrum. α and β fragment ions provide the sequence of the peptide, except for differentiation between isomeric Leu and Ile and the relative positions of Pro5 and Phe6. However, the position of these two residues is confirmed by the formation of the Phe δ ion at *m/z* 564.

Table 4.1. Negative ion fragmentation data for the signiferin 1 synthetic analogue RLCIPFIIPCA-OH. α and β fragmentations from the $[(M-H)^- -H_2S_2]^-$ ion are summarised in schematic form.

RLCIPFIIPCA-OH

$(M-H)^-$ *m/z* 1241.

$[(M-H)^- -H_2S_2]^-$ *m/z* 1175.

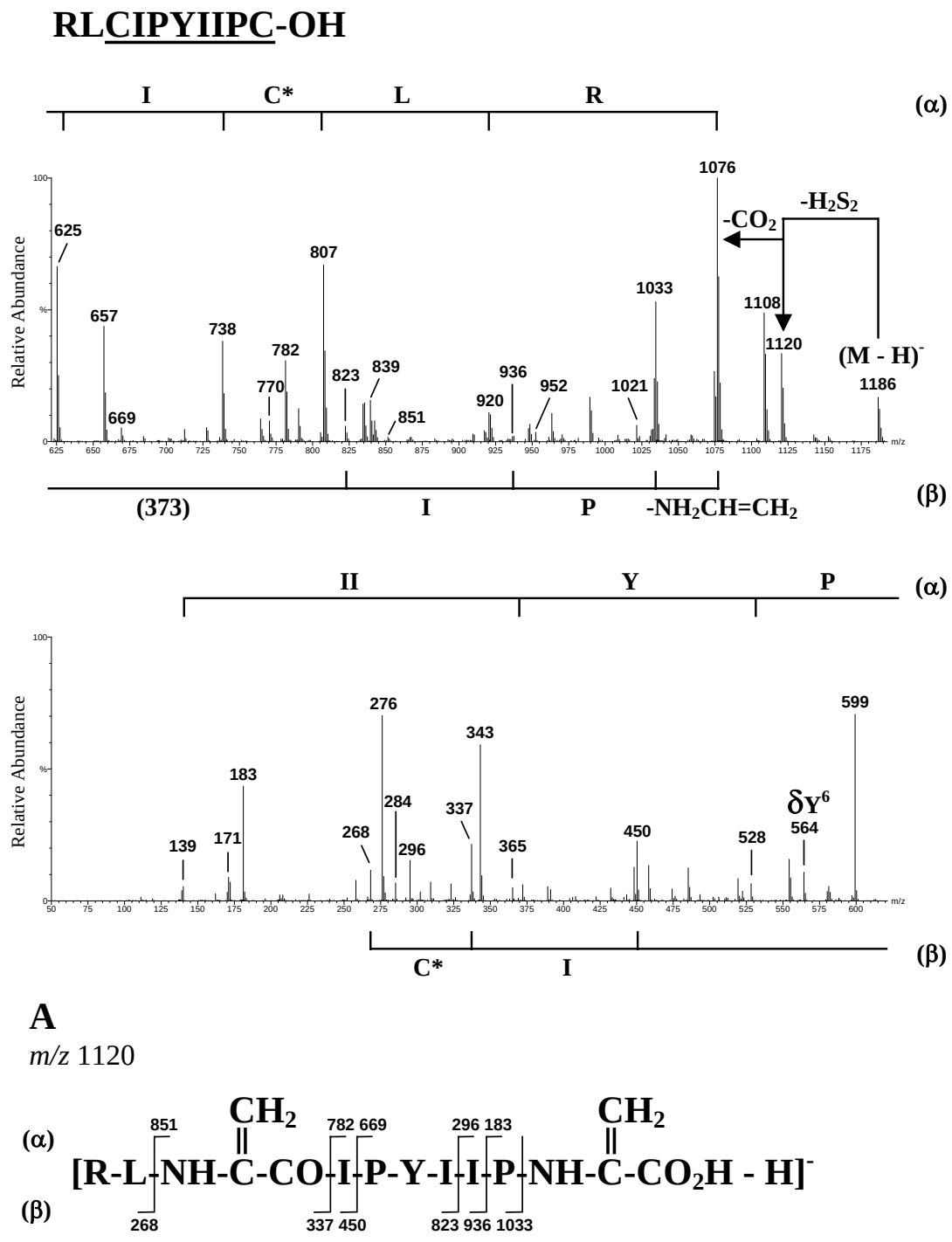


δF^6 from $[(M-H)^- -H_2S_2]^-$ *m/z* 564.

4.2.2 Loss of H_2S_2 and CO_2 from C-Terminal Intramolecular Disulfide Bridged Peptides

The positive ion mass spectrum of RLCIPYIIPC-OH **(8)**, a peptide containing an intramolecular disulfide bridge where one of the Cys residues involved in the bridge is in the C-terminal position, is shown in Figure 4.2. It has already been discussed how positive ion mass spectrometry is often unable to provide the full sequence of an intramolecular disulfide bridged peptide, as two amide bond cleavages are required to assign residues within the disulfide link. Negative ion mass spectrometry is better at identifying and sequencing intramolecular disulfide bridged peptides. The loss of H_2S_2 and formation of the $[(M-H)^- -H_2S_2]^-$ ion as the base peak in the spectrum immediately identifies the presence of a disulfide bridged peptide. The subsequent α and β cleavages from the $[(M-H)^- -H_2S_2]^-$ ion provide the sequence of the peptide, including the residues originally within the disulfide link.

The negative ion MS/MS data for **8**, an intramolecular disulfide bridged peptide where one of the Cys residues is in the C-terminal position are shown in Figure 4.6. The α and β fragment ions are shown above and below the spectrum respectively from the $[(M-H)^- -(H_2S_2+CO_2)]^-$ ion. "MS/MS/MS" data from the $[(M-H)^- -H_2S_2]^-$ ion are shown schematically below the spectrum.

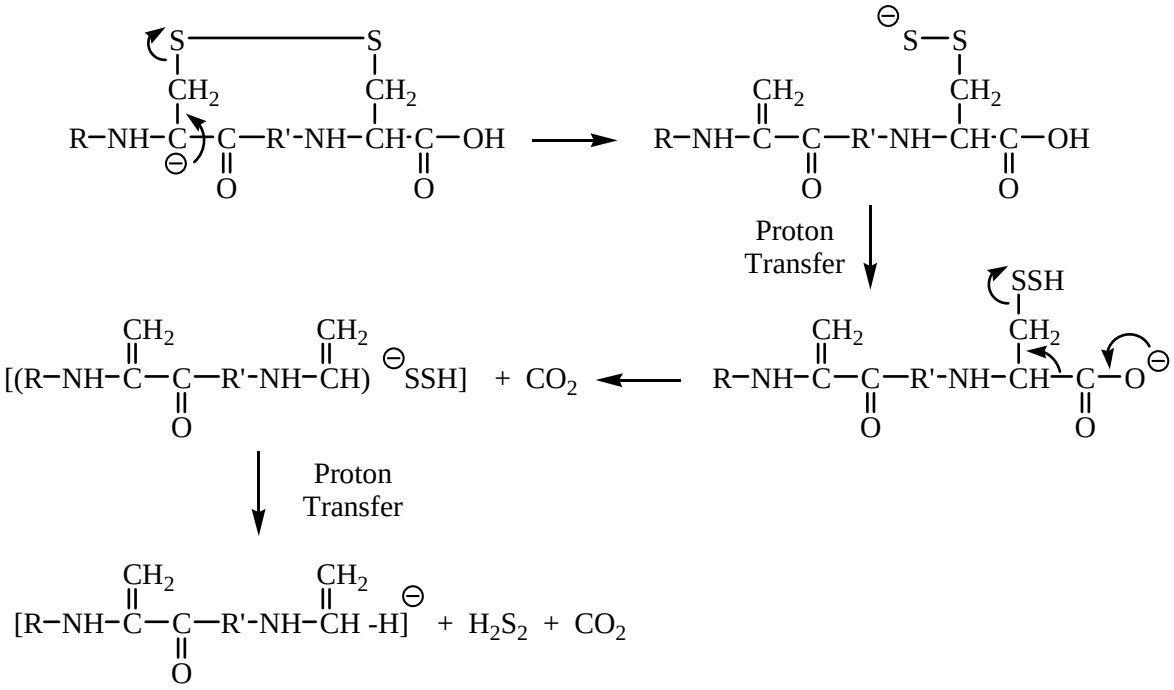


A

m/z 1120

Figure 4.6. Negative ion collision-induced MS/MS data of **8**. α and β fragmentation ions from $[(M-H)^- -(H_2S_2+CO_2)]^-$ are shown above and below the spectrum respectively. "MS/MS/MS" data of the $[(M-H)^- -H_2S_2]^-$ ion are shown schematically below the spectrum. C* = $NHC(=CH_2)CO$. Magnification ranges: *m/z* 100-600 (x6), *m/z* 650-800 (x4), *m/z* 900-1050 (x16), *m/z* 1100-1150 (x2).

The loss of H_2S_2 from $(M-H)^-$ forms the fragment ion at *m/z* 1120. However, this is not the base peak in the spectrum. The base peak corresponds to the loss of both H_2S_2 and CO_2. The α and β backbone cleavages originate from the $[(M-H)^- -(H_2S_2+CO_2)]^-$ ion, indicating that the loss of H_2S_2 and CO_2 from a Cys C-terminal intramolecular disulfide bridged peptide is a facile process. A proposed mechanism for the loss of H_2S_2 and CO_2 is shown in Scheme 4.6.



Scheme 4.6. A mechanism for loss of H_2S_2 and CO_2 from a Cys C-terminal intramolecular disulfide bridged peptide.

Theoretical calculations of the proposed mechanism were performed at the HF/6-31G(d)//AM1 level of theory on the model system HCO-CGC-O^-. The N-terminal portion of the peptide has been replaced with HCO and the residues within the disulfide link with a Gly residue. This is the same model system and level of theory used for the loss of H_2S_2.

A summary of the reaction coordinate profile for the losses of H_2S_2 and CO_2 is shown in Figure 4.7.

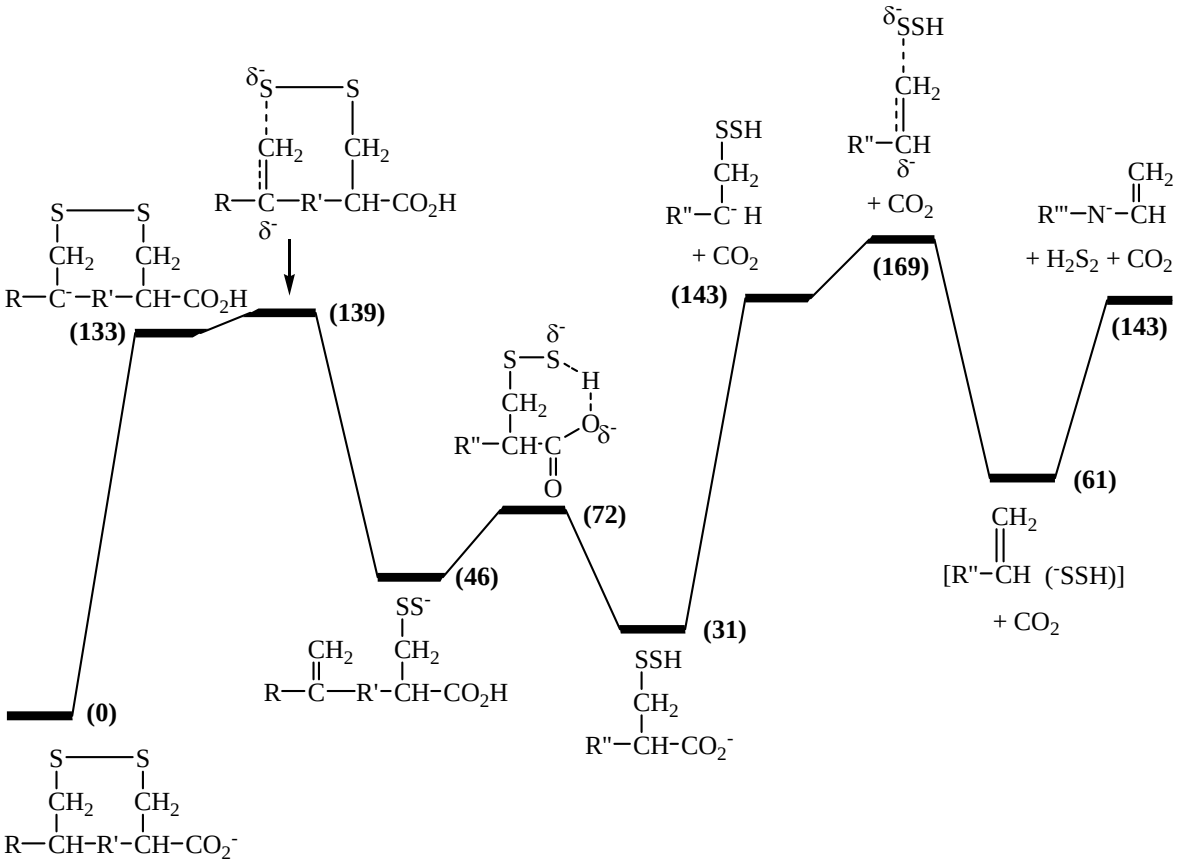


Figure 4.7. Losses of H_2S_2 and CO_2 from a Cys C-terminal intramolecular disulfide peptide on the model system HCO-CGC-O^-. R = HCONH, R' = $CONHCH_2CONH$, R'' = $HCONHC(=CH_2)CONHCH_2CONH$, R''' = $HCONHC(=CH_2)CONHCH_2CO$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory. Transition states confirmed by IRC calculations.

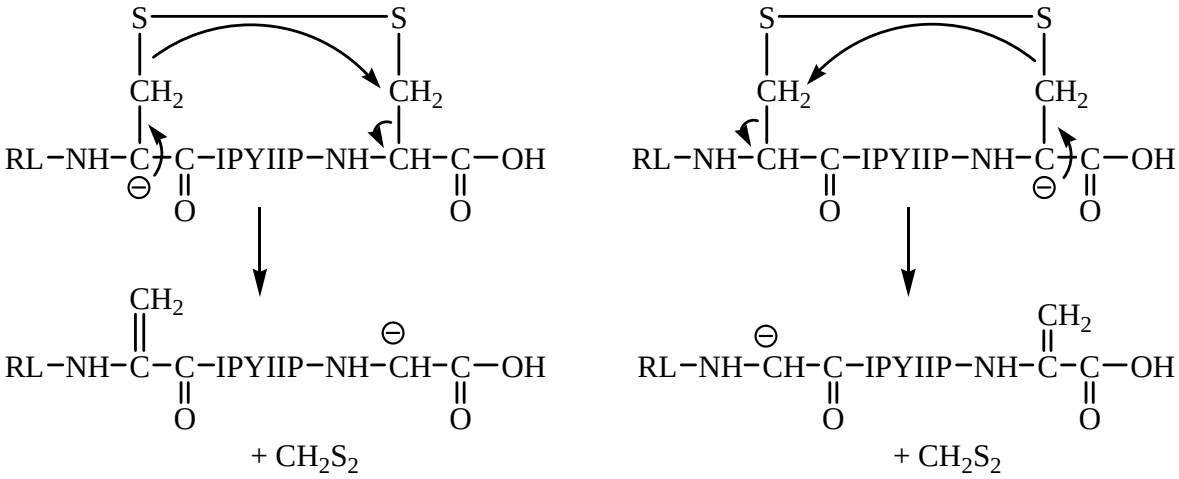
The initial steps in the combined loss of H_2S_2 and CO_2 are identical to the initial steps in the loss of H_2S_2 alone (Figure 4.5). These involve the cleavage of the N-terminal carbon-sulfur bond to open up the disulfide bridge. As mentioned previously, the enolate anion is one of many species which can be formed by direct deprotonation of the neutral peptide. Alternatively, the enolate anion can be formed following proton transfer to the

initially formed carboxylate anion. Only 6 kJ mol^{-1} of energy is required to open the disulfide link from the enolate anion. From this point on, the mechanisms for the combined loss of H_2S_2 and CO_2 differ from that of the loss of H_2S_2 alone.

For the loss of H_2S_2 and CO_2, the next step involves proton transfer to form a carboxylate anion (as opposed to the formation of a second enolate anion in the case of H_2S_2 loss). Once the carboxylate anion is formed, 112 kJ mol^{-1} of energy is required for loss of CO_2. The cleavage of the second C-S bond requires only 26 kJ mol^{-1} of energy from the enolate anion at the HF/6-31G(d)//AM1 level of theory. The overall process for the charge-induced loss of H_2S_2 and CO_2 is endothermic by 143 kJ mol^{-1} (or 10 kJ mol^{-1} from the first enolate anion) with the highest transition state barrier being 169 kJ mol^{-1} (or only 36 kJ mol^{-1} from the enolate anion). The loss of H_2S_2 and CO_2 is a facile process, thus explaining why α and β backbone cleavages originate from the $[(M-H)^- -(H_2S_2+CO_2)]^-$ fragment ion rather than from the parent $(M-H)^-$ ion.

Disulfide peptides which contain a C-terminal Cys residue produce both $[(M-H)^- -(H_2S_2+CO_2)]^-$ and $[(M-H)^- -H_2S_2]^-$ fragment ions. This experimental observation is consistent with the theoretical data in displaying similar energetics for the two processes. The "MS/MS/MS" cleavages from the $[(M-H)^- -H_2S_2]^-$ ion of **8** are summarised below the spectrum of Figure 4.6. There are several α and β backbone cleavages originating from the $[(M-H)^- -H_2S_2]^-$ ion but there is no accompanying loss of CO_2 to form an $[(M-H)^- -(H_2S_2+CO_2)]^-$ ion. This indicates that the loss of H_2S_2 along with CO_2 is an independent process from that which induces the loss of H_2S_2 alone.

The negative ion mass spectrum of RLCIPYIIPC-OH **(8)** shown in Figure 4.6 contains a pronounced peak at *m/z* 1108. This peak corresponds to the fragment ion $[(M-H)^- -CH_2S_2]^-$. Scheme 4.7 illustrates the processes by which two possible fragment anions are formed. The first structure originates from an N-terminal Cys enolate and the second from the C-terminal Cys enolate.



Scheme 4.7. Mechanisms for the formation of two $[(M-H)^- -CH_2S_2]^-$ fragment anions.

"MS/MS/MS" data from the $[(M-H)^- -CH_2S_2]^-$ ion are summarised in Figure 4.8. There are α and β backbone cleavages which originate from both possible structures of the $[(M-H)^- -CH_2S_2]^-$ ion. This confirms that both possible losses of CH_2S_2 are observed experimentally.

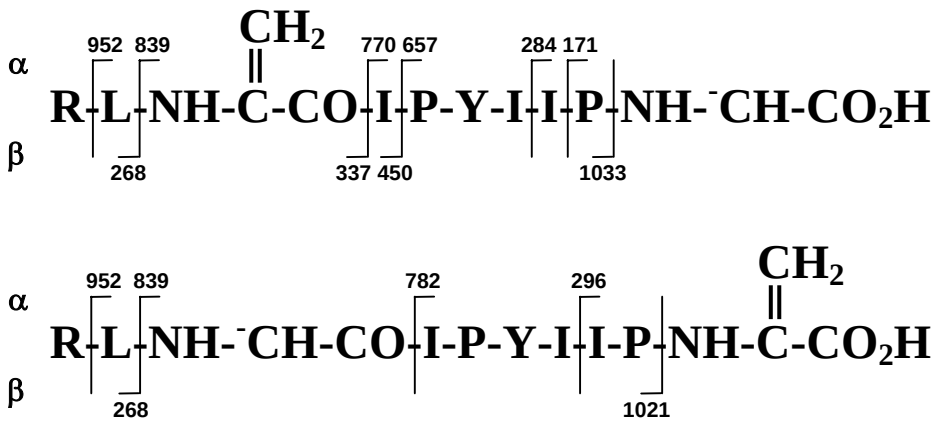


Figure 4.8. A summary of the "MS/MS/MS" data for the $[(M-H)^- -CH_2S_2]^-$ fragment anion.

The mechanism for loss of CH_2S_2 from the parent anion of an intramolecular disulfide bridged peptide was investigated using theoretical calculations. The same model system and level of theory was used for the loss of CH_2S_2 as for the formation of both the $[(M-H)^- -(H_2S_2+CO_2)]^-$ and $[(M-H)^- -H_2S_2]^-$ ions, so that the three processes could be compared directly. The reaction coordinate profile for the loss of CH_2S_2 from the C-terminal Cys enolate is shown in Figure 4.9.

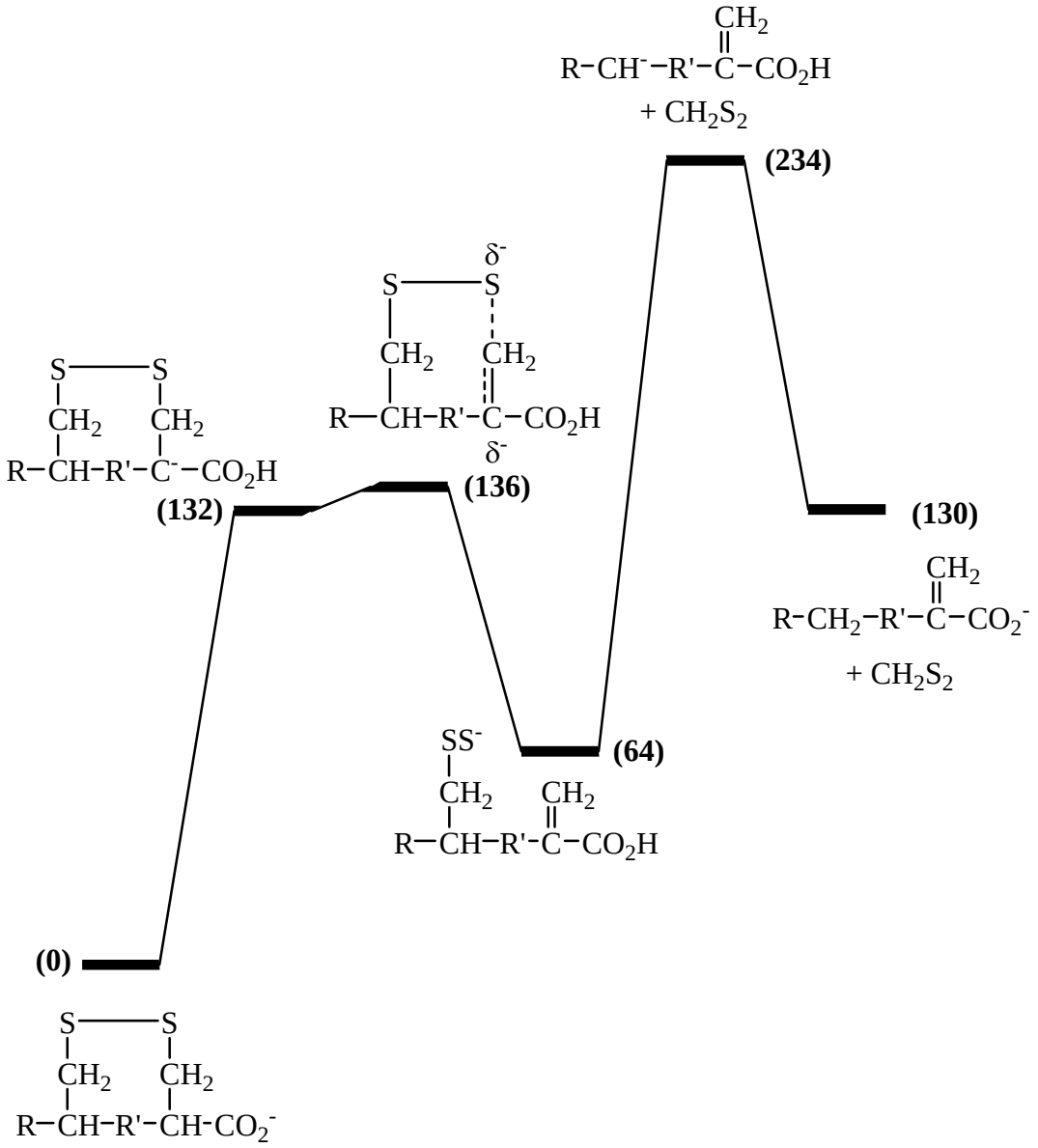


Figure 4.9. Loss of CH_2S_2 from the C-terminal Cys enolate of an intramolecular disulfide peptide on the model system HCO-CGC-O^-. R = HCONH, R' = $CONHCH_2CONH$. Relative energies in kJ mol^{-1}. Geometries at the AM1 level of theory. Energies at the HF/6-31G(d)//AM1 level of theory. Transition states were confirmed by IRC calculations.

The formation of the C-terminal Cys enolate from the carboxylate anion requires 132 kJ mol^{-1} of energy. Once formed, a further 4 kJ mol^{-1} is required to cleave the carbon-sulfur bond. No transition state was found for cleavage of the carbon-carbon bond, rather the CH_2S_2 molecule dissociates from the peptide to form an enolate ion. Proton transfer to

form the more stable carboxylate anion results in the overall process being endothermic by 130 kJ mol^{-1}. The highest transition state barrier from the enolate anion is 102 kJ mol^{-1}, with the overall process being slightly exothermic (-2 kJ mol^{-1}). Although the loss of CH_2S_2 cannot compete energetically with the losses of H_2S_2 and (H_2S_2+CO_2), it is still energetically favourable in comparison with the α and β backbone cleavage processes (Figure 2.8).

The negative ion MS/MS fragmentation data of several intramolecular disulfide linked peptides, where one Cys residue is in the C-terminal position, are summarised in Table 4.2. The α and β backbone cleavages are displayed in schematic form from either the $[(M-H)^-$ -(H_2S_2+CO_2)$]^-$ ion, or the $[(M-H)^-$ -(H_2S_2+CO_2+MeCHO)$]^-$ ion for peptides containing a Thr residue. [The side chain loss of MeCHO from Thr is a facile process (see Table 2.1) and effects the conversion of Thr to Gly].

Table 4.2. A summary of negative ion mass spectral data for C-terminal Cys intra-molecular disulfide bridged peptides.

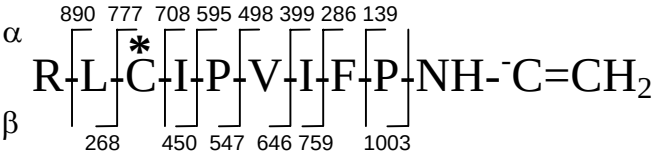
RL<u>CIPVIFPC</u>-OH

$(M-H)^-$ *m/z* 1156.

$[(M-H)^-$ -$H_2S_2]^-$ *m/z* 1090.

$[(M-H)^-$ -$CH_2S_2]^-$ *m/z* 1078.

$[(M-H)^-$ -(H_2S_2+CO_2)$]^-$ *m/z* 1046.



γF^8 from $[(M-H)^-$ -(H_2S_2+CO_2)$]^-$ *m/z* 269.

Table 4.2 (Continued).

FLPP<u>CAYKGTC</u>-OH

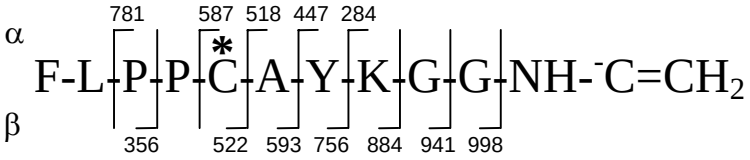
$(M-H)^-$ *m/z* 1195.

$[(M-H)^- -H_2S_2]^-$ *m/z* 1129.

$[(M-H)^- -CH_2S_2]^-$ *m/z* 1117.

$[(M-H)^- -(H_2S_2+CO_2)]^-$ *m/z* 1085.

$[(M-H)^- -(H_2S_2+CO_2+MeCHO)]^-$ *m/z* 1041.



δY^7 from $[(M-H)^- -(H_2S_2+CO_2+MeCHO)]^-$ *m/z* 610, δT^{10} from $[(M-H)^- -(H_2S_2+CO_2)]^-$ *m/z* 958.

FFLPP<u>CAHKGTC</u>-OH

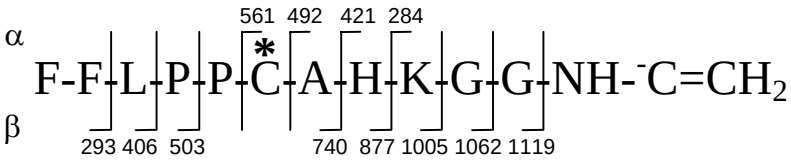
$(M-H)^-$ *m/z* 1316.

$[(M-H)^- -H_2S_2]^-$ *m/z* 1250.

$[(M-H)^- -CH_2S_2]^-$ *m/z* 1238.

$[(M-H)^- -(H_2S_2+CO_2)]^-$ *m/z* 1206.

$[(M-H)^- -(H_2S_2+CO_2+MeCHO)]^-$ *m/z* 1162.



δT^{11} from $[(M-H)^- -(H_2S_2+CO_2)]^-$ *m/z* 1079.

Table 4.2 (Continued).

FFLPP<u>CAYKGTC</u>-OH

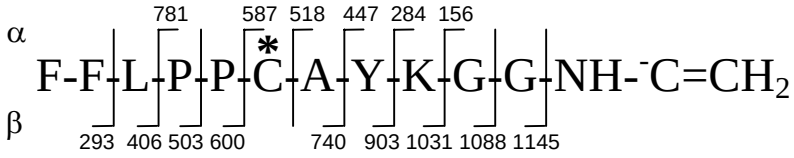
$(M-H)^-$ *m/z* 1342.

$[(M-H)^- -H_2S_2]^-$ *m/z* 1276.

$[(M-H)^- -CH_2S_2]^-$ *m/z* 1264.

$[(M-H)^- -(H_2S_2+CO_2)]^-$ *m/z* 1232.

$[(M-H)^- -(H_2S_2+CO_2+MeCHO)]^-$ *m/z* 1188.



γY^8 from $[(M-H)^- -(H_2S_2+CO_2+MeCHO)]^-$ *m/z* 430, δY^8 from $[(M-H)^- -(H_2S_2+CO_2+MeCHO)]^-$ *m/z* 757, δT^{11} from $[(M-H)^- -(H_2S_2+CO_2)]^-$ *m/z* 1105.

FPLP<u>CAYKGTYC</u>-OH

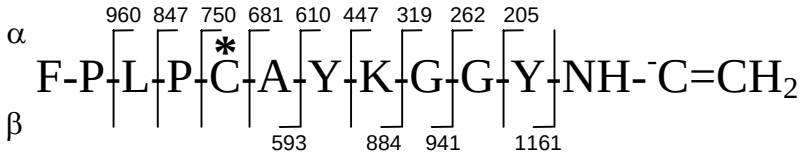
$(M-H)^-$ *m/z* 1358.

$[(M-H)^- -H_2S_2]^-$ *m/z* 1292.

$[(M-H)^- -CH_2S_2]^-$ *m/z* 1280.

$[(M-H)^- -(H_2S_2+CO_2)]^-$ *m/z* 1248.

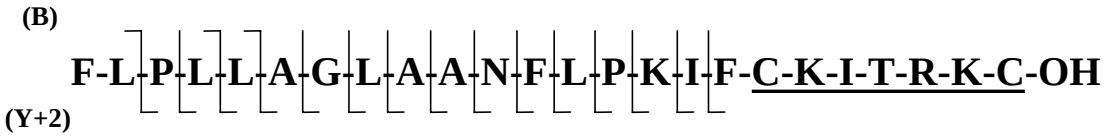
$[(M-H)^- -(H_2S_2+CO_2+MeCHO)]^-$ *m/z* 1204.



δY^7 from $[(M-H)^- -(H_2S_2+CO_2+MeCHO)]^-$ *m/z* 610, δT^{10} from $[(M-H)^- -(H_2S_2+CO_2)]^-$ *m/z* 958.

The final intramolecular disulfide peptide considered for this investigation is brevinin 1E **(10)**. This host defence peptide was isolated from the toad, *Rana ridibunda*, located in the Moscow region of Russia.[150] Many antibacterial disulfide containing peptides have been isolated previously from toads of the genus *Rana*.[133,151-153] Brevinin 1E, FLPLLAGLAANFLPKIF<u>CKITRKC</u>-OH **(10)** is a 24 residue peptide containing a C-

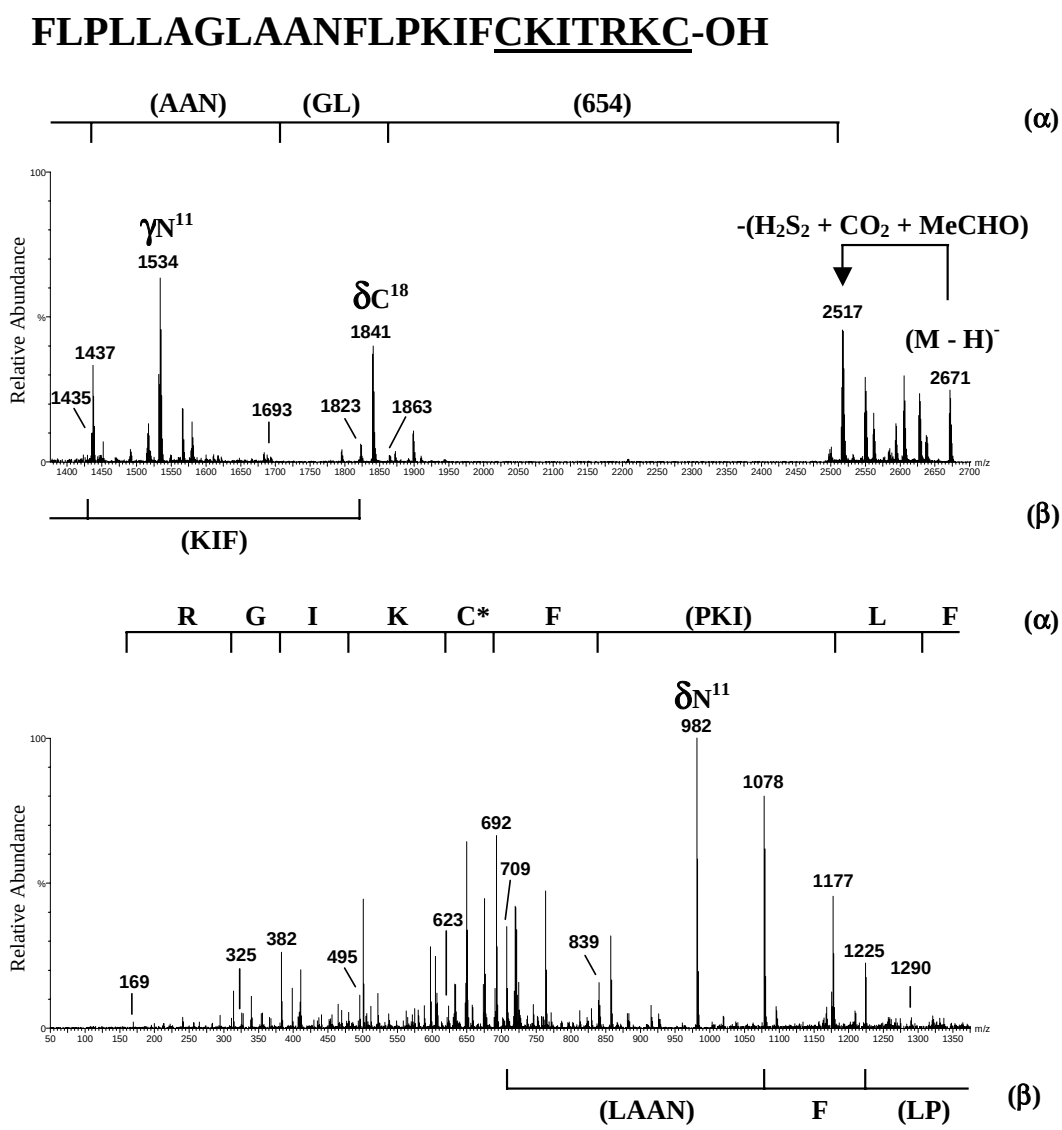
terminal Cys intramolecular disulfide bridge. The positive ion mass spectral data for **10** is summarised in schematic form in Scheme 4.8.



Scheme 4.8. A summary of positive ion MS/MS cleavages of **10**.

The positive ion spectra of $(MH)^+$ ions of the signiferin peptides show B and Y + 2 fragmentations of residues within the disulfide bridge. Such processes are not observed from the $(MH)^+$ ion of brevinin 1E. All the cleavages are from the N-terminal portion of the peptide. This is consistent with previous reports that positive ion cleavages are normally limited to residues outside the disulfide link.[51,109-115] The positive ion data for **10** provides the sequence of all residues outside the disulfide bridge except for Phe17 and the relative position of the first two residues (and differentiation between Leu and Ile, and Lys and Gln).

Negative ion MS/MS data for **10** are shown in Figure 4.10. The major peak at *m/z* 2517 corresponds to the loss of H_2S_2, CO_2 and MeCHO (the side chain loss from Thr). α and β cleavages from the $[(M-H)^- -(H_2S_2+CO_2+MeCHO)]^-$ ion are shown above and below the spectrum respectively and are summarised on the formula below the spectrum. The α and β cleavages along with the positive ion fragmentations described above provide most of the sequence of the peptide.



[(M-H)$^-$ -(H_2S_2 + CO_2 + MeCHO)]$^-$

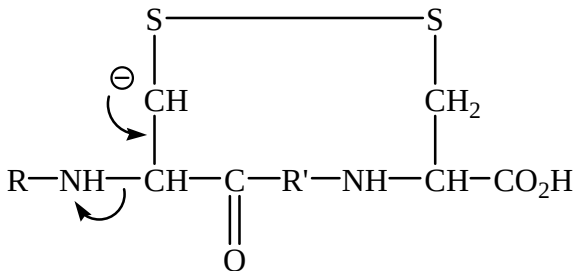
(α)

F-L-P-L-L-A-G-L-A-A-N-F-L-P-K-I-F-C*-K-I-G-R-K-(NHCHCH$_2$)

(β)

Figure 4.10. Collision-induced MS/MS data of **10**. Negative ion fragmentations of the $(M-H)^-$ ion. α and β cleavages are shown above and below the spectrum respectively and summarised on the formula below the spectrum. C* = NHC(=CH_2)CO. Magnification ranges: *m/z* 100-650 (x4), *m/z* 1000-1450 (x10), *m/z* 1450-1890 (x2), *m/z* 2490-2650 (x4).

The fragment ion *m/z* 1841 corresponds to a δ ion produced by the side chain induced backbone cleavage from Cys18. This type of fragmentation was not observed for any of the disulfide peptides considered thus far. The mechanism for the formation of the δ ion at *m/z* 1841 is shown in Scheme 4.9.



Scheme 4.9. Mechanism for formation of the δ ion at *m/z* 1841.

The "MS/MS/MS" data for *m/z* 1841 are shown in Figure 4.11. The α and β cleavages (shown above and below the spectrum respectively) provide additional sequencing information to the MS/MS spectrum (Figure 4.10) and confirm the structure of the ion at *m/z* 1841 as FLPLLAGLAANFLPKIF-NH^{-}.

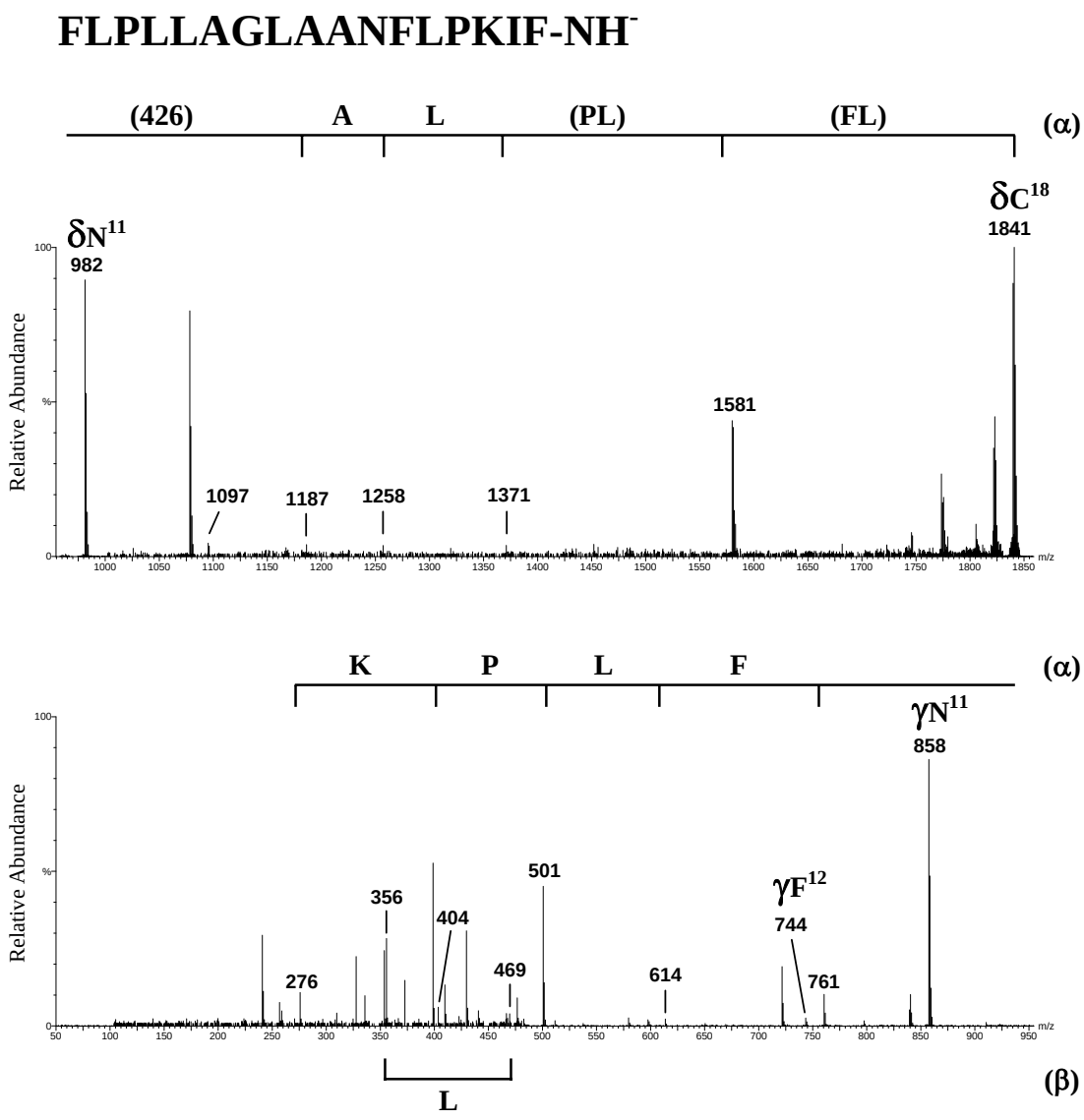
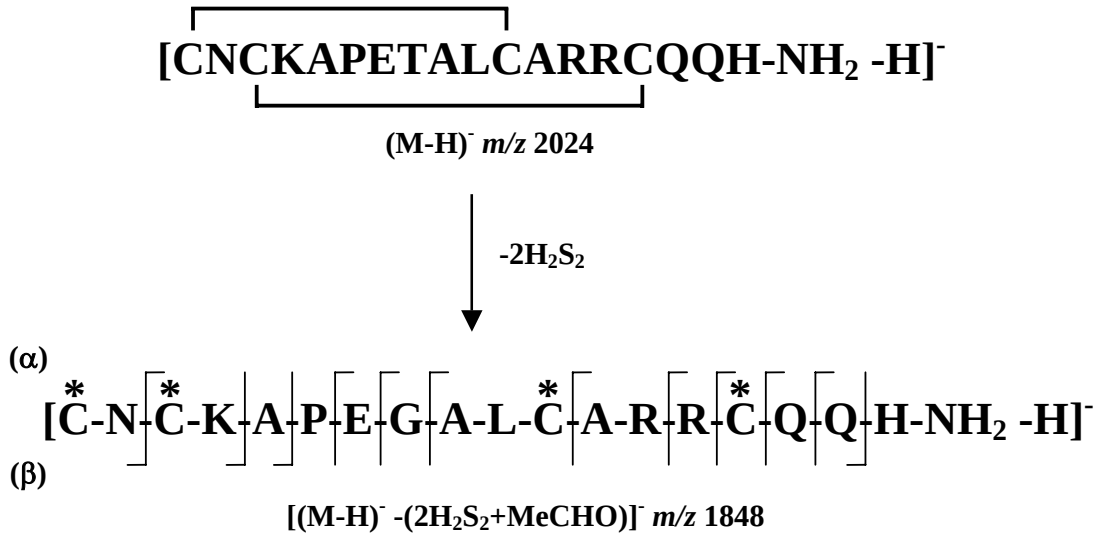


Figure 4.11. Negative ion "MS/MS/MS" of the δCys18 fragment ion from **10** at *m/z* 1841. α and β cleavages are indicated above and below the spectrum respectively. Magnification ranges: *m/z* 100-480 (x24), *m/z* 480-1000 (x2), *m/z* 1000-1100 (x16), *m/z* 1100-1830 (x24).

4.2.3 Multiple Disulfide Bridged Peptides

When a peptide contains an intramolecular disulfide bridge, a major fragmentation from the $(M-H)^-$ ion involves loss of H_2S_2. If a peptide contains two disulfide bridges, are there losses of $2H_2S_2$ from the $(M-H)^-$ parent ion? This was investigated by examining the

negative ion spectrum of the peptide apamin **(11)**. Apamin is an 18 residue peptide containing Cys1-Cys11 and Cys3-Cys15 intramolecular disulfide bridges. The negative ion mass spectral data for **11** are summarised in schematic form in Scheme 4.10.



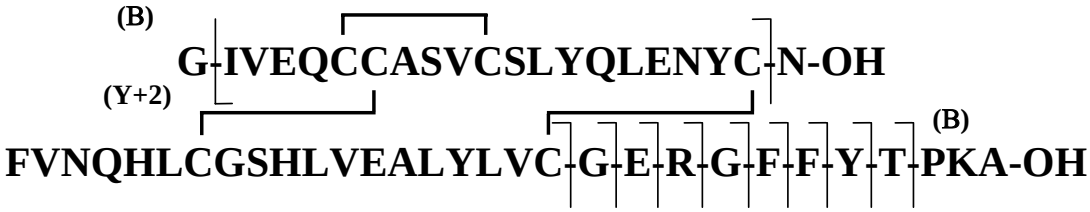
Scheme 4.10. A summary of the negative ion fragmentations from the $(M-H)^-$ parent ion of apamin **(11)**. C* = $NHC(=CH_2)CO$.

There is a pronounced peak formed at *m/z* 1848 which corresponds to the losses of $2H_2S_2$ and MeCHO. (The side chain loss of MeCHO from Thr is a facile process and converts Thr to Gly). The α and β backbone cleavages originate from the $[(M-H)^- -(2H_2S_2+MeCHO)]^-$ fragment ion and are shown schematically in Scheme 4.10. They provide much of the sequence of this peptide. The γGlu7 ion is present at *m/z* 1283 and also originates from the $[(M-H)^- -(2H_2S_2+MeCHO)]^-$ ion.

Bovine insulin is a protein consisting of two peptide chains linked together by a pair of intermolecular disulfide bridges. Chain A consists of 21 residues and chain B, 30 residues. The two intermolecular links are between chain A Cys7-chain B Cys7 and chain

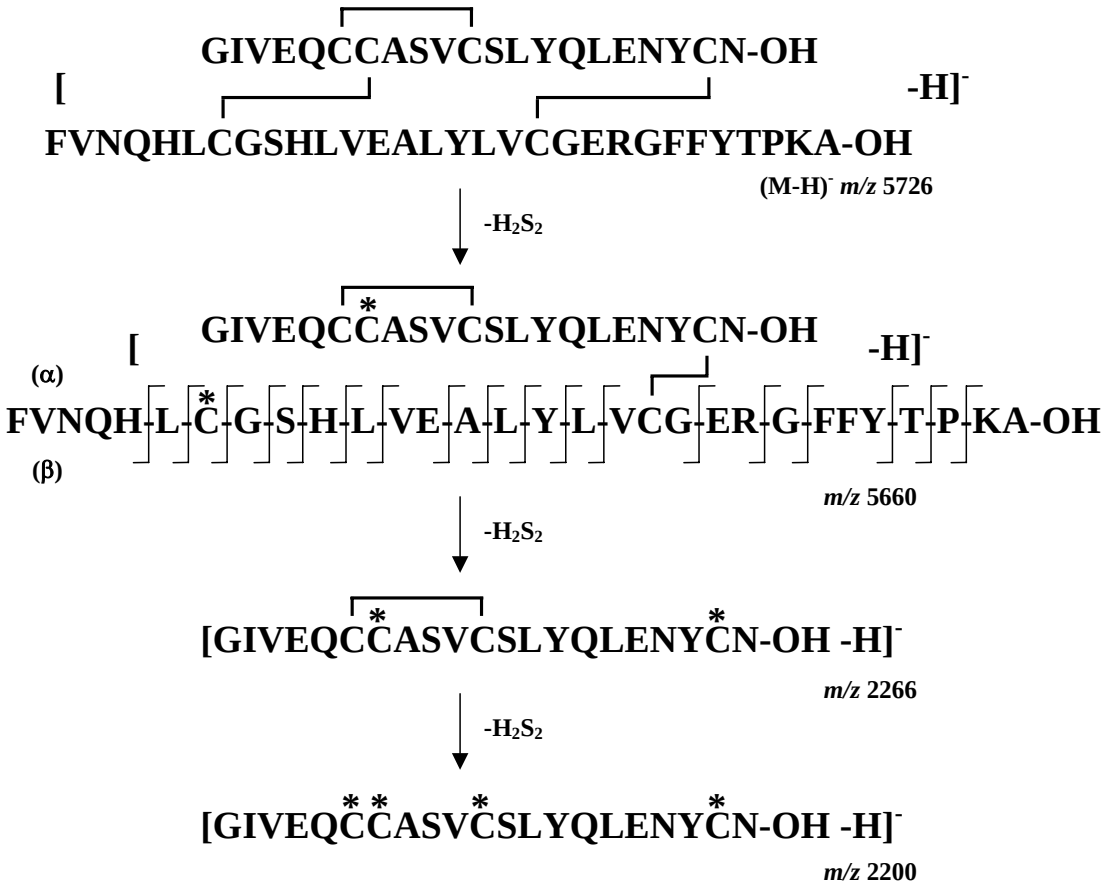
A Cys21-chain B Cys19. Chain A also contains one intramolecular disulfide link between Cys6-Cys11.

The positive ion mass spectrum of bovine insulin **(12)** has been reported previously[115] and is summarised in schematic form in Scheme 4.11. All backbone cleavages are restricted to residues outside the disulfide linkages.



Scheme 4.11. Positive ion cleavages from the $(MH)^+$ ion of bovine insulin **(12)**.[115]

The negative ion mass spectral data from bovine insulin **(12)** are shown in schematic form in Scheme 4.12. MS/MS of the $(M-H)^-$ parent ion produces a fragment anion corresponding to the loss of H_2S_2. α and β cleavages originate from the $[(M-H)^- -H_2S_2]^-$ ion and are summarised in Scheme 4.12. The backbone cleavages are consistent with the structure of bovine insulin, minus the N-terminal intermolecular disulfide link.



Scheme 4.12. A summary of negative ion cleavages from bovine insulin **(12)**. C* = NHC(=CH_2)CO.

A large peak at *m/z* 2266 is also present and corresponds to the cleavage of the second intermolecular disulfide link and formation of the chain A anion. A peak corresponding to the structure of chain B at *m/z* 3327 is also evident, but not intense enough to perform "MS/MS/MS". The "MS/MS/MS" of *m/z* 2266 (chain A) results in an additional loss of H_2S_2 from the intramolecular disulfide bridge.

4.3 Summary and Conclusions

1. Negative ion mass spectrometry is better at identifying and sequencing intramolecular disulfide bridged peptides than positive ion mass spectrometry.

2. The presence of an intramolecular disulfide is immediately evident by the facile loss of H_2S_2 from the $(M-H)^-$ parent ion. Formation of the $[(M-H)^- -H_2S_2]^-$ ion occurs via a charge-induced stepwise process. Theoretical calculations on a model system reveal that the loss of H_2S_2 is endothermic by 113 kJ mol^{-1} with the maximum transition state barrier of only 171 kJ mol^{-1} at the HF/6-31G(d)//AM1 level of theory. The loss of H_2S_2 from the parent anion opens the peptide from its cyclic form and allows for the sequencing of residues originally within the disulfide link. For peptides with two intramolecular disulfide bridges, there is a loss of $2H_2S_2$.

3. When one of the Cys residues involved in the disulfide link is in the C-terminal position, the negative ion spectrum is dominated by the loss of H_2S_2 and CO_2. (A peak corresponding to the formation of the $[(M-H)^- -H_2S_2]^-$ ion is also present). α and β fragment ions originate from the $[(M-H)^- -(H_2S_2+CO_2)]^-$ ion and identify the residues originally within the disulfide link. The loss of H_2S_2 and CO_2 first requires the cleavage of the N-terminal C-S bond. This is followed by dissociation of CO_2 and cleavage of the second carbon-sulfur bond. The highest transition state barrier for this process is only 169 kJ mol^{-1} with the product endothermic by 143 kJ mol^{-1}.

4. There is also a less abundant loss of CH_2S_2 from intramolecular disulfide bridged peptides. Loss of CH_2S_2 is initiated from the N-terminal and/or C-terminal Cys enolate anion to afford two possible structures for the $[(M-H)^- -CH_2S_2]^-$ fragment ion. Theoretical calculations reveal that the process from the C-terminal enolate is endothermic by 130 kJ mol^{-1} but has a transition state barrier of 234 kJ mol^{-1}.

4.4 Experimental

4.4.1 Peptide Synthesis

All peptides used in this study were synthesised by Mimotopes (Clayton, Victoria, Australia) using standard methods,[91] or obtained commercially from Sigma.

4.4.2 Mass Spectrometry

Electrospray MS/MS data were determined using a Micromass Q-Tof 2 orthogonal acceleration time of flight mass spectrometer with a mass range to 10 000 Da. The Q-Tof 2 is fitted with an electrospray source in an orthogonal configuration with a Z-Spray interface. Samples were dissolved in acetonitrile/water (1:1) and infused into the electrospray source at a flow rate of 5 μL min^{-1}. Conditions were as follows: capillary voltage, 3.10 kV; source temperature, 80ºC; desolvation temperature, 150ºC; cone voltage, 50 V. MS/MS data were acquired with the argon collision gas energy set between 40-80 eV to give optimal fragmentations.

"MS/MS/MS" data were determined using the Q-Tof 2 instrument by increasing the energy of the particular fragment anion under investigation in the ion source. This is essentially an MS/MS experiment on an energised source formed fragment ion.

4.4.3 Theoretical Calculations

The theoretical protocols adopted for this study have been described in Chapter 3, Section 3.4.3.

Chapter 5:

Internal Loss of Val from Citropin 1.1 and some Synthetic Analogues

5.1 Introduction

The positive ion MS/MS spectra of peptides are normally dominated by the B and Y + 2 backbone cleavages.[31,32] Under negative ion conditions, the α and β cleavages provide analogous sequencing information.[66,67] There are additional negative ion cleavages which do not have counterparts in the positive ion mode. These include the side chain cleavages which identify the presence of specific residues within the peptide.[70] There are side chain induced backbone cleavages which form γ and δ ions.[54] The formation of these ions identify the position of the residue within the peptide (and can differentiate between Lys and Gln).[78] Then there is the loss of H_2S_2 from intramolecular disulfide bridged peptides which identifies the presence of the disulfide bridges and allows for further amide cleavages to assist in sequence determination.

The final negative ion cleavage to be discussed involves the peptide anion having a specific conformational structure. This type of cleavage differs from all negative ion cleavages discusses thus far and has no positive ion counterpart.

5.1.1 Ile21/Glu23 Cyclisation and Backbone Cleavage

The negative ion mass spectrum of the peptide GLLKKLKKVAKKVLPKVVPVI-AEKL-NH_2 **(13)**[69] is shown in Figure 5.1. The α and β cleavages are shown on the formula below the spectrum with the $\beta^{|}$ cleavages displayed above the spectrum. The presence of the γ and δ cleavages from Glu23 are also seen in the spectrum. The fragment ion at *m/z* 2181 however cannot be identified by any of the fragmentations described so far. What is unusual about this fragment ion is that the peak intensity decreases as the pressure of the collision gas increases.[69] Normally backbone cleavages increase in intensity as the collision gas energy is increased.[54]

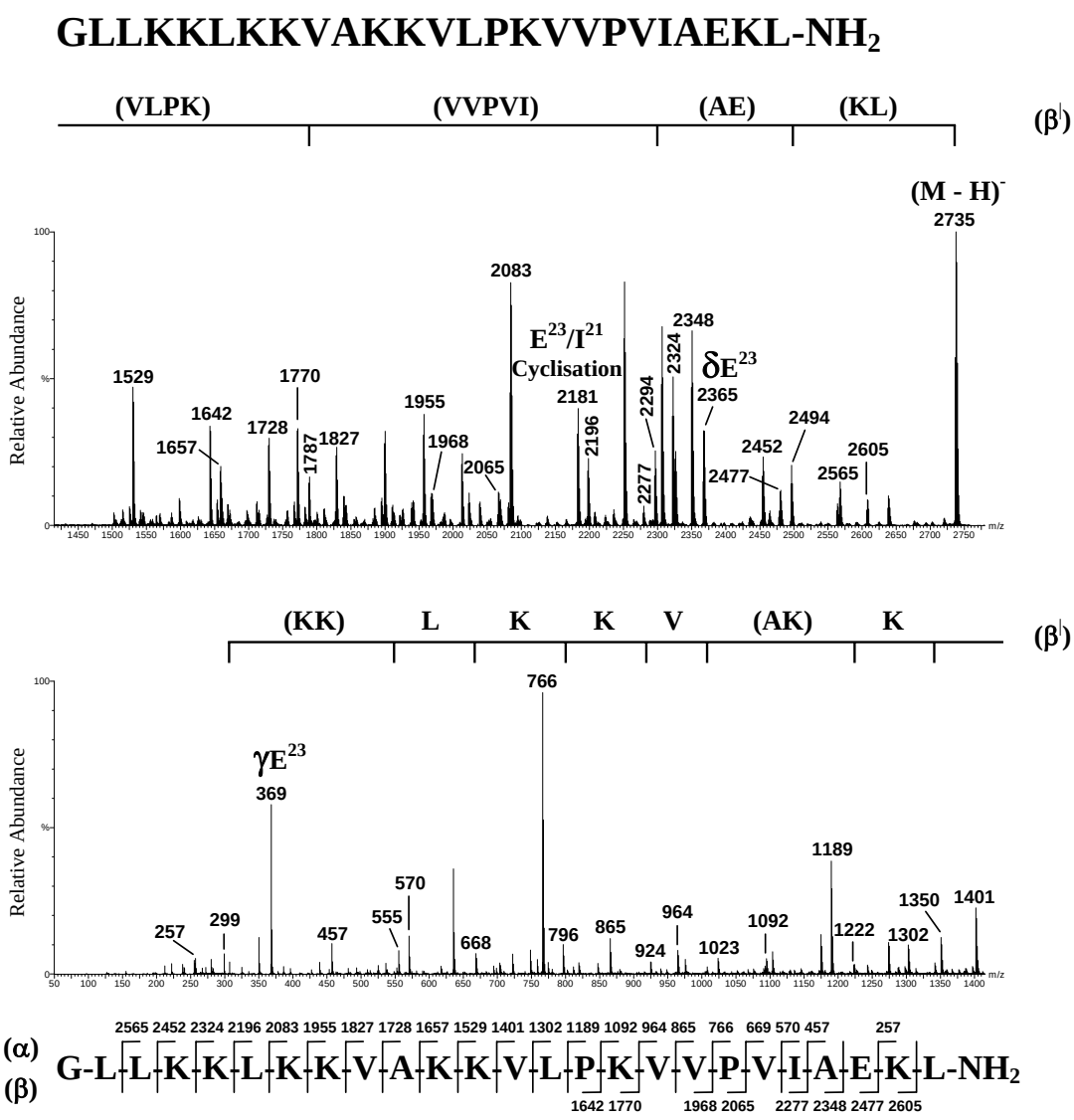


Figure 5.1. Collision-induced MS/MS data for the $(M-H)^-$ anion of **13**. α and β cleavages are shown on the formula below the spectrum with $\beta^|$ cleavages displayed above the spectrum. Magnification ranges: *m/z* 50-300 (x10), *m/z* 450-1425 (x6), *m/z* 1500-2100 (x10), *m/z* 2200-2600 (x6).[69]

The fragment ion at *m/z* 2181 indicates a loss of 554 Da from the parent anion. When the Lys residue at position 24 is replaced with His, the mass lost is 563 Da.[69] The difference of 9 Da corresponds to the difference in mass between the Lys and His residues. This means that the fragment ion at *m/z* 2181 involves the N-terminal portion of the

peptide. The formation of the fragment ion does not correspond to an amide backbone cleavage, but rather the cleavage of the NH-CH backbone bond of Ile at position 21. When the Glu23 residue is replaced with Ala23, the cleavage ion is not observed in the spectrum.[69] This means that it is the Glu23 residue that cyclises with Ile21 and not the C-terminal amide, as is the case with the Ser[78] and Cys backbone cleavages.

Ser (and Cys) induced backbone cleavages are a result of molecular recognition between the carbonyl of the C-terminal amide (or carboxylic acid) with the Ser OH (or Cys SH). This positions the NH^- (or CO_2^-) in close proximity to the electrophilic α-carbon where it can initiate nucleophilic attack. The Ile21/Glu23 cyclisation does not occur due to this form of molecular recognition. Whatever the interaction is, it is specific only for Ile21/Glu23, as a peptide ending with the sequence Ile21 Ala22 Glu23 Glu24 Leu25 (NH_2) shows cleavage due to the cyclisation of Ile21/Glu23 but not for Ala22/Glu24.[69]

Examination of the 3-dimensional structure of **13** provides a better understanding of the Ile21/Glu23 cleavage scenario. Figure 5.2 illustrates the α-helical structure of **13**, and also the proximity of the Glu side chain oxygen with the electrophilic α-carbon of the Ile residue at position 21.[69] This may also explain why the peak intensity resulting from this fragmentation decreases when the energy of the collision gas is increased. As the energy increases the peptide begins to lose its stable conformation and most likely pulls the Ile21 and Glu23 residues apart and into a conformation where the carboxylate anion is too far away to initiate nucleophilic attack.

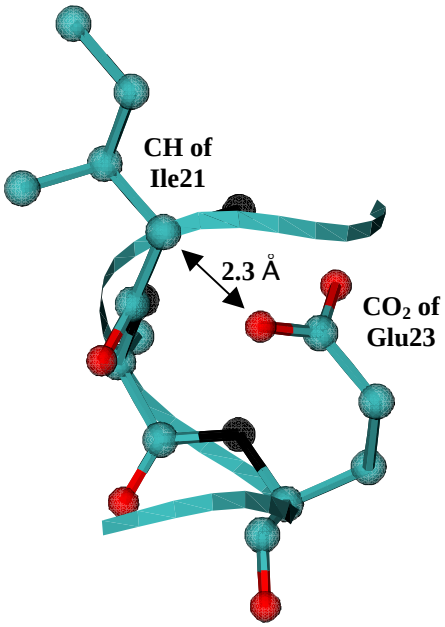


Figure 5.2. A computer simulation displaying the position of the Glu and Ile residues within the α-helical structure of the peptide.[69]

5.1.2 Internal Loss of Val from Citropin Peptides

The citropin 1.1 peptide, GLFDVIKKVASVIGGL-NH_2 was isolated from the amphibian *Litoria citropa*.[29] Its biological activity includes wide-spectrum antibacterial action against Gram positive organisms, significant activity against all major forms of human cancers, is a fungicide and inhibits the production of nitric oxide by neuronal nitric oxide synthase.[105]

The negative ion mass spectrum of citropin 1.1 contains a peak corresponding to a fragment ion formed by loss of 99 Da (Val) from the $(M-H)^-$ parent ion. This fragment ion is unusual because there is no Val present in either the N-terminal or C-terminal position of the peptide. Citropin 1.1 contains three Val residues. Which Val residue is lost in this unusual fragmentation?

5.2 Results and Discussion

To determine (i) which Val residue is lost from the $(M-H)^-$ parent ion of citropin 1.1 and (ii) the mechanism of this unusual fragmentation, a number of synthetic analogues were analysed. Table 5.1 contains the sequences of citropin 1.1 and the 17 modifications used in this investigation. The bold font in the table indicates the modified residues relative to citropin 1.1. Peptides which undergo loss of Val from the $(M-H)^-$ parent ion and other fragment ions are also indicated in the table.

Table 5.1. Loss of Val from Citropin 1.1 and synthetically modified analogues.

Peptide	Sequence	$(M-H)^-$ -Val	$[(M-H)^- - H_2O]^-$ -Val	$[(M-H)^- - CH_2O]^-$ -Val	$[(M-H)^- - (H_2O + CH_2O)]^-$ -Val	$[(M-H)^- - H_2S]^-$ -Val	*m/z* = 748 or analogue
Citropin 1.1	GLFDVIKKVASVIGGL-NH_2	X	X	X	X		X
Mod 1	GLF**E**VIKKVASVIGGL-NH_2	X	X	X	X		
Mod 2	GLFDVI**A**KVASVIGGL-NH_2	X	X	X	X		*m/z* 691
Mod 3	GLFDVIKKVASVI**K**GL-NH_2	X	X	X	X		X
Mod 4	GLFDVIKKVASVIGG**A**-NH_2	X	X	X	X		X
Mod 5	GLFDVIKKVA**A**VIGGL-NH_2	X	X				
Mod 6	GLFDVIKKVASVI**KK**L-NH_2	X	X	X			X
Mod 7	GLF**A**VIKKVA**C**VIGGL-NH_2	X				X	
Mod 8	GLF**A**VIKKVASVI**K**GL-NH_2	X		X			
Mod 9	GLF**A**VIKKVA**A**VI**K**GL-NH_2	X					
Mod 10	GLF**A**VIKKVA**A**VI**KK**L-NH_2	X					
Mod 11	GLF**A**VIKKVA**K**VI**KK**L-NH_2	X					
Mod 12	**K**LF**A**VIKKVA**A**VIGGL-NH_2	X					
Mod 13	**K**LF**A**VIKKVA**A**VI**RRK**-NH_2	X					
Mod 14	GLFDVI**QQ**VASVIGGL-NH_2	X					X
Mod 15	GLFD**II****QQ**VASVIGGL-NH_2	X					*m/z* 762
Mod 16	GLFDVI**QQI**ASVIGGL-NH_2	X					*m/z* 762
Mod 17	GLFDVI**QQ**VASI**I**GGL-NH_2						X

The abundance of the $[(M-H)^- - Val]^-$ fragment ion varies from peptide to peptide. It is usually between 5 and 50% of the intensity of the parent anion. For peptides containing an Asp or Glu residue, an $[(M-H)^- - (H_2O+Val)]^-$ fragment ion is also observed. For Ser peptides there is an $[(M-H)^- - (CH_2O+Val)]^-$ ion formed, and for the Cys containing peptide an $[(M-H)^- - (H_2S+Val)]^-$ fragment ion. Loss of Val is also evident from the $[(M-H)^- - (H_2O+CH_2O)]^-$ fragment ions of peptides containing both Asp/Glu and Ser.

The largest $[(M-H)^-$ -Val$]^-$ fragment ion is formed from citropin 1.1 synthetic modification 14 with the sequence GLFDVIQQVASVIGGL-NH_2 **(14)**. This peptide was therefore chosen to investigate the mechanism for the internal loss of the Val residue. Modifications 15-17 were next synthesised, where the Val residue was sequentially replaced by Ile to determine which Val is lost from the $(M-H)^-$ parent ion. (Ile was chosen rather than Gly, in case the side chain itself is a prerequisite for the loss). The negative ion mass spectral data for modifications 15-17 are summarised in Table 5.1. When Val12 is replaced with Ile12, there is no loss of Val.

The negative ion spectrum of **14** is shown in Figure 5.3. The α and β cleavages are shown on the formula above the spectrum. The γ ions from Asp4, Gln7, Gln8 and Ser11 are all present. The Asp4 γ ion further fragments in accord with the side chain losses of H_2O from Asp4 and CH_2O from Ser11. The loss of Val is evident from the parent ion and also from the γAsp4-H_2O and γAsp4-(H_2O+CH_2O) fragment ions.

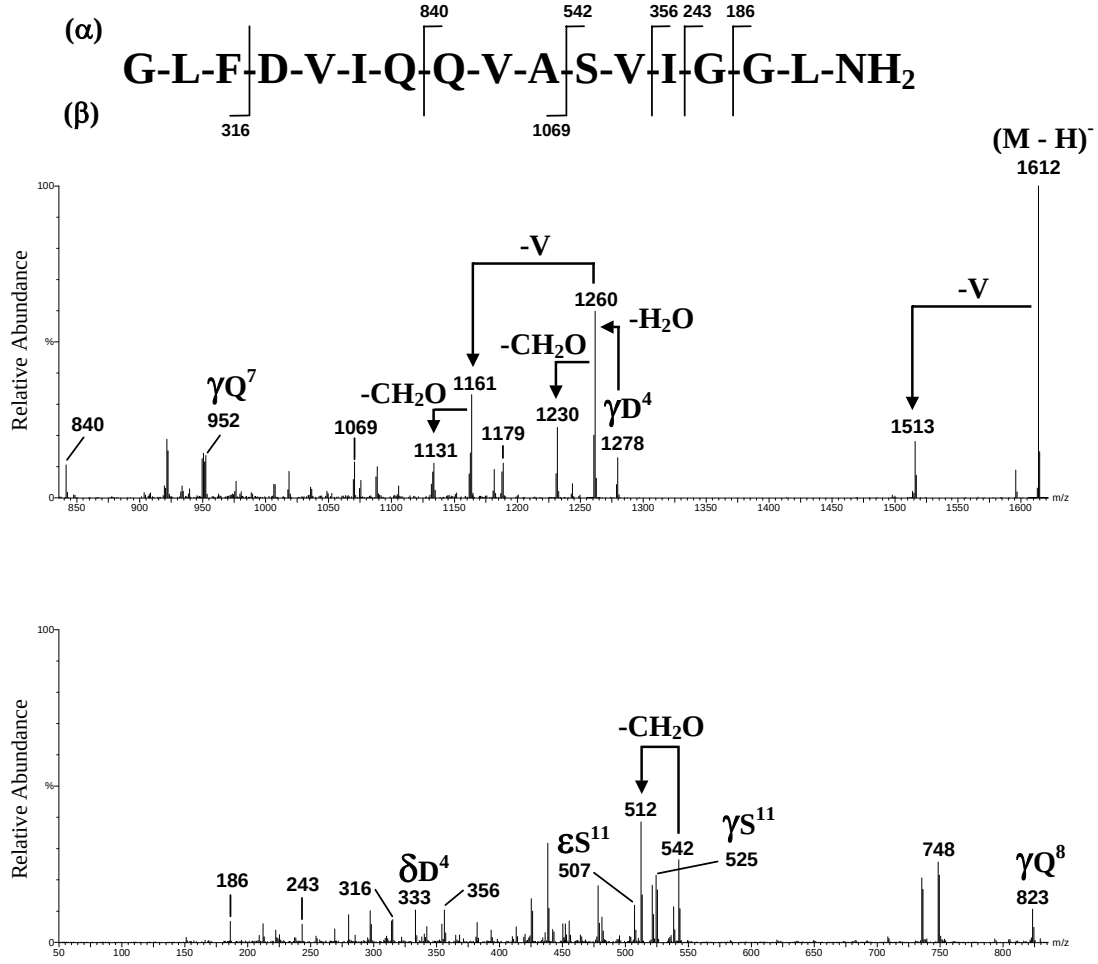


Figure 5.3. Collision-induced MS/MS data of **14**. Negative ion fragmentations of the $(M-H)^-$ ion. α and β cleavages are shown on the formula above the spectrum. Magnification ranges: *m/z* 150-550 (x16), *m/z* 800-1200 (x10), *m/z* 1200-1300 (x2).

For the Val12 residue to be lost from the $(M-H)^-$ ion, a cyclisation must occur which places Val12 in a terminal position. The MS/MS/MS spectrum of the $[(M-H)^-$ -Val$]^-$ anion (*m/z* 1513) shown in Figure 5.4 identifies the structure of this anion, and provides insight into the mechanism of the process.

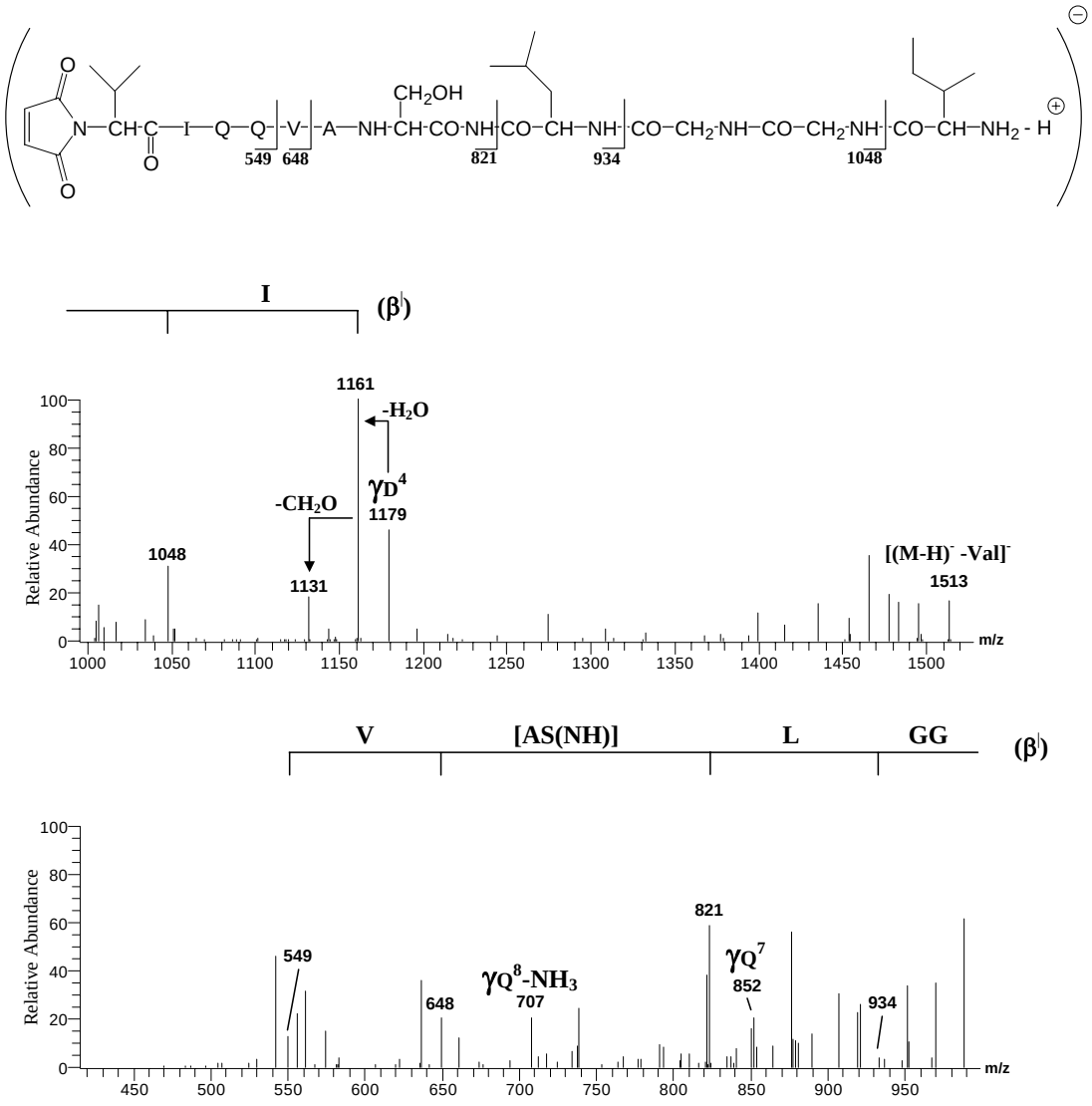


Figure 5.4. MS/MS/MS of the $[(M-H)^- -Val]^-$ fragment ion. The structure above the spectrum corresponds to the fragment ion at *m/z* 1161. $\beta^|$ cleavages are shown above the spectrum and summarised on the structure. Magnification ranges: *m/z* 525-575 (x100), *m/z* 575-850 (x20), *m/z* 850-1000 (x100), *m/z* 1000-1050 (x50).

The γAsp4-H_2O fragment ion (*m/z* 1161) dominates the spectrum shown in Figure 5.4 with $\beta^|$ cleavages subsequently formed from this ion. There are no α and β cleavages. γGln7 and γGln8-NH_3 fragmentations are also observed but there is no characteristic γSer11 fragment ion. This indicates that there may be some form of structural modification of the peptide in the vicinity of Ser11. This is confirmed by an unusual

feature of the $\beta^{|}$ cleavages: namely SerNH is lost instead of Ser. These data suggest that the fragment ion at *m/z* 1161 has the structure shown above the spectrum in Figure 5.4. This implies that the loss of Val is preceded by attack of the C-terminal $CONH^-$ moiety at the Ser11 carbonyl group, followed by ring opening to place Val12 in the C-terminal position.

There are two possible mechanisms to account for why cyclisation at the Ser residue is preferred over cyclisation at other residues of the peptide. The first is that the C-terminal amide hydrogen bonds with the Ser OH and places the $CONH^-$ in a position where is can undergo nucleophilic attack on the Ser carbonyl group. This mechanism is excluded from a consideration of the mass spectral data of other peptides in Table 5.1. Peptides that have Ser11 replaced with Ala11 or Lys11 still produce the $[(M-H)^- -Val]^-$ fragment ion. Additionally, even when Ser11 is present, the loss of Val does not require hydrogen bonding with Ser OH as loss of Val can occur from the $[(M-H)^- -CH_2O]^-$ fragment ion. (The side chain loss of CH_2O from Ser converts Ser to Gly).

The second mechanism involves the peptide acquiring a specific conformation where the C-terminal $CONH^-$ and Ser carbonyl are in close proximity, thus allowing nucleophilic attack to proceed. Figure 5.5 illustrates the amphipathic α-helical structure of citropin 1.1, as determined by 2D nuclear magnetic resonance.[29]

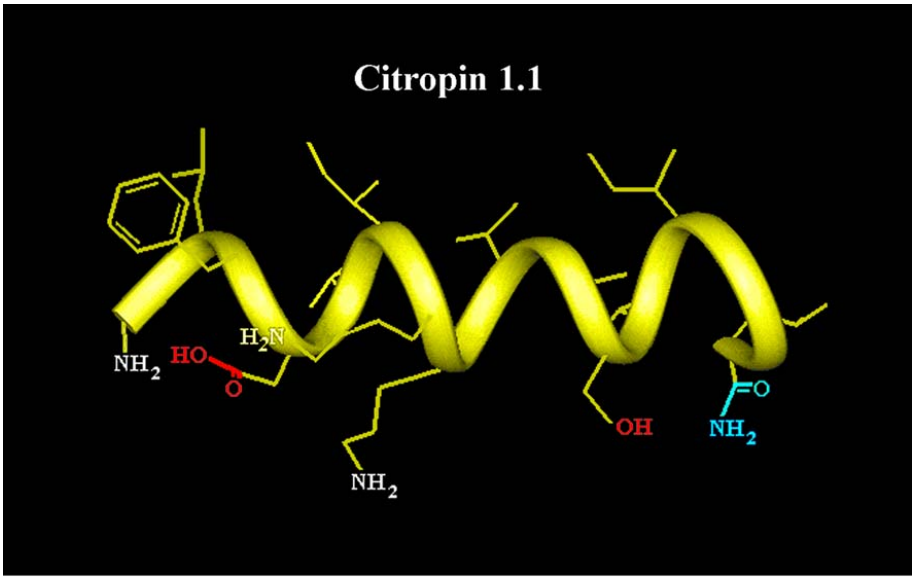


Figure 5.5. The secondary structure of citropin 1.1 as determined by 2D NMR.[29]

The region in the vicinity of Ser11 and the C-terminal amide is enhanced in Figure 5.6. This illustrates the α-helical conformation of the peptide and indicates that the C-terminal $CONH^-$ moiety may approach to within 2.5 Å of the Ser carbonyl group, a position from which it can initiate nucleophilic attack on the Ser11 carbonyl group.

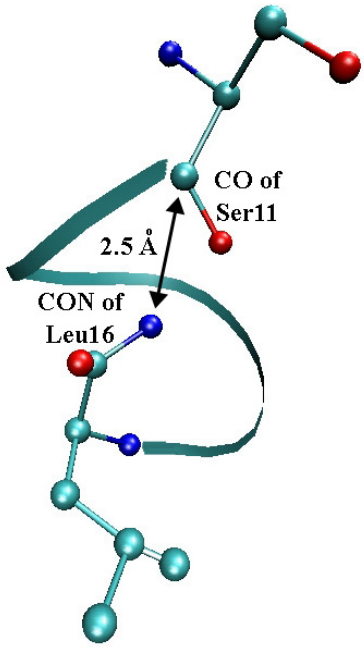


Figure 5.6. Computer simulation of the Ser11 and C-terminal region of citropin 1.1.

When the collision gas energy is slowly increased during the MS/MS experiment, one of the first fragment ions to form is the loss of Val from the parent ion. As the collision gas energy is increased further, the α, β and γ ions increase in abundance and intensity. When the collision gas energy is further increased, the intensity of the $[(M-H)^-$ -Val$]^-$ ion decreases while other backbone cleavage ions retain or increase their intensity. These results are consistent with the loss of Val occurring from the parent anion of an α-helical structured peptide. At low collision energy, a large proportion of the peptide is in the α-helical conformation allowing for nucleophilic attack of the C-terminal amide on the Ser carbonyl group. Increasing the collision gas energy increases the internal energy of the system, unravelling the α-helical conformation of the peptide.

Changes to the pH of the electrospray solution have an interesting effect on the intensity of the $[(M-H)^-$ Val$]^-$ fragment ion. The results are summarised in Table 5.2. The

ratios of the abundances of $[(M-H)^-$ Val$]^-$ and $(M-H)^-$ are normalised to 1:1 when the acetonitrile/water (1:1) solvent is at the usual pH of 3.46. When the pH is decreased to 1.5, the abundance of $[(M-H)^-$ Val$]^-$ drops relative to the $(M-H)^-$ ion. The same outcome is observed when the pH is increased, and at pH 13.4, the $[(M-H)^-$ Val$]^-$ fragment ion is no longer present. The abundances of the α, β and γ fragment ions remain virtually unchanged at the differing pH levels.

Table 5.2. Ratio of $[(M-H)^-$ Val$]^-$ and $(M-H)^-$ under differing solvent conditions.

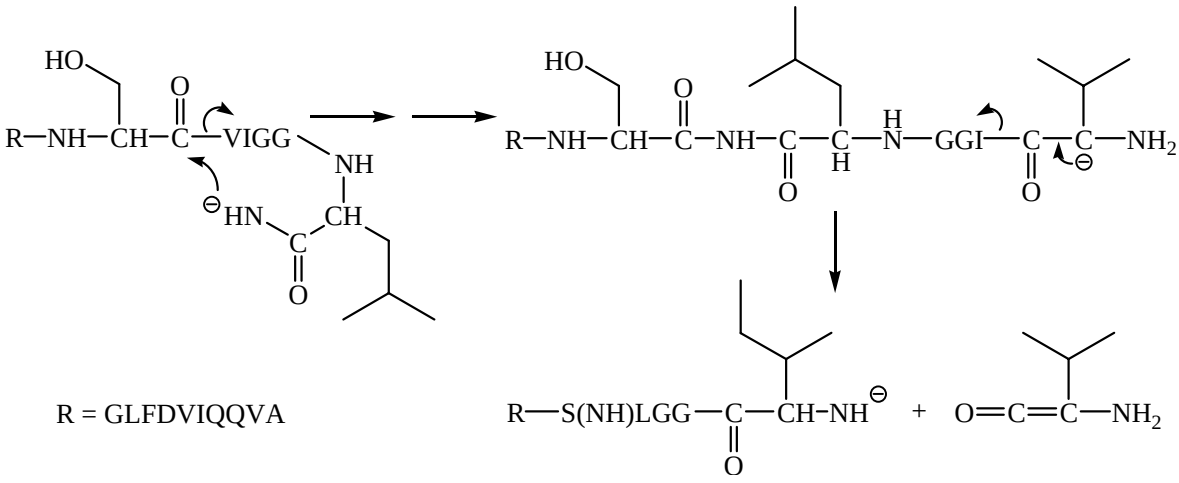
Solvent	pH	Ratio of relative abundances of $[(M-H)^-$ -Val$]^-$: $(M-H)^-$
Acetonitrile/Water (1:1)	1.5	1:16
	3.46	1:1
	7.0	1:4
	10.1	1:12
	13.4	Not present
Isopropanol		30:1

At both low and high pH levels, the peptide hydrogen bonds more readily with the solvent rather than other residues along the peptide chain. This causes the peptide to unravel from its α-helical conformation, pulling the C-terminal amide away from the Ser carbonyl group. The increased distance between these functional groups reduces the likelihood of nucleophilic attack taking place, and results in a reduction of the abundance of the $[(M-H)^-$ -Val$]^-$ ion relative to the parent $(M-H)^-$ ion. α, β and γ cleavages may occur from both the α-helical and unravelled conformation of the peptide, and hence these fragmentation ions remain virtually unchanged at low and high pH levels.

When the spectrum was measured in pure isopropanol, the loss of Val was enhanced relative to the intensity of the parent ion and all other backbone cleavage ions.

Isopropanol hydrogen bonds less readily than water with the peptide, thus allowing a larger proportion of the α-helical conformation in solution. These results are also consistent with the loss of Val occurring from an $(M-H)^-$ anion of helical structure.

The mechanism for loss of Val from the $(M-H)^-$ parent ion of **14** is shown in Scheme 5.1. The C-terminal $CONH^-$ moiety effects nucleophilic attack at the Ser carbonyl group. The amide bond between Ser and Val is cleaved, placing Val in the C-terminal position of the peptide.

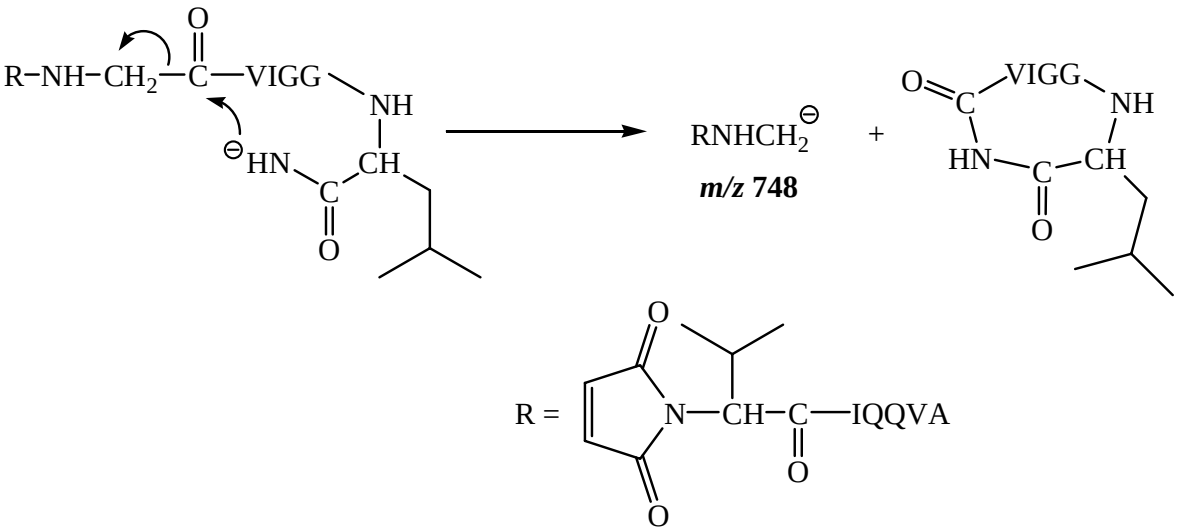


Scheme 5.1. The mechanism by which Val is placed in the C-terminal position, and lost to form the $[(M-H)^- -Val]^-$ fragment ion for peptide **14**.

There is a second fragment anion shown in Figure 5.3 which (like the loss of Val) also appears anomalous, namely that at *m/z* 748. The *m/z* 748 fragment ion originates from the γAsp4-(H_2O+CH_2O) ion. No *m/z* 748 fragment ion is seen in the MS/MS/MS spectrum of $[(M-H)^- -Val]^-$ (Figure 5.4). Analogous fragmentations from citropin 1.1 and other synthetic modifications are highlighted in Table 4.1. Just like the loss of Val, the peak

intensity of *m/z* 748 is high when the collision gas energy is low, but decreases as the collision gas energy is increased.

The mechanism for formation for the fragment ion at *m/z* 748 from the γAsp4-(H_2O +CH_2O) ion is shown in Scheme 5.2. This reaction competes with the loss of Val from γAsp4-(H_2O+CH_2O). The two reaction pathways have a common intermediate structure. Following cyclisation, cleavage of the amide bond places Val in the C-terminal position of the peptide (the precursor for loss of Val) while cleavage of the bond N-terminal to the carbonyl group produces the fragment ion at *m/z* 748.



Scheme 5.2. Mechanism of formation of *m/z* 748 from the γAsp4-(H_2O+CH_2O) ion.

The two competing reaction pathways have been examined using theoretical calculations on a simplified model system. The model consists of an N-terminal region with the sequence $HCONHCH_2CONHCH_2CHO$ (to represent the Ser and Val portion of the peptide) and a C-terminal $HCONH^-$. The reaction coordinate profile is illustrated in Figure 5.7.

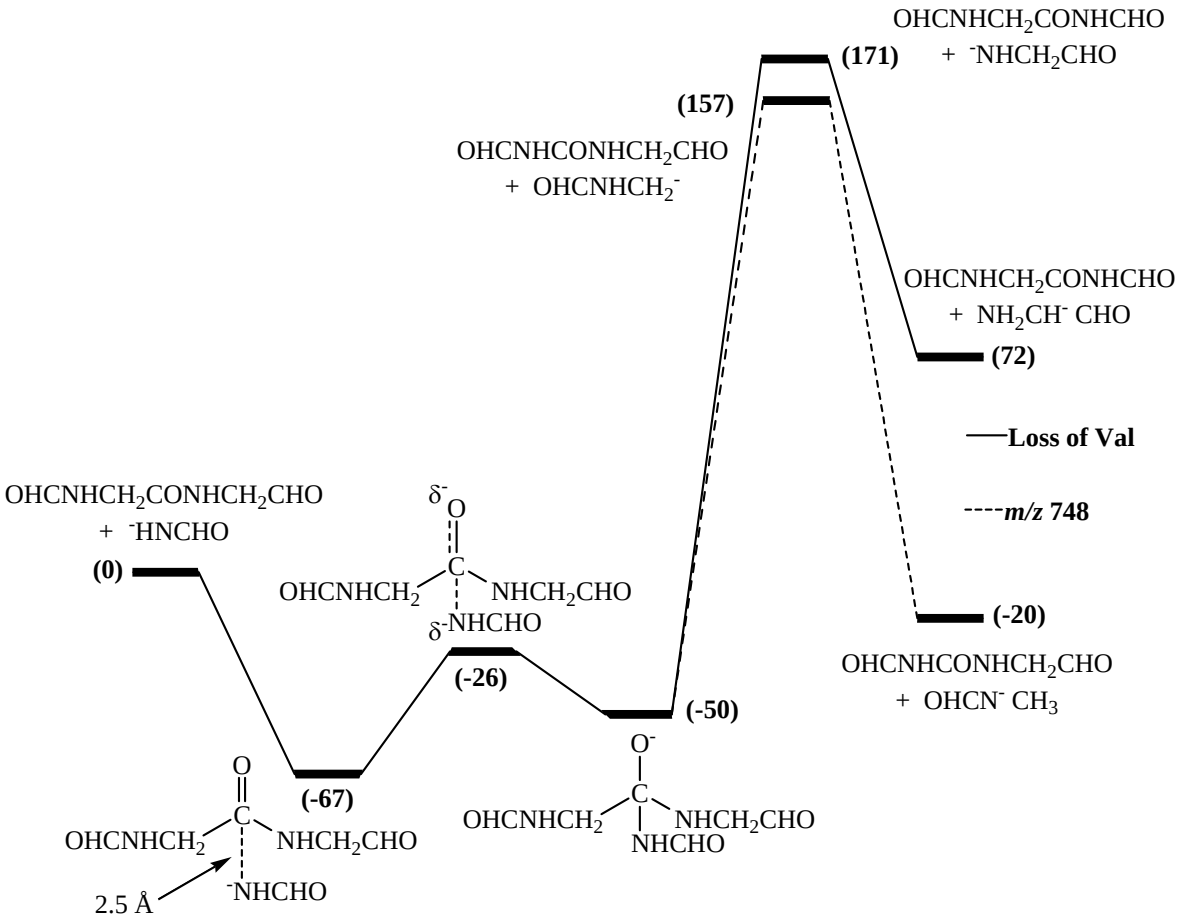


Figure 5.7. The cyclisation and formation of the fragment ion at *m/z* 748 (or analogue) and formation of C-terminal Val (the precursor for loss of Val) on the model system $HCONHCH_2CONHCH_2CHO$ and $HCONH^-$. Relative energies in kJ mol^{-1}. Geometries and energies at the HF/6-31G(d)//HF/6-31G(d) level of theory. Transition states confirmed by IRC calculations.

The α-helical conformation of citropin peptides positions the C-terminal amide to within 2.5 Å from the Ser carbonyl group. In the model system, fixing these two groups 2.5 Å apart furnished a complex 67 kJ mol^{-1} lower in energy than the starting reagents. From this complex, only 41 kJ mol^{-1} of energy is required to overcome the transition state barrier of the S_N2 reaction.

Once the tetrahedral intermediate is formed, the reaction can take one of two pathways. Cleavage of the amide bond (and formation of the precursor which loses Val) requires 171 kJ mol^{-1} of energy to effect dissociation. Additional proton transfer forms the more stable anion, with the overall process endothermic by 72 kJ mol^{-1}. The alternative pathway requires 157 kJ mol^{-1} to form the *m/z* 748 fragment ion (or analogue) with the overall process exothermic by 20 kJ mol^{-1}. Both reaction pathways are energetically favourable compared to α and β backbone cleavages (Figure 2.8).

The relative energies of the penultimate products for the two processes in Figure 5.7 differ by only 14 kJ mol^{-1} at the HF/6-31G(d)//HF/6-31G(d) level of theory. Because the difference in energy is small, and determines the relative kinetics of the two processes, relative energies of the penultimate products were calculated at higher levels of theory. The results are as follows: HF/6-31G(d)//HF/6-31G(d) [+14 kJ mol^{-1}], HF/6-31+G(d)//HF/6-31G(d) [+30 kJ mol^{-1}], HF/6-311G(d)//HF/6-31G(d) [+30 kJ mol^{-1}], HF/6-311+G(d)//HF/6-31G(d) [+18 kJ mol^{-1}], B3LYP/6-31G(d)//HF/6-31G(d) [-16 kJ mol^{-1}], B3LYP/6-31+G(d)//HF/6-31G(d) [-8 kJ mol^{-1}], B3LYP/6-31G(d)//B3LYP/6-31G(d) [-21 kJ mol^{-1}], B3LYP/aug-cc-pVDZ//B3LYP/6-31G(d) [-10 kJ mol^{-1}]. The only conclusion that can be made from these data is that the energies of the penultimate products of the two processes in Figure 5.7 are very similar.

5.3 Summary and Conclusions

1. The internal loss of Val from the $(M-H)^-$ parent ion (and other fragment ions) of citropin 1.1 involves Val12.

2. The loss of Val requires the peptide anion to be in an α-helical conformation. In this conformation the C-terminal $CONH^-$ moiety may adopt a position only some 2.5 Å from the Ser11 carbonyl group. Nucleophilic attack then cleaves the amide bond, placing Val12 at the C-terminal position of the peptide. The mechanism for loss of Val shows a barrier of 171 kJ mol^{-1}, with the overall process endothermic by 72 kJ mol^{-1} at the HF/6-31G(d)// HF/6-31G(d) level of theory.

3. Citropin 1.1 and analogues which contain both an Asp and Ser residue may yield a pronounced peak corresponding to the γAsp4-(H_2O+CH_2O) fragment ion. This ion can undergo cyclisation (cf. 2 above) and (i) lose Val or (ii) cleave the bond N-terminal to the carbonyl as shown in Scheme 5.2. The second reaction pathway competes with loss of Val, requiring 157 kJ mol^{-1} of energy to overcome the highest energy barrier with the overall process exothermic (-20 kJ mol^{-1}).

5.4 Experimental

5.4.1 Peptide Synthesis

Citropin 1.1 and the synthetic modifications used in this study were synthesised by Mimotopes (Clayton, Victoria, Australia) using standard methods.[91]

5.4.2 Mass Spectrometry

For full details of MS/MS experimental conditions used on the Micromass Q-Tof 2 mass spectrometer, see Chapter 4, section 4.4.2.

The MS/MS/MS experiments were carried out using the Finnigan LCQ ion trap mass spectrometer. Samples were dissolved in acetonitrile/water (1:1) and infused into the electrospray source at a flow rate of 8 μL min^{-1}. Conditions were as follows: source voltage, 4.3 kV; source current, 18 μA; capillary temperature, 200ºC; capillary voltage, 3 V, sheath gas flow, 30 psi. MS/MS/MS data were acquired with the automatic gain control on, a maximum time of 500 ms and using three microscans per scan, averaging over a total of 20 scans.

5.4.3 Theoretical Calculations

Geometry optimisations and energies were calculated at the HF/6-31G(d) level of theory using the GAUSSIAN 98 suite of programs[93] on the Alpha Server SC [Australian Partnership for Advanced Computing (APAC) National Facility, Canberra, Australia]. Stationary points were characterised as either minima (no imaginary frequencies) or transition states (one imaginary frequency) by calculation of frequencies using analytical gradient procedures. Transition states were confirmed using Intrinsic Reaction Coordinate (IRC) calculations.

Chapter 6:

Studies of the Dermal Secretions of *Litoria peronii*

6.1 Introduction

Litoria peronii, also known as Peron's Tree Frog is one of approximately 200 Australian amphibians identified thus far.[7] Its body is predominantly light brown or grey in colour with small pale green spots (Figure 6.1). The ventral surfaces are white and the rear of the thighs consist of black on yellow or orange. Males are generally 44-53 mm in length with females slightly larger at 46-65 mm. Similar species include *Litoria rothii* and *Litoria tyleri*.[7,154]



Figure 6.1. *Litoria peronii*.

Litoria peronii inhabits a large area which includes south-eastern Queensland, eastern and southern New South Wales, northern Victoria and along the Murray River into South Australia (Figure 6.2). They can be generally found under loose bark of eucalyptus trees beside rivers and waterholes, and on lower vegetation near water at night. They call from vegetation around temporary ponds or permanent waterholes, often well above ground level. The call is a long series of 29-50 explosive notes or bursts of sound like laughter, described as "maniacal cackle". *Litoria peronii* tadpoles are pale golden yellow with a dark lateral stripe and can reach a maximum recorded length of 44 mm.[7]



Figure 6.2. The distribution of *Litoria peronii* throughout Australia.[7]

The dermal peptide secretions of Australian amphibians studied to date generally remain consistent throughout the year. However, *Litoria splendida* is one example where two additional peptides are secreted in winter, while relative concentrations of other peptides remain similar to summer levels.[106] *Litoria rothii* on the other hand shows a large seasonal variation in peptide secretions (see Table 6.1). A total of 15 peptides are secreted,

but only five are observed in both summer and winter. Six peptides are secreted only in summer and four only in winter.[101]

Table 6.1. Peptide skin secretions of *Litoria rothii*.[101] Peptides isolated in summer and winter are assigned S+W while those isolated only in summer or winter are assigned S and W respectively.

Name		Sequence	MW	Season
Rothein	1	SVSNIPESIGF-OH	1148	S+W
	2.1	AGGLDDLLEPVLNSADNLVHGL-OH	2230	S+W
	2.2	AGDDETLLKPVLNSLDNLVSGL-NH_2	2280	S+W
	3.1	ASAAGAVRAGDDETLLKPVLNSLDNLVSGL-NH_2	2963	W
	3.2	DSAAGAVRAGDDETLLKPVLNSLDNLVSGL-NH_2	3007	W
	3.3	ASAAGAVRAGDDETLLNPVLNSLDNLVSGL-NH_2	2949	W
	3.4	ASAAGAVREDDDETLLNPVLNSLDNLVSGL-NH_2	3065	W
	4.1	AEILFGDVRPPWMPPPIFPEMP-OH	2534	S
Caerulein	1.1	pEQDY(SO_3)TGWMDF-NH_2	1351	S+W
	1.2	pEQDY(SO_3)TGWFDF-NH_2	1367	S+W
Caerin	1.1	GLLSVLGSVAKHVLPHVVPVIAEHL-NH_2	2581	S
	1.6	GLFSVLGAVAKHVLPHVVPVIAEKL-NH_2	2590	S
	1.10	GLLSVLGSVAKHVLPHVVPVIAEKL-NH_2	2572	S
	2.1	GLVSSIGRALGGLLADVVKSKGQPA-OH	2391	S
	3.1	GLWQKIKDKASELVSGIVEGVK-NH_2	2381	S

Does *Litoria peronii*, a species very similar to *Litoria rothii*, (i) secrete similar peptides to *Litoria rothii*, and (ii) alter the secreted peptide profile seasonally?

6.2 Results and Discussion

6.2.1 Isolation of Peptides

Skin secretions of *Litoria peronii* were collected from three individual animals in both summer and winter (January and July respectively). The secretions were purified by high performance liquid chromatography (HPLC). The summer and winter HPLC traces of *Litoria peronii* are shown in Figure 6.3. Fractions containing peptide material are labelled in the figure. A total of 12 peptides were identified with ten present in the summer secretion and eight in the winter secretion.

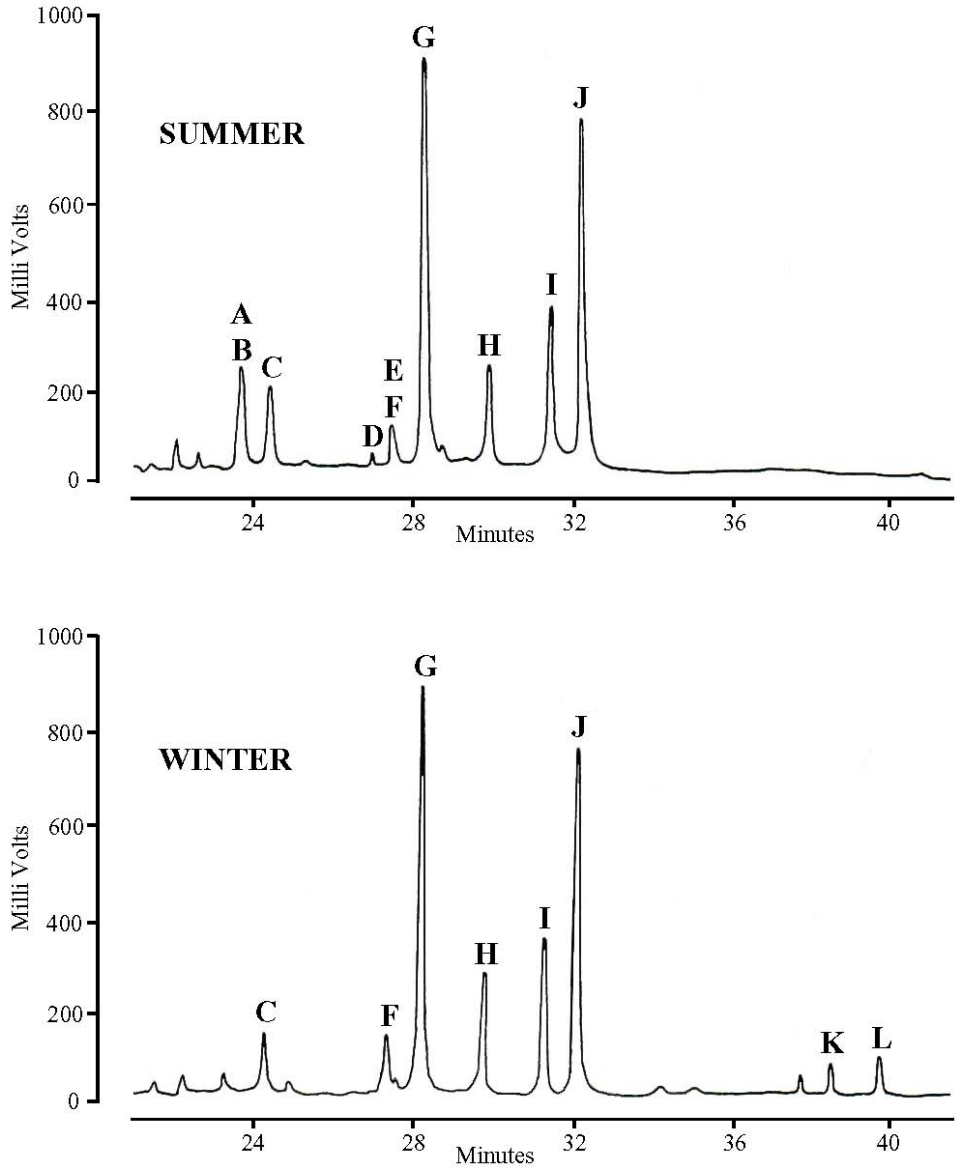


Figure 6.3. HPLC traces of summer and winter skin peptide secretions of *Litoria peronii*.

A combination of methods, including positive and negative ion mass spectrometry and automated Edman degradation was utilised to identify the sequence of the peptides. The sequences of the peptides secreted by *Litoria peronii* are listed in Table 6.2.

Table 6.2. Peptides isolated from *Litoria peronii*. Peptides isolated in summer and winter are assigned S+W while those isolated only in summer or winter are assigned S and W respectively.

Name		Sequence	MW	HPLC Fraction	Season
Peronein	1.1	pEPWLPFG-NH_2	825	G	S+W
	1.2	pEPWIPFV-NH_2	867	I	S+W
	1.3	pEPWLPFV-NH_2	867	J	S+W
	1.4	pETWLPFV-NH_2	871	H	S+W
	1.5	pEPWLPFR-NH_2	924	F	S+W
pro-Peronein	1.1A	DAQEKRQPWLPFG-NH_2	1569	A	S
	1.1B	SEEEKRQPWLPFG-NH_2	1600	B	S
	1.2A	DAQEKRQPWIPFV-NH_2	1611	D	S
	1.3A	DAQEKRQPWLPFV-NH_2	1611	E	S
Caerulein	1.1	pEQDY(SO_3)TGWMDF-NH_2	1351	C	S+W
Caerin	1.1	GLLSVLGSVAKHVLPHVVPVIAEHL-NH_2	2582	L	W
	2.1	GLVSSIGRALGGLLADVVKSKGQPA-OH	2391	K	W

Caerulein 1.1 is present in both the summer and winter secretions. It is a neuropeptide and has been isolated previously from several species of Australian amphibians.[2,11,49,101-107] Post translational modifications of caerulein 1.1 include an N-terminal pGlu residue, a C-terminal amide and a sulfated Tyr. The caerin peptides were present only in the winter secretions. Both caerin 1.1[2,11,101,106,107,155,156] and 2.1[2,101,106,107] have been isolated previously from Australian amphibians.

The major components of the skin secretions of *Litoria peronii*, in both summer and winter are the peroneins. These peptides have not been isolated previously. The final four peptides (isolated in summer only) are not active peptides, but rather, propeptides of the peroneins.

6.2.2 Peptide Sequence Determination

The first group of peptides to be considered are the peroneins. The positive ion mass spectrum of peronein 1.1 is shown in Figure 6.4. The B fragmentations are shown above the spectrum and Y + 2 fragmentations below the spectrum. A combination of B and Y + 2 cleavage data provide the sequence of the peptide, except for identifying the first two residues and differentiating between isomeric Leu and Ile at residue four. The mass of 208 Da for the first two residues corresponds to three possible N-terminal sequences, pyro-Glu Pro, Ala His and His Ala. The absence of Edman sequencing data suggests that peronein 1.1 has a post translationally modified pyro-Glu N-terminal residue. The loss of 17 Da (NH_3) from the $(MH)^+$ parent ion indicates that peronein 1.1 contains a post-translationally modified C-terminal amide.

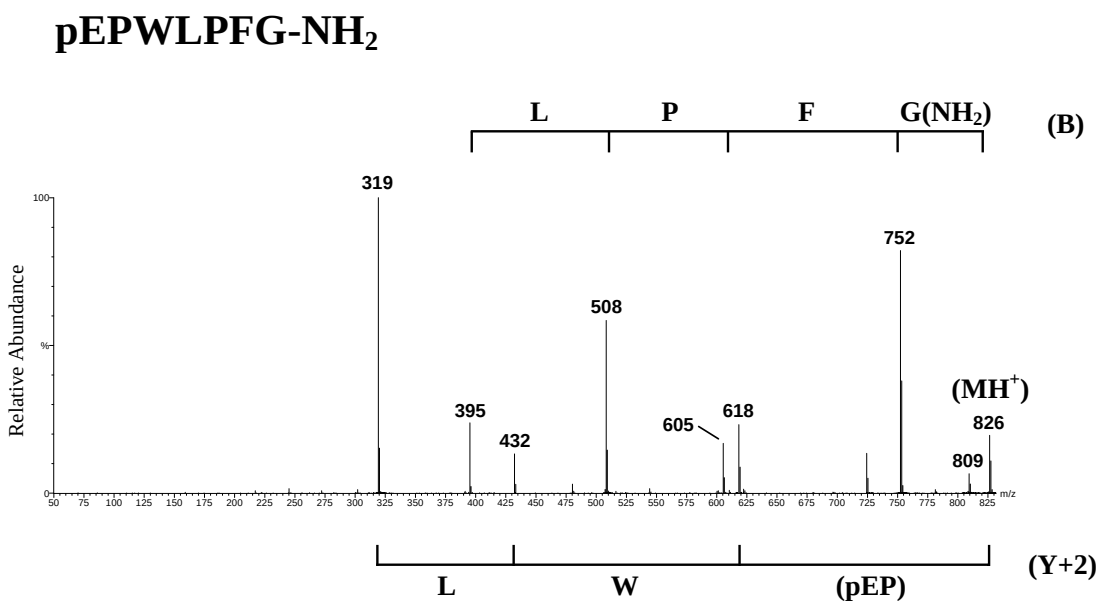


Figure 6.4. Collision-induced MS/MS data of peronein 1.1. Positive ion fragmentations of the $(MH)^+$ species, B and Y + 2 cleavage ions are shown above and below the spectrum respectively. Magnification range: *m/z* 400-800 (x4).

Additional sequencing data was provided by the negative ion mass spectrum of peronein 1.1 (Figure 6.5). The α and β fragmentations are shown above and below the spectrum respectively. The α cleavages from the $(M-H)^-$ parent ion identify the sequence of the first two residues as pyro-Glu Pro. The large fragment ion at *m/z* 695 corresponds to the side chain loss of C_9H_7N from Trp (see Table 2.1), converting Trp to Gly. Several α fragmentation ions and one β fragment ion originate from the $[(M-H)^- - C_9H_7N]^-$ fragment anion. The γPhe6, δPhe6 and δTrp3 backbone cleavages are also observed in the spectrum. The peak at *m/z* 491 corresponds to either (i) the side chain loss of C_9H_7N from the δPhe6 ion (*m/z* 620), or (ii) the formation of a δPhe6 fragment ion from the $[(M-H)^- - C_9H_7N]^-$ ion (*m/z* 695), or (iii) a combination of both processes. α and β cleavages provide the sequence of peronein 1.1, except for differentiating between isomeric Leu and Ile, and the

relative position of the last two residues. The formation of the γPhe6 and δPhe6 fragment ions confirm residue six as Phe, and therefore residue seven as Gly.

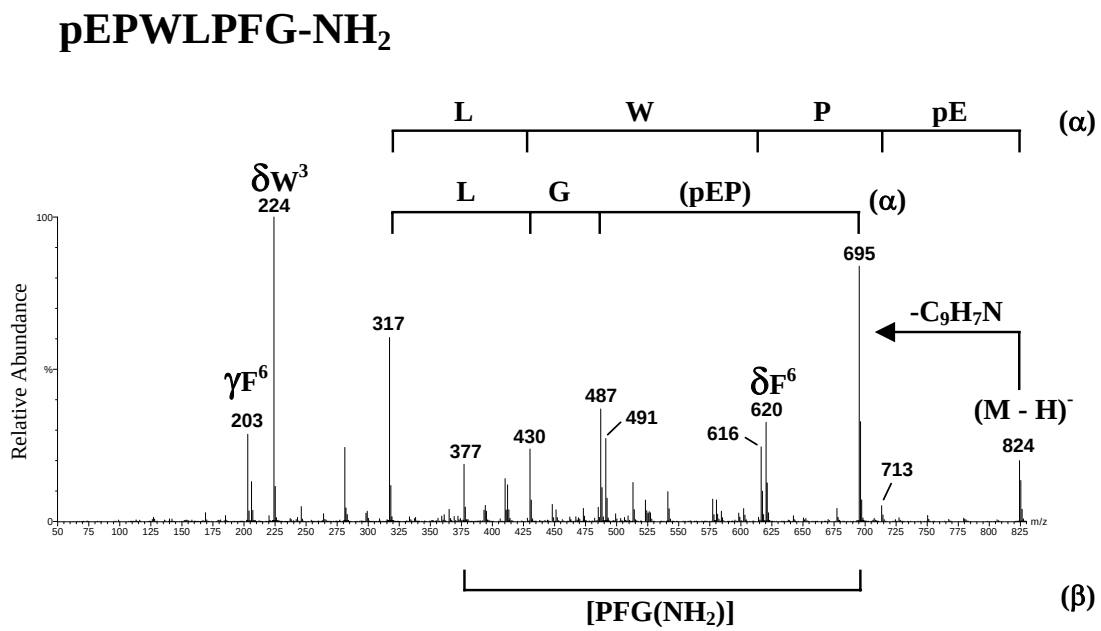


Figure 6.5. Negative ion collision-induced MS/MS data of peronein 1.1. α fragmentations from the $(M-H)^-$ and $[(M-H)^- -C_9H_7N]^-$ ions are shown above the spectrum and the β fragment ion from the $[(M-H)^- -C_9H_7N]^-$ ion shown below the spectrum. Magnification range: *m/z* 400-800 (x4).

Methylation of peronein 1.1 yields a methyl ester which gives an $(MH)^+$ ion at *m/z* 873 in the positive ion spectrum. The increase of 46 Da is consistent with a molecule containing a C-terminal amide (increase of 15 Da) and an N-terminal pyro-Glu residue (increase of 32 Da). This is in agreement with data obtained in the positive and negative ion mass spectra.

Differentiating between isomeric Leu and Ile at residue four was more difficult. The presence of the N-terminal pyro-Glu prevented the utilisation of automated Edman sequencing to achieve this. However, the summer peptide secretions of *Litoria peronii* include two propeptides of peronein 1.1 (see Table 6.2). The Edman degradation data for these two propeptides reveal that residue ten (and therefore residue four of peronein 1.1) is Leu. Peronein 1.1 was synthesised with Leu at position four and the HPLC retention time of the peptide was measured using identical HPLC conditions as for the natural peptide secretions. The matching HPLC retention times of natural and synthetic peronein 1.1 indicate that it has the sequence pEPWLPFG-NH_2.

The sequencing data of the four additional peroneins isolated from *Litoria peronii* are summarised in Table 6.3. The only difference between peronein 1.2 and 1.3 is an Ile and Leu respectively at residue four. The summer secretions of *Litoria peronii* include a propeptide for each of these two peroneins. Peronein 1.2 (pEPWIPFV-NH_2) and 1.3 (pEPWLPFV-NH_2) were commercially synthesised and their HPLC retention times measured. Peronein 1.2 (containing an Ile4) had the shorter retention time and corresponded to HPLC fraction I (Figure 6.3). Peronein 1.3 (containing a Leu4) had a matching retention time to HPLC fraction J (Figure 6.3). No propeptides were isolated in the summer secretions for peronein 1.4 and 1.5. To determine which isomeric residue was present at position four, each peptide was synthesised with either a Leu4 or Ile4. The HPLC retention times of the peptides were measured and compared to the naturally isolated peptide retention times. The sequence of peronein 1.4 and 1.5 was concluded to be pETWLPFV-NH_2 and pEPWLPFR-NH_2 respectively.

Table 6.3. Positive and negative ion mass spectral data for peroneins 1.2-1.5.

Peronein 1.2 pEPWIPFV-NH_2

$(MH)^+$ *m/z* 868.

B ions *m/z* 752, 605, 508, 395 [IPFV-NH_2].

Y + 2 ions *m/z* 660, 74, 316 [(pEP)WI].

Sequence: (pEP)WIPFV-NH_2.

$(M-H)^-$ *m/z* 866.

$[(M-H)^- -C_9H_7N]^-$ *m/z* 737.

α ions from $(M-H)^-$ *m/z* 755, 658, 472, 359 [pEPWI].

α ions from $[(M-H)^- -C_9H_7N]^-$ *m/z* 529, 472, 359 [(pEP)WI].

β ions from $[(M-H)^- -C_9H_7N]^-$ *m/z* 474, 377 [P(FV)-NH_2].

δW^3 *m/z* 224, γF^6 *m/z* 245, δF^6 *m/z* 620, δF^6-C_9H_7N and/or δF^6 from $[(M-H)^- -C_9H_7N]^-$ *m/z* 491.

Sequence: pEPWIPFV-NH_2.

Methylation gives a methyl ester, $(MH)^+$ *m/z* 915, (C-terminal amide and N-terminal pE).

Table 6.3 (Continued).

Peronein 1.3 pEPWLPFV-NH_2

$(MH)^+$ *m/z* 868.

B ions *m/z* 752, 605, 508, 395 [LPFV-NH_2].

Y + 2 ions *m/z* 660, 74, 316 [(pEP)WL].

Sequence: (pEP)WLPFV-NH_2.

$(M-H)^-$ *m/z* 866.

$[(M-H)^- -C_9H_7N]^-$ *m/z* 737.

α ions from $(M-H)^-$ *m/z* 755, 658, 472, 359 [pEPWL].

α ions from $[(M-H)^- -C_9H_7N]^-$ *m/z* 529, 472, 359 [(pEP)WL].

β ions from $[(M-H)^- -C_9H_7N]^-$ *m/z* 474, 377 [P(FV)-NH_2].

δW^3 *m/z* 224, γF^6 *m/z* 245, δF^6 *m/z* 620, δF^6-C_9H_7N and/or δF^6 from $[(M-H)^- -C_9H_7N]^-$ *m/z* 491.

Sequence: pEPWLPFV-NH_2.

Methylation gives a methyl ester, $(MH)^+$ *m/z* 915, (C-terminal amide and N-terminal pE).

Table 6.3 (Continued).

Peronein 1.4 pETWLPFV-NH_2

$(MH)^+$ *m/z* 872.

B ions *m/z* 756, 609, 512, 399 [LPFV-NH_2].

Y + 2 ions *m/z* 660, 474, 361 [(pET)WL].

Sequence: (pET)WLPFV-NH_2.

$(M-H)^-$ *m/z* 870.

$[(M-H)^- -CH_3CHO]^-$ *m/z* 826.

$[(M-H)^- -C_9H_7N]^-$ *m/z* 741.

α ions from $[(M-H)^- -CH_3CHO]^-$ *m/z* 658, 472, 359, 115 [(pET)WL(PF)V-NH_2].

β ions from $[(M-H)^- -CH_3CHO]^-$ *m/z* 710, 466, 167 [(WL)(PF)V-NH_2].

δT^2 *m/z* 127, δW^3 from $[(M-H)^- -CH_3CHO]^-$ *m/z* 184, δF^6 from $[(M-H)^- -CH_3CHO]^-$ *m/z* 580.

Sequence: pETWLPFV-NH_2.

Methylation gives a methyl ester, $(MH)^+$ *m/z* 919, (C-terminal amide and N-terminal pE).

Table 6.3 (Continued).

Peronein 1.5 pEPWLPFR-NH_2

$(MH)^+$ *m/z* 925.

B ions *m/z* 605, 508, 395, 209 [WLP(FR)-NH_2].

Y + 2 ions *m/z* 717, 531, 418 [(pEP)WL].

Sequence: (pEP)WLP(FR)-NH_2.

$(M-H)^-$ *m/z* 923.

$[(M-H)^- -C_9H_7N]^-$ *m/z* 794.

α ions from $(M-H)^-$ *m/z* 812, 715, 529, 416, 319 [pEPWLP].

α ions from $[(M-H)^- -C_9H_7N]^-$ *m/z* 586, 529, 416, 319 [(pEP)WLP].

β ions from $(M-H)^-$ *m/z* 750, 602 [FR-NH_2].

β ion from $[(M-H)^- -C_9H_7N]^-$ *m/z* 474 [(FR)-NH_2].

δW^3 *m/z* 224, γF^6 *m/z* 302, δF^6 *m/z* 620, δF^6-C_9H_7N and/or δF^6 from $[(M-H)^- -C_9H_7N]^-$ *m/z* 491.

Sequence: pEPWLPFR-NH_2.

Methylation gives a methyl ester, $(MH)^+$ *m/z* 972, (C-terminal amide and N-terminal pE).

The next group of peptides to be considered are the peronein propeptides. The four propeptides isolated were minor components in the summer secretions only. Two peronein 1.1 propeptides were identified. The positive ion mass spectrum of pro-peronein 1.1B is shown in Figure 6.6. The B and Y + 2 fragmentation ions are displayed above and below the spectrum respectively. The B and Y + 2 cleavages provide the sequence of the peptide except for the relative orientation of the first two residues and differentiation between isomeric Leu and Ile, and isobaric Lys and Gln.

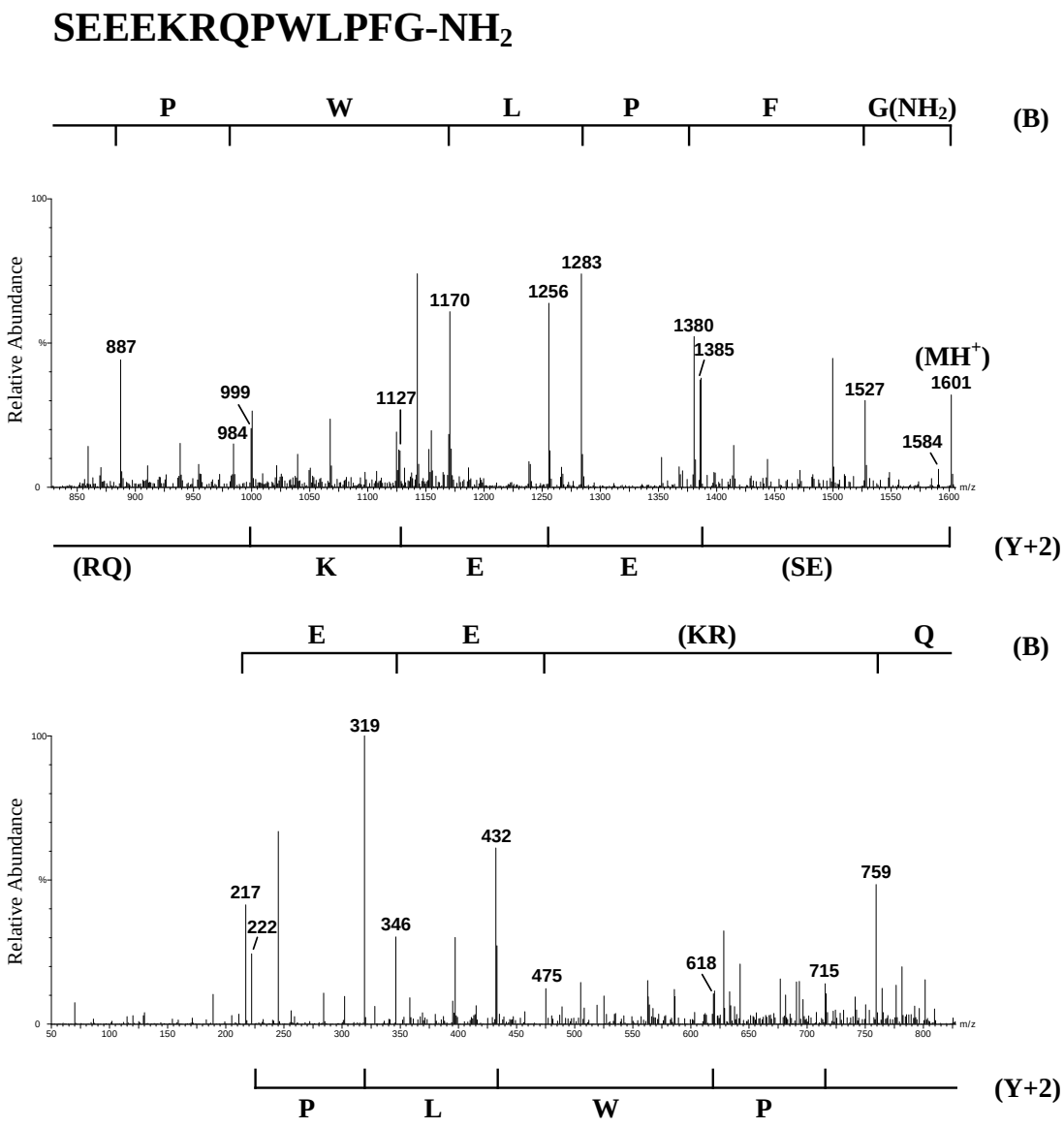


Figure 6.6. Collision-induced MS/MS data of pro-peronein 1.1B. Positive ion fragmentations of the $(MH)^+$ species, B and Y + 2 cleavage ions are shown above and below the spectrum respectively. Magnification ranges: *m/z* 325-800 (x6), *m/z* 850-1200 (x10), *m/z* 1350-1550 (x6).

The negative ion mass spectrum of pro-peronein 1.1B is shown in Figure 6.7. The loss of 17 Da from the $(M-H)^-$ parent ion is consistent with the loss of NH_3 from Gln7. Three α fragmentation ions originating from the $[(M-H)^- -NH_3]^-$ fragment ion are shown above the spectrum with one β fragment ion (also from the $[(M-H)^- -NH_3]^-$ fragment ion)

shown below the spectrum. Two of the three α fragment ions have the same *m/z* as some γ ions.

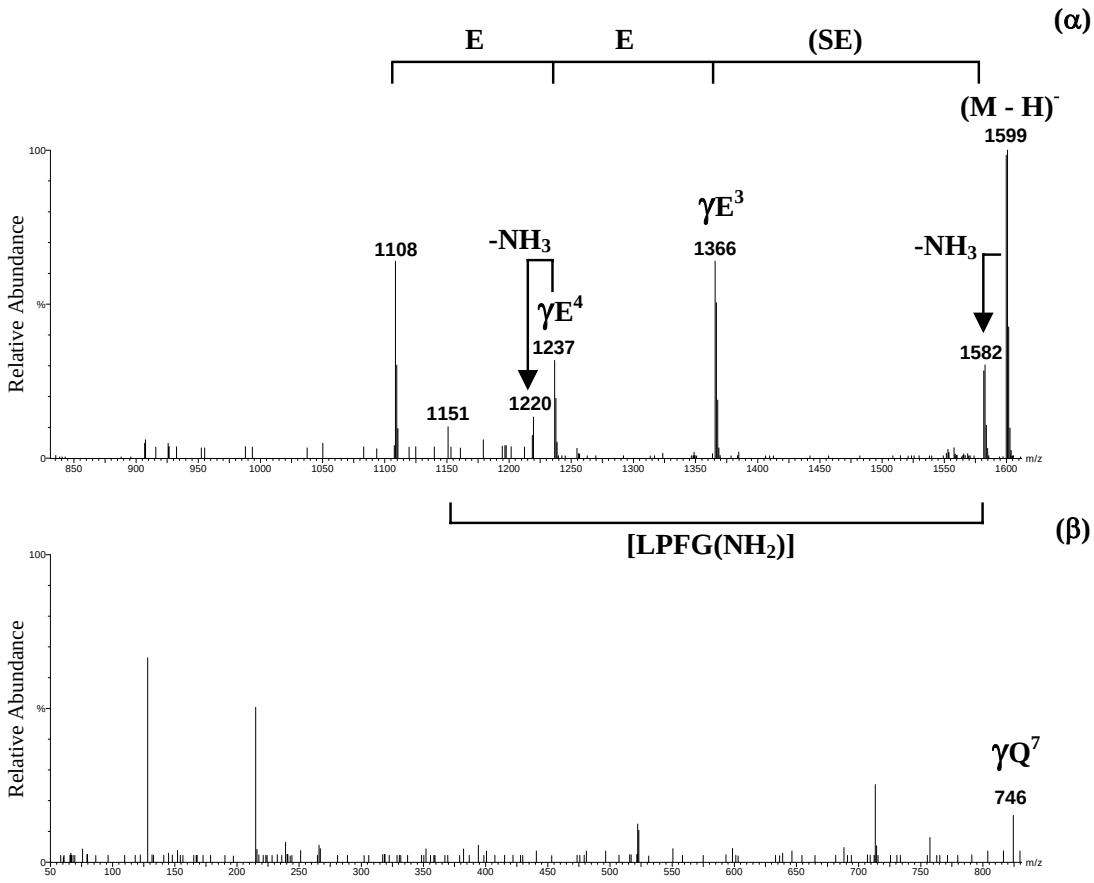


Figure 6.7. Collision-induced MS/MS data of pro-peronein 1.1B. α and β fragmentations are shown from the $[(M-H)^- - NH_3]^-$ fragment ion above and below the spectrum respectively. Magnification ranges: *m/z* 50-900 (x6), *m/z* 900-1225 (x10), *m/z* 1225-1590 (x2).

Lys-C enzymatic digest followed by positive ion mass spectrometric analysis yielded ions at *m/z* 621 and 999. This is consistent with Lys5 and Gln7. This was confirmed by the automated Edman sequencing data, along with Leu10. Insufficient

material was available for methylation, however the loss of 17 Da (the first B ion) from the $(MH)^+$ parent ion in the positive ion spectrum indicates a C-terminal amide. The overall sequence of pro-peronein 1.1B is SEEEKRQPWLPFG-NH_2.

The second propeptide of peronein 1.1 is pro-peronein 1.1A and has the sequence DAQEKRQPWLPFG-NH_2. The positive and negative ion mass spectral data of this peptide and the two other pro-peronein peptides are summarised in Table 6.4. Their sequences were confirmed by automated Edman sequencing.

Table 6.4. Positive and negative ion mass spectral data of pro-peronein 1.1A-1.3A.

Pro-Peronein 1.1A DAQEKRQPWLPFG-NH_2

$(MH)^+$ *m/z* 1570.

B ions *m/z* 1496, 1349, 1252, 1139, 953, 856, 728, 444, 315 [(DAQ)E(KR)QPWL-PFG-NH_2].

Y + 2 ions *m/z* 1455, 1384, 1127, 999, 715, 432, 319, 222 [DA(QE)K(RQ)(PW)LP-(FG)-NH_2].

Sequence: DAQEKRQPWLPFG-NH_2.

$(M-H)^-$ *m/z* 1568.

$[(M-H)^- -NH_3]^-$ *m/z* 1551.

$[(M-H)^- -2NH_3]^-$ *m/z* 1534.

α ions from $[(M-H)^- -2NH_3]^-$ *m/z* 1348, 1220 [(DA)Q].

γQ^3 *m/z* 1365, γE^4 *m/z* 1327, γQ^7 *m/z* 824.

Lys-C digestion yields peptides of $(MH)^+$ *m/z* 590 and 999 (Lys5).

Table 6.4 (Continued).

Pro-Peronein 1.2A DAQEKRQPWIPFV-NH_2

$(MH)^+$ *m/z* 1612.

B ions *m/z* 1496, 1349, 1252, 1139, 953, 856, 728, 444, 315 [(DAQ)E(KR)QPWI-PFV-NH_2].

Y + 2 ions *m/z* 1169, 1041, 757, 474, 361, 264 [(DAQE)K(RQ)(PW)IP(FV)-NH_2].

Sequence: (DAQ)EKRQPWIPFV-NH_2.

$(M-H)^-$ *m/z* 1610.

$[(M-H)^- -NH_3]^-$ *m/z* 1593.

$[(M-H)^- -2NH_3]^-$ *m/z* 1576.

α ions from $[(M-H)^- -2NH_3]^-$ *m/z* 1390, 1262 [(DA)Q].

γQ^3 *m/z* 1407, γE^4 *m/z* 1279, γQ^7 *m/z* 866.

Lys-C digestion yields peptides of $(MH)^+$ *m/z* 590 and 1041 (Lys5).

Pro-Peronein 1.3A DAQEKRQPWLPFV-NH_2

$(MH)^+$ *m/z* 1612.

B ions *m/z* 1496, 1349, 1252, 1139, 953, 856, 728, 444, 315 [(DAQ)E(KR)QPWL-PFV-NH_2].

Y + 2 ions *m/z* 1169, 1041, 757, 474, 361, 264 [(DAQE)K(RQ)(PW)LP(FV)-NH_2].

Sequence: (DAQ)EKRQPWLPFV-NH_2.

$(M-H)^-$ *m/z* 1610.

$[(M-H)^- -NH_3]^-$ *m/z* 1593.

$[(M-H)^- -2NH_3]^-$ *m/z* 1576.

α ions from $[(M-H)^- -2NH_3]^-$ *m/z* 1390, 1262 [(DA)Q].

γQ^3 *m/z* 1407, γE^4 *m/z* 1279, γQ^7 *m/z* 866.

Lys-C digestion yields peptides of $(MH)^+$ *m/z* 590 and 1041 (Lys5).

The pro portion of the peronein propeptides consist of the sequence SEEEKR or DAQEKR. This is similar to the propeptide sequence SEEKR, identified previously by cDNA cloning of various other amphibian peptides.[157] Propeptides are not normally identified in amphibian secretions. It is not known why four propeptides were present in the summer secretions of *Litoria peronii*, but not in the winter secretions. Normally the pro portion is cleaved just prior to secretion, to afford the active peptide. Cleavage C-terminal of Arg [using pro-peronein 1.1A (DAQEKRQPWLPFG-NH_2) as an example] yields a peptide with the sequence QPWLPFG-NH_2. The N-terminal Gln residue cyclises to form a pyro-Glu residue, resulting in the final sequence: pEPWLPFG-NH_2.

Caerulein 1.1 was identified in both the summer and winter secretions of *Litoria peronii*. This peptide has been identified previously from several species of Australian amphibians,[2,11,49,101-107] including *Litoria rothii*.[101] The presence of caerulein 1.1 was confirmed by a combination of positive and negative ion mass spectral data. The negative ion mass spectra of caerulein peptides contain pronounced $(M-H)^-$ and $[(M-H)^- -SO_3]^-$ ions.[104] Under positive ion conditions, the $[MH^+ -SO_3]^+$ ions dominate with little $(MH)^+$ ions.[104] Therefore, the sequence of caerulein 1.1 was confirmed by the positive ion MS/MS spectrum from the $[MH^+ -SO_3]^+$ ion, with the sulfated Tyr confirmed by the $(M-H)^-$ parent anion at *m/z* 1350 (Table 6.5). Although no propeptide of caerulein 1.1 was identified, cDNA cloning of this peptide has revealed that just like the peronein peptides, the N-terminal pyro-Glu originates from a Gln residue.[158]

Table 6.5. A summary of positive ion MS/MS data of caerulein 1.1 from $[MH^+ -SO_3]^+$.

Caerulein 1.1 pEQDT(SO_3)TGWMDF-NH_2	
$(M-H)^-$ *m/z* 1350.	
$[MH^+ -SO_3]^+$ *m/z* 1272.	
B ions from $[MH^+ -SO_3]^+$	*m/z* 1108, 993, 862, 676, 619, 518, 355, 240 [(pEQ)-DYTGWMDF-NH2].
Y + 2 ions from $[MH^+ -SO_3]^+$	*m/z* 1033, 755, 654, 597 [(pEQ)(DT)TG].
Sequence: pEQDY(SO_3)TGWMDF-NH_2.	

The two caerin peptides secreted by *Litoria peronii* have been identified previously from other species of Australian amphibians[2,11,101,106,107,155,156] including *Litoria rothii*.[101] The isolation of caerin 1.1 and 2.1 from *Litoria peronii* was confirmed by examination of the positive ion MS/MS spectra of these two peptides. The data are summarised in Table 6.6.

Table 6.6. Positive ion mass spectral data of caerin 1.1 and 2.1.

Caerin 1.1 GLLSVLGSVAKHVLPHVVPVIAEHL-NH_2	
$(MH)^+$ *m/z* 2583.	
B ions	*m/z* 2453, 2316, 2187, 2116, 2003, 1904, 1807, 1708, 1609, 1472, 1375, 1262, 1163, 1026, 728, 584, 471, 372, 285, 172 [(GL)LSVL(GS)(VAK)HV-LPHVVPVIAEHL-NH_2].
Y + 2 ions	*m/z* 2413, 2300, 2213, 2114, 2001, 1944, 1857, 1758, 1687, 1559, 1422, 1323, 1210, 1113, 976, 877, 778, 582, 469, 397, 269 [(GL)LSVLGSVAKH-VLPHVV(PV)IAE(HL)-NH_2].
Sequence: (GL)LSVLGSVAKHVLPHVVPVIAEHL-NH_2.	

Table 6.6 (Continued).

Caerin 2.1	**GLVSSIGRALGGLLADVVKSKGQPA-OH**
$(MH)^+$ *m/z* 2392.	
B ions	*m/z* 2303, 2206, 2078, 2021, 1893, 1806, 1678, 1579, 1480, 1365, 1294, 1181, 1068, 1011, 954, 841, 770, 557, 444, 357, 270, 171 [(GL)VSSI(GR)-ALGGLLADVVKSKGQPA-OH].
Y + 2 ions	*m/z* 1382, 1212, 1099, 913, 814, 715, 587, 500, 187 [(GL)L(AD)VVKSKG-QPA-OH].
Sequence: (GL)VSSI(GR)ALGGLLADVVKSKGQPA-OH.	

6.2.3 Peptide Activity

Caerulein 1.1 is a neuropeptide involved in both the defence of the animal and with the regulation of dermal physiological action. The wide range of biological activities of caerulein 1.1 include smooth muscle contraction, enhancement of blood circulation, thermoregulation, and this compound is an analgesic thousands of times more potent than morphine.[22-24]

Caerin 1.1 is a wide spectrum antibacterial peptide active mainly against Gram positive organisms.[12,159] NMR studies reveal that the solution structure of this amphipathic peptide contains two α-helical regions separated by a central more flexible hinge region (Figure 6.8).[159] The similar activities of the natural L-caerin 1.1 and synthetic D-caerin 1.1 suggest that it is a membrane active peptide.[160] The antibiotic activities of caerin 1.1 to various bacteria are summarised in Table 6.7. (MIC is the minimum inhibitory concentration required to kill the organism.)

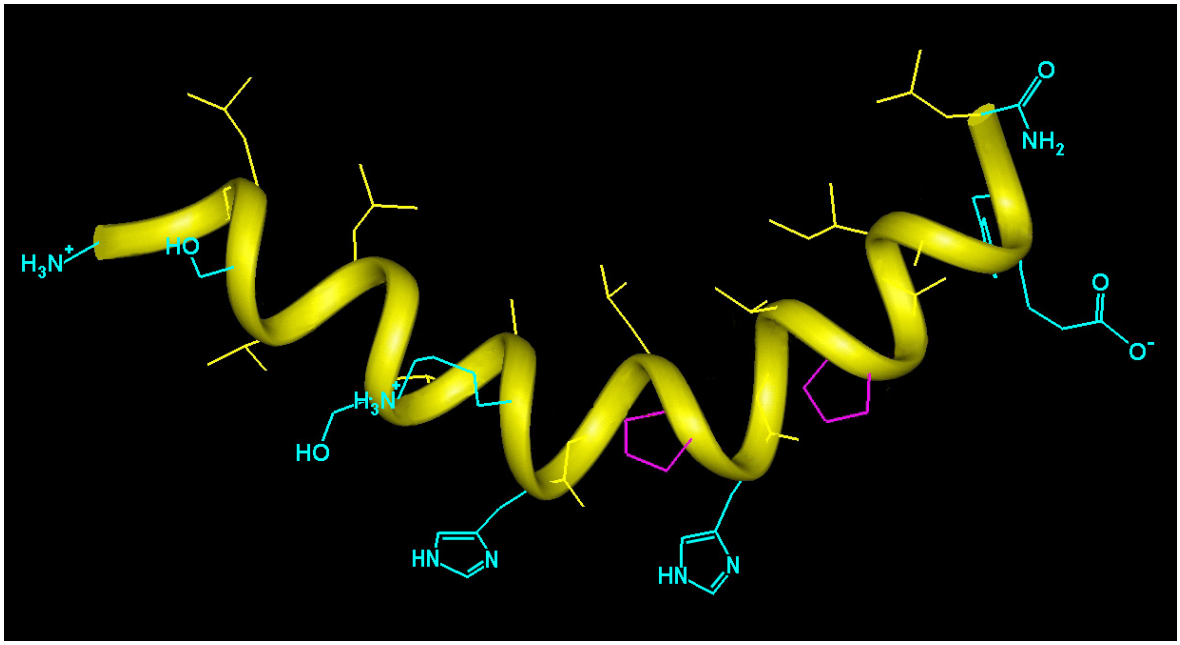


Figure 6.8. The secondary structure of caerin 1.1 as determined by 2D NMR.[159]

Table 6.7. Antibiotic activities of caerin 1.1 and 2.1 against Gram positive (and *Gram negative) organisms.[12,159] No figure indicates MIC > 100 μg mL^{-1}.

Organism	MIC (μg mL^{-1})	
	Caerin 1.1	Caerin 2.1
Bacillus cereus	50	
Leuconostoc lactis	1.5	
Listeria innocua	25	
Micrococcus luteus	12	
Staphylococcus aureus	3	
Staphylococcus epidermidis	12	
Streptococcus uberis	12	
*Escherichia coli**		
*Pasteurella multocida**	25	25

Caerin 1.1 also displays activity against all the major human cancer types in the IC_{50} 10^{-6} M range.[12] (IC_{50} is the concentration of inhibitor required for 50% inhibition.) It also has activity toward viruses such as HIV and *Herpes simplex 1* at MIC 20 and 24 μg mL^{-1} respectively. It can kill nematodes at 10^{-6} M concentrations and is active against the malarial parasite *Plasmodium falciparum* (MIC 10 μg mL^{-1}).[12] Caerin 1.1 can also inhibit the formation of nitric oxide (NO) from neuronal nitric oxide synthase (nNOS) at an IC_{50} concentration of 37 μM. Red blood cells are lysed by caerin 1.1 at > 250 μg mL^{-1}, a concentration greater than those listed above.[12]

Caerin 2.1 displays minimal antibiotic activity[12,159] (see Table 6.7) and shows no anticancer activity against the cell lines tested.[12] Although the solution structure of caerin 2.1 has not yet been determined, the Edmundson projection does not display an amphipathic nature when the peptide is in the α-helical conformation. This may be a possible explanation for the poor antibacterial activity. The main activity of caerin 2.1 is the inhibition of nNOS, the regulator of the cell signalling agent NO.[12]

Full biological testing of the peronein peptides has not been conducted. At the time of writing of this thesis, smooth muscle activity tests have only been completed for two peronein peptides. Peronein 1.1 displayed no smooth muscle activity while peronein 1.3 exhibited contraction at the modest 10^{-6} M concentration.

6.3 Summary and Conclusions

1. The major peptide secretions of *Litoria peronii* are the five peronein peptides. These peptides have not been isolated previously from other Australian amphibians. Initial results on the biological activity of peronein 1.3 suggest it has smooth muscle activity. Peronein 1.1 is not smooth muscle active.

2. Four pro-peroneins were isolated in the summer secretions only.

3. Caerulein 1.1 was present in both the summer and winter peptide secretions. Many Australian amphibians secrete caerulein 1.1. It has a wide range of biological activities, including smooth muscle contraction, enhancement of blood circulation and thermoregulation.

4. Caerin 1.1 and 2.1 were present as minor components only in the winter secretions of *Litoria peronii*. Both peptides are inhibitors of nNOS, the regulator of the neurotransmitter NO. Caerin 1.1 has additional activities which include wide spectrum antibiotic activity, anticancer and antifungal activity.

5. *Litoria peronii* has some peptides also produced by *Litoria rothii* (a related species). These are caerulein 1.1, caerin 1.1 and 2.1.

6.4 Experimental

6.4.1 Collection of Glandular Secretions

Litoria peronii were held by the back legs, the skin moistened with deionised water and the dorsal skin glands stimulated using a bipolar electrode of 21 G platinum attached to a Palmer Student Model electrical stimulator. The electrode was rubbed in a circular manner using 5 V and a pulse rate of 25 Hz. The resulting secretion was washed from the frog with deionised water and the mixture diluted with an equal volume of methanol. The sample was filtered through a Millex HV filter unit (0.45 μm) and then concentrated to approximately 1 mL. Secretions were obtained with the assistance of Assoc. Prof. Michael J. Tyler from the School of Earth and Environmental Sciences, The University of Adelaide. This procedure does not harm the animal and is approved by The University of Adelaide Animal Ethics Committee.

6.4.2 HPLC Separations of Peptide Secretions

Peptide secretions were separated using high performance liquid chromatography (HPLC). The VYDAC C_{18} HPLC column (5 μm, 300 Å, 4.6 x 250 mm) (Separations Group, Hesperia, CA, USA) was first equilibrated to 10% acetonitrile/aqueous 0.1% trifluoroacetic acid. The crude peptide mixture was filtered through a Millex HV filter unit (0.45 μm) and then 250-500 μL injected onto the column. The elution profiles were generated with a linear gradient produced by an ICI DP 800 data station (ICI Instruments Australia, Melbourne, Victoria, Australia) controlling both a Waters 501 and a Waters 510 HPLC pump (Millipore-Waters Chromatographie, Eschborn, Federal Republic of

Germany) increasing the acetonitrile concentration from 10 to 75% over a period of 60 min at a flow rate of 1 mL min^{-1}. The eluent was monitored by ultraviolet absorbance at 214 nm using an ICI LC-1200 variable wavelength detector (ICI Instruments Australia, Melbourne, Victoria, Australia).

6.4.3 Mass Spectrometry

For full details of MS/MS experimental conditions used on the Micromass Q-Tof 2 mass spectrometer, see Chapter 4, section 4.4.2.

6.4.4 Lys-C Enzymatic Digestions

An aqueous ammonium hydrogen carbonate buffer solution (0.1 M, 5 μL, pH 8) and endoprotease Lys-C (1 μL made from 3 units dissolved in 200 μL of water) (*Lysobacter enzymogenes*, Sigma, St. Louis, MO, USA) was added to the lyophilised peptide in an Eppendorf tube. The resulting solution was incubated at 45 °C for 60 min before dilution with water (500 μL) followed by analysis by mass spectrometry.

6.4.5 Automated Edman Sequencing

Automated Edman sequencing was performed by a standard procedure[36] using an Applied Biosystems 492 Procise Sequencer equipped with a 900A data analysis module. Best results were obtained using a disc of immobilon film treated with bioprene in ethanol, onto which the peptide was absorbed from aqueous acetonitrile (90%). The disc was pierced several times with a razor blade to aid the flow of solvent. Automated Edman

sequencing was performed by Chris Cursaro in the Department of Molecular Biosciences, The University of Adelaide.

6.4.6 C-Terminal End Group Determination

Acidified methanol (100 µL) was added to lyophilised peptide and the resulting mixture heated at 45 °C for 30 min. The solvent was removed in a stream of N_2 and the product analysed by mass spectrometry. Acidified methanol was prepared by cooling methanol (858 µL) in dry ice for 5 min and then adding acetyl chloride (142 µL) under N_2. The tube was sealed and cooled again in dry ice for 5 min. The solution was then allowed to warm to 20 °C over a period of 60 min, flushing with N_2 every 15 min. The acidified methanol was stored at -4 °C for a maximum of 5 days.

6.4.7 Synthetic Peptides

The peronein peptides were synthesised by GenScript Corporation (Scotch Plains, NJ, USA) using standard methods.[91]

6.5.8 Smooth Muscle Activity Testing

Immediately prior to the experiment, guinea-pigs (weighing approximately 300 g) were killed by stunning and subsequent decapitation. The ileum was dissected free and cleansed by rinsing with physiological salt solution (KCl 2.7 mM, $CaCl_2$ 1.0 mM, $NaHCO_3$ 13.0 mM, NaH_2PO_4 3.2 mM, NaCl 137 mM, glucose 5.5 mM, pH 7.4) and the mesenteric tissue removed. Segments of approximately 3 cm were cut and suspended in 20 mL organ

baths containing the physiological salt solution at 37 °C and were gassed with 95% O_2 and 5% CO_2. Segments were connected to a tissue holder and an isometric force-displacement transducer. Tension was recorded via a MACLAB v 3.0. Segments were washed thoroughly by replacing the physiological salt solution repeatedly, and were then allowed to equilibrate for a period of 30 min under 2 g of resting tension.

Following the 30 min equilibration period, the tissue bathing solution was replaced repeatedly with fresh drug-free physiological salt solution until a stable baseline tension was achieved. The tension was readjusted to 2 g. All segment preparations were then constricted with acetylcholine (0.01-1 μM). Following washout, acetylcholine (1 μM) was used once again to check that the response was stable. After 5 min washout and achievement of a stable baseline, a cumulative concentration response curve to the peptide (10^{-9}-10^{-6} M) was performed. Smooth muscle activity tests were performed by Vita M. Maselli, Rebecca J. Jackway and Dr. Ian F. Musgrave in the Department of Experimental and Clinical Pharmacology, The University of Adelaide. This procedure is approved by The University of Adelaide Animal Ethics Committee.

References

1. Zasloff, M. *Proc. Natl. Acad. Sci.* **1987**, *84*, 8449-5453.
2. Stone, D. J. M.; Waugh, R. J.; Bowie, J. H.; Wallace, J. C.; Tyler, M. J. *J. Chem. Soc. Perkin Trans. I* **1992**, 3173-3178.
3. Hiramatsu, K.; Aritaka, N.; Hanaki, H.; Kawasaki, S.; Hosoda, Y.; Hori, S.; Fukuchi, Y.; Kobayashi, I. *Lancet* **1997**, *350*, 1670-1673.
4. Hancock, R. E. W. *Lancet* **1997**, *349*, 418-422.
5. Barra, D.; Simmaco, M. *TIBTECH* **1995**, *13*, 205-209.
6. Brinkworth, C. S.; Bowie, J. H.; Tyler, M. J.; Wallace, J. C. *Aust. J. Chem.* **2002**, *55*, 605-610.
7. Barker, J.; Grigg, G.; Tyler, M. J. *A Field Guide to Australian Frogs*; Surrey Beatty & Sons; NSW, Australia, 1995.
8. Nicolas, P.; Mor, A. *Annu. Rev. Microbiol.* **1995**, *49*, 277-304.
9. Giovannini, M. G.; Poulter, L.; Gibson, B. W.; Williams, D. H. *Biochemical J.* **1987**, *243*, 113-120.
10. Gibson, B. W.; Poulter, L.; Williams, D. H.; Maggio, J. E. *J. Biol. Chem.* **1986**, *261*, 5341-5349.
11. Stone, D. J. M.; Waugh, R. J.; Bowie, J. H.; Wallace, J. C.; Tyler, M. J. *J. Chem. Res.* **1993**, *S138*, M910-936.
12. Apponyi, M. A.; Pukala, T. L.; Brinkworth, C. S.; Maselli, V. M.; Bowie, J. H.; Tyler, M. J.; Booker, G. W.; Wallace, J. C.; Carver, J. A.; Separovic, F.; Doyle, J.; Llewellyn, L. E. *Peptides* **2004**, *25*, 1035-1054.
13. Wabnitz, P. A.; Bowie, J. H.; Tyler, M. J.; Wallace, J. C.; Smith, B. P. *Nature* **1999**, *401*, 444-445.

14. Ehrenstein, G.; Lecar, H. *Quart. Rev. Biophys.* **1977**, *10*, 1-34.

15. Epand, R. M.; Shai, Y. C.; Segrest, J. P.; Anantharamaiah, G. M. *Biopolymers* **1995**, *37*, 319-338.

16. Sansom, M. S. P. *Prog. Biophys. Mol. Biol.* **1991**, *55*, 139-235.

17. Wegener, K. L. *Amphibian Peptides: Their Structures and Bioactivities*; PhD Thesis, The University of Adelaide, 2001.

18. Pouny, Y.; Rapaport, D.; Mor, A.; Nicolas, P.; Shai, Y. C. *Biochemistry* **1992**, *31*, 12416-12423.

19. Shai, Y. C. *Biochem. Biophys. Acta* **1999**, *1462*, 55-70.

20. Shai, Y. C.; Oren, Z. *Peptides* **2001**, *22*, 1629-1641.

21. Resnick, N.; Maloy, W.; Guy, H.; Zasloff, M. *Cell* **1991**, *66*, 541-554.

22. Bevins, C. L.; Zasloff, M. *Annu. Rev. Biochem.* **1990**, *59*, 395-414.

23. Erspamer, V. *Bioactive secretions of the amphibian integument. In: Heatwole H, editor. Amphibian biology: the integument. vol. 1.*; Beatty and Sons; NSW, Australia, 1994, pp 178-350.

24. Lazarus, L. H.; Attila, M. *Prog. Neurobiol.* **1993**, *41*, 473-507.

25. Anastasi, A.; Erspamer, V.; Endean, R. *Arch. Biochem. Biophys.* **1968**, *125*, 57-68.

26. Carraway, R. E.; Cochrane, D. E. *J. Biol. Chem.* **1987**, *261*, 15886-15889.

27. Erspamer, V.; Erspamer, G. F.; Mazanti, G.; Endean, R. *Comp. Biochem. Physiol. C.* **1984**, *77*, 99-108.

28. Tyler, M. J.; Stone, D. J. M.; Bowie, J. H. *J. Pharm. Toxicol. Methods* **1992**, *28*, 199-200.

29. Wegener, K. L.; Wabnitz, P. A.; Carver, J. A.; Bowie, J. H.; Chia, B. C. S.; Wallace, J. C.; Tyler, M. J. *Eur. J. Biochem.* **1999**, *265*, 627-637.

30. Blondelle, S. E.; Lohner, K.; Aguilar, M. I. *Biochem. Biophys. Acta (Biomembranes)* **1999**, *1462*, 89-108.

31. Biemann, K.; Martin, S. A. *Mass Spectrom. Rev.* **1987**, *6*, 1-76.

32. Roepstorff, P.; Hojrup, P.; Moller, J. *Biomed. Mass Spectrom.* **1985**, *12*, 181-189.

33. Mueller, D. R.; Eckersley, M.; Richter, W. J. *Org. Mass Spectrom.* **1988**, *23*, 217-222.

34. Bradford, A. M.; Raftery, M. J.; Bowie, J. H.; Tyler, M. J.; Wallace, J. C.; Adams, G. W.; Severini, C. *Aust. J. Chem.* **1996**, *49*, 475-484.

35. Bodanszky, M. *Peptide Chemistry: A Practical Textbook; Springer-Verlag: Berlin* **1988**, 28-31.

36. Hunkapiller, M. W.; Hewick, R. M.; Drewer, W. J.; Hood, L. E. *Methods Enzymology* **1983**, *91*, 399-406.

37. Wabnitz, P. A.; Walters, H.; Tyler, M. J.; Wallace, J. C.; Bowie, J. H. *J. Peptide Res.* **1998**, *52*, 477-483.

38. Wabnitz, P. A.; Bowie, J. H.; Wallace, J. C.; Tyler, M. J. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1724-1732.

39. Werner, H. W.; Venema, A.; Linssen, A. J. *J. Vac. Sci. and Technol.* **1972**, *9*, 216-219.

40. Neumann, S.; Ewald, H. Z. *Physik* **1962**, *169*, 224-238.

41. Johnson, E. G.; Nier, A. O. *Phys. Rev.* **1953**, *91*, 10-17.

42. March, R. E. *Int. J. Mass Spectrom. Ion Processes* **1992**, *118*, 71-135.

43. Cooks, R. G.; Kaiser, R. E. *Acc. Chem. Res.* **1990**, *23*, 213-219.

44. Jonscher, K. R.; Yates, J. R. *Analyt. Biochem.* **1997**, *244*, 1-15.

45. Williams, D. H.; Fleming, I.; Spectroscopic Methods in Organic Chemistry; McGraw-Hill Book Co.: London, England: 1995, pp 9.

46. Smith, R. D.; Loo, J. A.; Loo, R. R.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.* **1991**, *10*, 359-451.

47. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.

48. Williams, D. H.; Bradley, C. V.; Santikarn, S.; Boesen, G. *Biochem. J.* **1982**, *201*, 105-107.

49. Brinkworth, C. S.; Pukala, T. L.; Bowie, J. H.; Tyler, M. J. *Aust. J. Chem.* **2004**, *57*, 693-701.

50. Johnson, R. S.; Martin, S., A.; Biemann, K. *Int. J. Mass Spectrom. Ion Processes* **1988**, *86*, 137-154.

51. Bean, M. F.; Carr, S. A. *Anal. Biochem.* **1992**, *201*, 216-226.

52. Papayannopoulos, I. A. *Mass Spectrom. Rev.* **1995**, *14*, 49-73.

53. Paizs, B.; Suhai, S. *Mass Spectrom. Rev.* **2005**, *24*, 508-548.

54. Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Mass Spectrom. Rev.* **2002**, *21*, 87-107.

55. Bowie, J. H. *Mass Spectrom.* **1984**, *7*, 151-167.

56. Bowie, J. H. *Mass Spectrom.* **1985**, *8*, 161-183.

57. Bowie, J. H. *Mass Spectrom.* **1987**, *9*, 172-195.

58. Voigt, B.; Schmidt, J. *Mass Spectrom.* **1978**, *5*, 44-46.

59. Kulik, W.; Heerma, W. *Biomed. Environ. Mass Spectrom.* **1988**, *15*, 419.

60. Bradley, C. V.; Howe, I.; Beynon, J. H. *J. Chem. Soc. Chem. Commun.* **1980**, 562-564.

61. Bradley, C. V.; Howe, I.; Beynon, J. H. *Biomed. Mass Spectrom.* **1981**, *8*, 85-89.

62. Buko, A. M.; Phillips, L. R.; Fraser, B. A. *Biomed. Mass Spectrom.* **1983**, *10*, 387-393.

63. Kulik, W.; Heerma, W. *Biomed. Environ. Mass Spectrom.* **1989**, *18*, 910-917.

64. Kulik, W.; Heerma, W. *Biol. Mass Spectrom.* **1991**, *20*, 883-558.

65. Waugh, R. J.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 169-173.

66. Bradford, A. M.; Waugh, R. J.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 677-685.

67. Steinborner, S. T.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1243-1247.

68. Brinkworth, C. S.; Dua, S.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 713-721.

69. Brinkworth, C. S.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1339-1351.

70. Boontheung, P.; Brinkworth, C. S.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 287-292.

71. Eckersley, M.; Bowie, J. H.; Hayes, R. N. *Org. Mass Spectrom.* **1989**, *24*, 597-602.

72. Waugh, R. J.; Bowie, J. H.; Gross, M. L. *Aust. J. Chem.* **1993**, *46*, 693-702.

73. Waugh, R. J.; Bowie, J. H.; Gross, M. L.; Vollmer, D. *Int. J. Mass Spectrom. Ion Processes* **1994**, *133*, 165-174.

74. Steinborner, S. T.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 253-258.

75. Maselli, V. M.; Brinkworth, C. S.; Bowie, J. H.; Tyler, M. J. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2155-2161.

76. Brinkworth, C. S.; Dua, S.; McAnoy, A. M.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1965-1973.

77. Wabnitz, P. A.; Bowie, J. H.; Wallace, J. C.; Tyler, M. J. *Aust. J. Chem.* **1999**, *52*, 639-645.

78. Brinkworth, C. S.; Dua, S.; Bowie, J. H. *Eur. J. Mass Spectrom.* **2002**, *8*, 53-66.

79. Foresman, J. B.; Frisch, Æ. *Exploring Chemistry with Electronic Structure Methods*; Gaussian Inc.: Pittsburgh PA, USA, 1993.

80. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902-3909.

81. Hehre, W. J.; Yu, J.; Klunzinger, P. E. *A guide to molecular mechanics and molecular orbital calculations in SPARTAN. Wavefunction Inc.; 18401 von Karman Ave., Suite 370, Irvine, CA 92715, USA* **1997**.

82. Eckersley, M.; Bowie, J. H.; Hayes, R. N. *Int. J. Mass Spectrom. Ion Processes* **1989**, *93*, 199-213.

83. Bradford, A. M.; Waugh, R. J.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1082-1085.

84. Maselli, V. M.; Bowie, J. H.; Tyler, M. J. *unpublished observations.*

85. Adams, J. *Mass Spectrom. Rev.* **1990**, *9*, 141-186.

86. Gross, M. L. *Int. J. Mass Spectrom. Ion Processes* **1992**, *118/119*, 137-165.

87. Cheng, C.; Gross, M. L. *Mass Spectrom. Rev.* **2000**, *19*, 398-420.

88. Jensen, N. J.; Tomer, K. B.; Gross, M. L. *J. Am. Chem. Soc.* **1985**, *107*, 1863-1868.

89. Adams, J.; Gross, M. L. *J. Am. Chem. Soc.* **1989**, *111*, 435-440.

90. Gross, M. L. *Int. J. Mass Spectrom.* **2000**, *200*, 611-624.

91. Maeji, N. J.; Bray, A. M.; Valerio, R. M.; Wang, W. *Pept. Res.* **1995**, *8*, 33-38.

92. Wavefunction Inc.: 18401 Von Karmen Ave., S., Irvine, CA 92715, USA.

93. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Peterson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.;

Stewart, J. P.; Head-Gordon, M.; Gonzales, C.; Pople, J. A. *Gaussian 98; Gaussian Inc.: Pittsburgh PA* **1999**.

94. Badock, V.; Raida, M.; Adermann, K.; Forssmann, W.-G.; Schrader, M. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1952-1956.

95. Solouki, T.; Emmett, M. R.; Guan, S.; Marshall, A. G. *Anal. Chem.* **1997**, *69*, 1163-1168.

96. Hogg, P. J. *Trends Biochem. Sci.* **2003**, *28*, 210-214.

97. Mhatre, R.; Woodard, J.; Zeng, C. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2503-2510.

98. Neitz, S.; Jurgens, M.; Kellmann, M.; Schulz-Knappe, P.; Schrader, M. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1586-1592.

99. Yen, T.-Y.; Joshi, R. K.; Yan, H.; Seto, N. O., L.; Palcic, M. M.; Macher, B. A. *J. Mass Spectrom.* **2000**, *35*, 990-1002.

100. Steinborner, S. T.; Wabnitz, P. A.; Waugh, R. J.; Bowie, J. H.; Gao, C.; Tyler, M. J.; Wallace, J. C. *Aust. J. Chem.* **1996**, *49*, 955-963.

101. Brinkworth, C. S.; Bowie, J. H.; Bilusich, D.; Tyler, M. J. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2716-2724.

102. Steinborner, S. T.; Waugh, R. J.; Bowie, J. H.; Wallace, J. C.; Tyler, M. J.; Ramsay, S. L. *J. Peptide Sci.* **1997**, *3*, 181-185.

103. Rozek, T.; Waugh, R. J.; Steinborner, S. T.; Bowie, J. H.; Tyler, M. J.; Wallace, J. C. *J. Peptide Sci.* **1998**, *4*, 111-115.

104. Wabnitz, P. A.; Bowie, J. H.; Tyler, M. J. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2498-2502.

105. Doyle, J.; Llewellyn, L. E.; Brinkworth, C. S.; Bowie, J. H.; Wegener, K. L.; Rozek, T.; Wabnitz, P. A.; Wallace, J. C.; Tyler, M. J. *Eur. J. Biochem.* **2002**, *269*, 100-109.

106. Wabnitz, P. A.; Bowie, J. H.; Tyler, M. J.; Wallace, J. C.; Smith, B. P. *Eur. J. Biochem.* **2000**, *267*, 269-275.

107. Maclean, M. J.; Brinkworth, C. S.; Bilusich, D.; Bowie, J. H.; Doyle, J. R.; Llewellyn, L. E.; Tyler, M. J. *Toxicon* **2006**, *47*, 664-675.

108. Pukala, T. L.; Bowie, J. H. *unpublished results.*

109. Chrisman, P. A.; Pitteri, S. J.; Hogan, J. M.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1020-1030.

110. Gorman, J. J.; Ferguson, B. L.; Speelman, D.; Mills, J. *Protein Sci.* **1997**, *6*, 1308-1315.

111. Jakubowski, J. A.; Keays, D. A.; Kelley, W. P.; Sandall, D. W.; Bingham, J.-P.; Livett, B. G.; Gayler, K. R.; Sweedler, J. V. *J. Mass Spectrom.* **2004**, *39*, 548-557.

112. Jones, M. D.; Patterson, S. D.; Lu, H. S. *Anal. Chem.* **1998**, *70*, 136-143.

113. Legros, C.; Celerier, M.-L.; Guette, C. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 1317-1323.

114. Pitt, J. J.; Da Silva, E.; Gorman, J. J. *J. Biol. Chem.* **2000**, *275*, 6469-6478.

115. Stephenson, J. L.; Cargile, B. J.; McLuckey, S. A. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2040-2048.

116. Morris, H. R.; Pucci, P. *Biochem. Biophys. Res. Comm.* **1985**, *126*, 1122-1128.

117. Wallis, T. P.; Pitt, J. J.; Gorman, J. J. *Protein Sci.* **2001**, *10*, 2251-2271.

118. Gorman, J. J.; Wallis, T. P.; Pitt, J. J. *Mass Spectrom. Rev.* **2002**, *21*, 183-216.

119. Hugli, T. E. *J. Biol. Chem.* **1973**, *248*, 1712-1718.

120. Martini, F.; Aceto, A.; Sacchetta, P.; Bucciarelli, T.; Dragani, B.; Di Ilio, C. *Eur. J. Biochem.* **1993**, *218*, 845-851.

121. Despeyroux, D.; Bordas-Nagy, J.; Jennings, K. R. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 156-159.

122. Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1990**, *62*, 693-698.

123. Xu, G.; Kiselar, J.; He, Q.; Chance, M. R. *Anal. Chem.* **2005**, *77*, 3029-3037.

124. Thevis, M.; Loo, R. R.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 635-647.

125. Burns, J. A.; Butler, J. C.; Moran, J.; Whitesides, G. M. *J. Org. Chem.* **1991**, *56*, 2648-2650.

126. Lapko, V. N.; Smith, D. L.; Smith, J. B. *J. Mass Spectrom.* **2000**, *35*, 572-575.

127. Glocker, M. O.; Arbogast, B.; Deinzer, M. L. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 638-643.

128. Batista, C. V. F.; Scaloni, A.; Rigden, D. J.; Silva, L. R.; Rodrigues Romero, A.; Dukor, R.; Sebben, A.; Talamo, F.; Bloch, C. *FEBS Lett.* **2001**, *494*, 85-89.

129. Krokhin, O. V.; Cheng, K.; Sousa, S. L.; Ens, W.; Standing, K. G.; Wilkins, J. A. *Biochemistry* **2003**, *42*, 12950-12959.

130. Boja, E. S.; Fales, H. M. *Anal. Chem.* **2001**, *73*, 3576-3582.

131. Huang, R.-H.; Xiang, Y.; Liu, X.-Z.; Zhang, Y.; Hu, Z.; Wang, D.-C. *FEBS Lett.* **2002**, *521*, 87-90.

132. Jones, M. D.; Hunt, J.; Liu, J. L.; Patterson, S. D.; Kohno, T.; Lu, H. S. *Biochemistry* **1997**, *36*, 14914-14923.

133. Basir, Y. J.; Conlon, J. M. *Peptides* **2003**, *24*, 379-383.

134. Bingham, J.-P.; Broxton, N. M.; BLivett, B. G.; Down, J. G.; Jones, A.; Moczydlowski, E. G. *Anal. Biochem.* **2005**, *338*, 48-61.

135. Jakubowski, J. A.; Sweedler, J. V. *Anal. Chem.* **2004**, *76*, 6541-6547.

136. Bures, E. J.; Hui, J. O.; Young, Y.; Chow, D. T.; Katta, V.; Rohde, M. F.; Zeni, L.; Rosenfeld, R. D.; Start, K. L.; Haniu, M. *Biochemistry* **1998**, *37*, 12172-12177.

137. Yen, T.-Y.; Pal, S.; de la Maza, L. M. *Biochemistry* **2005**, *44*, 6250-6256.

138. Schnaible, V.; Wefing, S.; Resemann, A.; Suckau, C.; Bucker, A.; Wolf-Kummeth, S.; Hoffmann, D. *Anal. Chem.* **2002**, *74*, 4980-4988.

139. Qi, J.; Wu, J.; Somkuti, G. A.; Watson, J. T. *Biochemistry* **2001**, *40*, 4531-4538.

140. Wolf, S. M.; Biemann, K. *Int. J. Mass Spectrom. Ion Processes* **1997**, *160*, 317-329.

141. Yen, T.-Y.; Yan, H.; Macher, B. A. *J. Mass Spectrom.* **2002**, *37*, 15-30.

142. Brake, J. M.; Wold, F. *Biochemistry* **1962**, *1*, 386-391.

143. Tashjian Jr., A. H.; Ontjes, D. A.; Munson, P. L. *Biochemistry* **1964**, *3*, 1175-1182.

144. Lundell, N.; Schreitmuller, T. *Anal. Biochem.* **1999**, *266*, 31-47.

145. Sun, Y.; Smith, D. L. *Anal. Biochem.* **1988**, *172*, 130-138.

146. Dai, J.; Zhang, Y.; Wang, J.; Li, X.; Lu, Z.; Cai, Y.; Qian, X. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 1130-1138.

147. Raftery, M. J. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 674-682.

148. Xu, G.; Chance, M. R. *Anal. Chem.* **2005**, *77*, 2437-2449.

149. Crimmins, D., L; Saylor, M.; Rush, J.; Thoma, R. S. *Anal. Biochem.* **1995**, *226*, 355-361.

150. Lebedev, A. T.; Samguina, T.; Bowie, J. H. *unpublished observations.*

151. Morikawa, N.; Hagiwara, K.; Nakajima, T. *Biochem. Biophys. Res. Comm.* **1992**, *189*, 184-190.

152. Simmaco, M.; Mignogna, G.; Barra, D.; Bossa, F. *J. Biol. Chem.* **1994**, *269*, 11956-11961.

153. Simmaco, M.; Mignogna, G.; Barra, D. *Biopolymers* **1998**, *47*, 435-450.

154. Clyne, D. *Australian Frogs*; Periwinkle Books; Vic, Australia, 1969, pp 68-71.

155. Steinborner, S. T.; Bowie, J. H.; Tyler, M. J.; Wallace, J. C. *Aust. J. Chem.* **1997**, *50*, 889-894.

156. Waugh, R. J.; Stone, D. J. M.; Bowie, J. H.; Wallace, J. C.; Tyler, M. J. *J. Chem. Res.* **1993**, *S139*, M937-961.

157. Vanhoye, D.; Bruston, F.; Nicolas, P.; Amiche, M. *Eur. J. Biochem.* **2003**, *270*, 2068-2081.

158. Richter, K.; Egger, R.; Kreil, G. *J. Biol. Chem.* **1986**, *261*, 2676-2680.

159. Bowie, J. H.; Wegener, K. L.; Chia, B. C. S.; Wabnitz, P. A.; Carver, J. A.; Tyler, M. J.; Wallace, J. C. *Protein Pept. Lett.* **1999**, *6*, 259-269.

160. Wong, H.; Bowie, J. H.; Carver, J. A. *Eur. J. Biochem.* **1997**, *247*, 545-557.

161. O'Hair, R. A. J. *J. Mass Spectrom.* **2000**, *35*, 1377-1381.

162. Schlosser, A.; Lehmann, W. D. *J. Mass Spectrom.* **2000**, *35*, 1382-1390.

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

164. Wysocki, V. H.; Tsaprilis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399-1406.

165. Harrison, A. G.; Csizmadia, I. G.; Tang, T.-H. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 427-436.

166. Farrugia, J. M.; Taverner, T.; O'Hair, R. A. J. *Int. J. Mass Spectrom.* **2001**, *209*, 99-112.

167. Farrugia, J. M.; O'Hair, R. A. J. *Int. J. Mass Spectrom.* **2001**, *210/211*, 71-87.

168. Clipston, N. L.; Jai-nhuknan, J.; Cassady, C. J. *Int. J. Mass Spectrom.* **2003**, *222*, 363-381.

169. Ewing, N. P.; Cassady, C. J. *J. Am. Soc. Mass Spectrom.* **2000**, *12*, 105-116.

170. Jai-nhuknan, J.; Cassady, C. J. *Anal. Chem.* **1998**, *70*, 5122-5128.

171. Marzluff, E. M.; Campbell, S.; Rodgers, M. T.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 6947-6948.

172. Marzluff, E. M.; Campbell, S.; Rodgers, M. T.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 7787-7796.

173. Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1-13.

174. Harrsion, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1242-1249.

175. Harrison, A. G.; Siu, K. W. M.; El Aribi, H. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 869-875.

176. Harrison, A. G.; Young, A. B. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 466-456.

177. Harrison, A. G. *J. Mass Spectom.* **2004**, *39*, 136-144.

178. Chass, G. A.; Marai, C. N. J.; Setiadi, D. H.; Csizmadia, I. G.; Harrison, A. G. *J. Mol. Struct. (Theochem)* **2004**, *675*, 149-162.

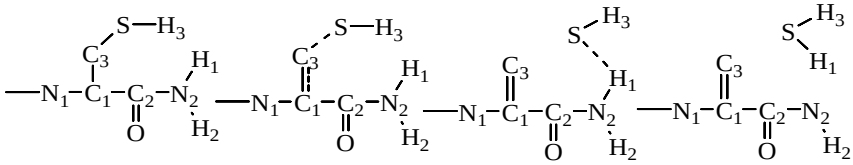
179. Harrison, A. G.; Young, A. B. *J. Mass Spectrom.* **2005**, *40*, 1173-1186.

180. Men, L.; Wang, Y. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 777-784.

181. Grossert, J. S.; Cook, M. C.; White, R. L. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1511-1516.

Appendix

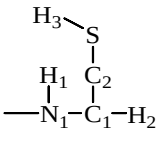
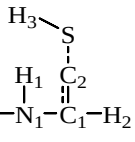
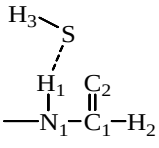
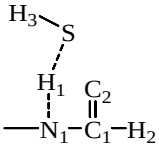
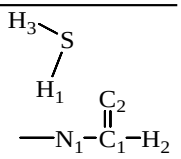
Energies and partial geometries of intermediates and transition states in Figure 3.2.

				
Relative Energies[a]	**79**	**84**	**-14**	**31**
N1C1 (Å)	1.428	1.436	1.428	1.409
C1C2	1.423	1.439	1.516	1.536
C2O	1.274	1.220	1.268	1.285
C2N2	1.242	1.401	1.352	1.322
N2H1	0.995	0.995	1.066	2.627
N2H2	0.996	0.997	0.991	0.997
C1C3	1.461	1.383	1.346	1.348
C3S	1.769	2.331	-	-
SH3	1.316	1.333	1.303	1.324
SH1	-	-	1.835	1.316
N1C1C2 (°)	120.0	122.5	115.3	114.2
C1C2O	124.3	125.8	119.1	114.7
C1C2N2	120.0	117.0	120.1	116.8
C2N2H1	113.8	111.6	124.2	139.2
C2N2H2	112.6	117.2	116.4	111.3
C2C1C3	122.7	118.9	122.0	121.1
C1C3S	113.0	117.3	-	-
C3SH3	101.7	94.3	-	-
H3SH1	-	-	104.4	99.5
SH1N2	-	-	175.9	190.8
N1C1C2O (°)	-0.4	4.8	-15.0	-46.2
N1C1C3N2	178.4	-176.8	163.7	133.4
C1C2N2H1	32.9	24.1	11.7	-32.6
C1C2N2H2	162.0	158.2	176.0	179.0
N1C1C3S	98.0	81.0	-	-
C1C3SH3	-20.9	59.3	-	-

HF/6-31G(d)//AM1 kJ mol^{-1}

a) Relative energies with respect to $CH_3NHCH(CH_2SH)CONH^-$ -737.9560961 hartrees (0 kJ mol^{-1}).

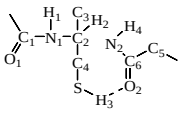
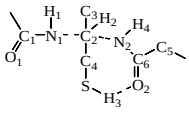
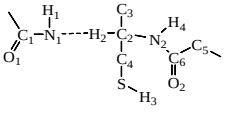
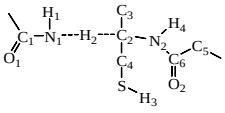
Energies and partial geometries of intermediates and transition states in Figure 3.6.

					
Relative Energies[a]	**278**	**281**	**85**	**103**	**89**
N1C1	1.394	1.443	1.385	1.377	13.700
N1H1	0.996	1.003	1.181	1.387	2.696
C1H2	1.092	1.085	1.113	1.116	1.121
C1C2	1.427	1.409	1.347	1.350	1.355
C2S	1.812	2.128	-	-	-
SH3	1.312	1.332	1.305	1.308	1.330
SH1	-	-	1.660	1.500	1.316
N1C1H2	115.3	112.3	116.0	116.9	118.3
N1C1C2	124.0	115.7	125.4	125.6	125.6
C1N1H1	118.3	120.4	118.6	119.8	-
C1C2S	114.2	118.7	-	-	-
C2SH3	101.5	95.8	-	-	-
H3SH1	-	-	101.3	99.9	99.4
N1C1C2S	87.0	65.7	-	-	-
C1C2SH3	0.5	51.9	-	-	-
C2C1N1H1	-72.7	-60.4	0.0	0.0	0.0
SC2C1H2	-86.5	-68.5	-	-	-
H2C1N1H1	101.1	75.3	180.0	180.0	180.0
H3SH1N1	-	-	0.1	0.1	0.0
SH1N1C1	-	-	179.8	179.8	180.0

HF/6-31G(d)//AM1 kJ mol^{-1}

a) Relative energies with respect to $HCONHCH(CH_2SH)CO_2^-$ -831.5518252 hartrees (0 kJ mol^{-1}).

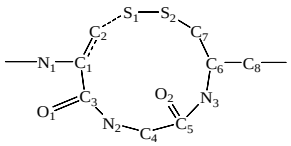
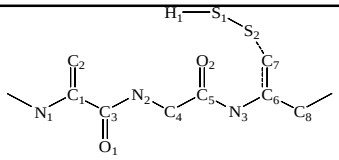
Energies and partial geometries of intermediates and transition states in Figure 3.8.

				
Relative Energies[a]	**-60**	**132**	**-84**	**-4**
C1N1 (Å)	1.378	1.343	1.317	1.342
C1O1	1.249	1.267	1.303	1.276
N1H1	0.993	0.989	0.998	0.994
N1H2	-	-	2.877	1.522
N1C2	1.447	2.033	-	-
C2H2	1.135	1.110	1.140	1.271
C2C4	1.535	1.501	1.535	1.520
C4S	1.725	1.728	1.725	1.729
SH3	1.327	1.329	1.341	1.320
C2C3	1.532	1.503	1.526	1.501
C2N2	3.778	1.998	1.444	1.438
N2C6	1.323	1.338	1.371	1.367
N2H4	0.995	0.988	1.010	0.996
C6C5	1.558	1.548	1.539	1.542
C6O2	1.279	1.267	1.254	1.252
H3O2	2.285	-	-	-
N1C1O1 (°)	122.9	122.8	126.2	125.6
C1N1H1	117.8	115.5	112.0	112.6
C1N1C2	123.1	122.2	-	-
C1N1H2	-	-	80.7	116.2
N1C2C4	112.8	97.0	-	-
N1H2C2	-	-	155.7	160.2
C3C2H2	106.4	121.0	107.1	103.6
C2C4S	113.6	114.3	114.8	114.8
C4SH3	100.5	102.2	103.3	102.4
C4C2C3	106.4	117.9	107.8	110.5
C3C2N2	51.1	90.0	113.2	114.2
C2N2C6	100.1	121.2	123.4	126.6
C6N2H4	112.9	116.7	120.0	118.8
N2C6C5	121.4	120.7	118.8	118.1
N2C6O2	124.4	122.9	122.5	123.5
SH3O2	127.3	-	-	-
C6O2H3	150.2	-	-	-
H1N1C1O1 (°)	-167.4	174.0	0.8	1.6
C1N1C2C3	-102.7	-67.7	-	-
C1N1H2C2	-	-	21.4	34.8
N1H2C2C3	-	-	-154.5	-124.2
C3C2C4S	155.2	175.4	-168.7	-163.9
C2C4SH3	-99.0	-34.4	8.1	33.9
C3C2N2C6	-164.8	136.2	-56.5	-71.9
C2N2C6C5	-164.6	-159.1	-171.9	-177.3

HF/6-31G(d)//AM1 kJ mol^{-1}

a) Relative energies with respect to $CH_3CONHCH(CH_2SH)CHO$ + $^{-}HNCOCH_2NHCH_3$ -1097.6584565 hartrees (0 kJ mol^{-1}).

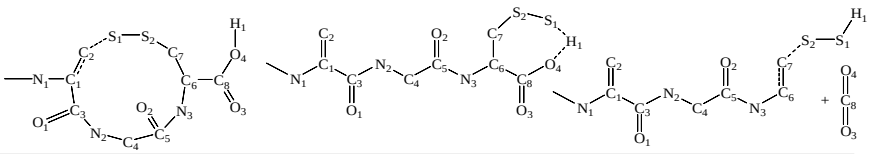
Energies and partial geometries of transition states in Figure 4.5.

		
Relative energies[a]	**139**	**171**
N1C1 (Å)	1.417	1.405
C1C2	1.361	1.346
C2S1	2.539	-
S1S2	1.898	1.899
C1C3	1.495	1.518
C3O1	1.258	1.250
C3N2	1.395	1.378
N2C4	1.434	1.439
C4C5	1.541	1.546
C5O2	1.248	1.251
C5N3	1.388	1.369
N3C6	1.445	1.414
C6C7	1.535	1.355
C7S2	1.739	2.599
C6C8	1.529	1.492
S1H1	-	1.321
N1C1C2 (°)	123.1	125.9
C1C2S1	106.7	-
C2S1S2	119.6	-
N1C1C3	114.5	113.4
C1C3O1	122.0	120.5
C1C3N2	117.2	117.0
C3N2C4	119.3	121.1
N2C4C5	112.8	112.9
C4C5O2	120.1	120.2
C4C5N3	117.3	115.2
C5N3C6	122.9	125.4
N3C6C7	115.1	125.4
N3C6C8	112.8	114.5
C6C7S2	112.1	102.1
C7S2S1	107.4	118.6
S2S1H1	-	101.3
N1C1C2S1 (°)	-117.4	-
C1C2S1S2	52.1	-
C2S1S2C7	-126.5	-
N1C1C3O1	-14.4	-37.6
N1C1C3N2	164.1	137.8
C1C3N2C4	-161.7	167.4
C3N2C4C5	94.3	-123.1
N2C4C5O2	126.3	-7.4
N2C4C5N3	-55.3	175.4
C4C5N3C6	148.0	-178.0
C5N3C6C8	77.6	157.5
C5N3C6C7	-48.2	-26.2
N3C6C7S2	-76.3	-112.3
C6C7S2S1	101.3	-149.1
C7S2S1H1	-	-120.0

HF/6-31G(d)//AM1 kJ mol^{-1}

a) Relative energies with respect to HCO-CGC-O$^-$ -1680.5358238 hartrees (0 kJ mol^{-1}).

Energies and partial geometries of transition states in Figure 4.7.

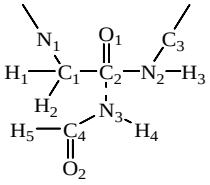
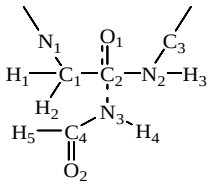
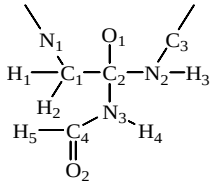


Relative energies[a]	**139**	**72**	**169**
N1C1 (Å)	1.417	1.405	1.405
C1C2	1.361	1.346	1.346
C2S1	2.539	-	-
S1S2	1.898	1.919	1.894
C1C3	1.495	1.518	1.519
C3O1	1.258	1.250	1.250
C3N2	1.395	1.378	1.378
N2C4	1.434	1.439	1.440
C4C5	1.541	1.544	1.544
C5O2	1.248	1.250	1.254
C5N3	1.388	1.375	1.368
N3C6	1.445	1.454	1.412
C6C7	1.535	1.537	1.347
C7S2	1.739	1.736	2.595
C6C8	1.529	1.556	-
C8O3	1.234	1.253	1.190
C8O4	1.357	1.290	1.190
O4H1	0.973	1.375	-
S1H1	-	1.500	1.324
N1C1C2 (°)	123.1	125.9	126.0
C1C2S1	106.7	-	-
C2S1S2	119.6	-	-
N1C1C3	114.5	113.4	113.4
C1C3O1	122.0	120.4	120.3
C1C3N2	117.2	116.9	116.9
C3N2C4	119.3	121.2	121.2
N2C4C5	112.8	113.1	113.0
C4C5O2	120.1	120.3	120.2
C4C5N3	117.3	115.7	115.9
C5N3C6	122.9	121.8	122.5
N3C6C7	115.1	111.8	122.4
N3C6C8	112.8	111.0	-
C6C7S2	112.1	113.2	109.6
C7S2S1	107.4	105.3	121.0
C6C8O3	124.5	120.7	-
C6C8O4	121.5	118.6	-
C8O4H1	110.2	116.2	-
S2S1H1	-	100.1	101.9
S1H1O4	-	170.4	-
N1C1C2S1 (°)	-117.4	-	-
C1C2S1S2	52.1	-	-
C2S1S2C7	-126.5	-	-
N1C1C3O1	-14.4	-37.9	-38.4
N1C1C3N2	164.1	137.5	136.8
C1C3N2C4	-161.7	167.6	166.8
C3N2C4C5	94.3	-124.0	-121.9
N2C4C5O2	126.3	-10.8	-8.5
N2C4C5N3	-55.3	166.7	169.5
C4C5N3C6	148.0	163.9	166.6
C5N3C6C8	77.6	-117.8	-
C5N3C6C7	-48.2	119.8	-179.9
N3C6C7S2	-76.3	-132.7	67.9
C6C7S2S1	101.3	-69.7	95.0
N3C6C8O3	109.9	-6.3	-
N3C6C8O4	-73.5	172.8	-
C6C8O4H1	6.5	11.2	-
C7S2S1H1	-	20.7	-88.2

HF/6-31G(d)//AM1 kJ mol^{-1}

a) Relative energies with respect to HCO-CGC-O$^-$ -1680.5358238 hartrees (0 kJ mol^{-1}).

Energies and partial geometries of key species in Figure 5.7.

			
Relative Energies[a]	**-67**	**-26**	**-50**
N1C1 (Å)	1.436	1.440	1.444
C1H1	1.077	1.082	1.087
C1H2	1.089	1.087	1.084
C1C2	1.518	1.529	1.550
C2O1	1.201	1.231	1.288
C2N2	1.379	1.428	1.485
N2H3	0.997	0.999	1.001
N2C3	1.447	1.443	1.443
C2N3	2.500	1.910	1.528
N3H4	1.009	1.003	0.997
N3C4	1.310	1.315	1.325
C4O2	1.229	1.220	1.209
C4H5	1.115	1.106	1.096
N1C1C2 (°)	112.4	111.1	108.4
C1C2O1	121.3	117.5	112.0
C1C2N2	115.2	110.4	105.4
C1C2N3	86.4	98.3	108.4
C2C1H1	110.8	112.4	113.6
C2C1H2	106.3	106.4	106.5
C2N2C3	115.0	115.7	114.2
C2N2H3	110.9	111.4	111.7
C2N3C4	134.9	134.2	130.7
N3C4O2	131.7	130.2	127.3
N3C4H5	111.7	112.1	113.0
C4N3H4	107.6	112.9	117.2
N1C1C2N2 (°)	169.8	-179.8	-171.3
N1C1C2O1	-25.5	-39.2	-47.1
N1C1C2N3	82.4	73.5	72.1
C1C2N2C3	-173.0	175.4	163.6
C1C2N2H3	-45.1	-56.0	-69.9
C1C2N3C4	82.8	78.5	60.2
N2C2C1H1	46.5	56.9	66.7
N2C2C1H2	-70.7	-61.0	-52.3
C2N3C4O2	156.3	166.8	179.6
C2N3C4H5	-23.4	-13.5	-0.2
O2C4N3H4	-0.5	1.9	-1.0

HF/6-31G(d)//HF/6-31G(d) kJ mol^{-1}

a) Relative energies with respect to $HCONH^-$ + $OCHNHCH_2CONHCH_2CHO$ -695.8236671 hartrees (0 kJ mol^{-1})

Publications

"Determination of disulfide functionality in underivatised peptides using negative ion mass spectrometry. An aid to structure determination."

Bilusich, D.; Bowie, J. H. *Curr. Anal. Chem.* **2006**, *accepted*.

"New caerin antibiotic peptides from the skin secretion of the Dainty Green Tree Frog *Litoria gracilenta*. Identification using positive and negative ion electrospray mass spectrometry."

Maclean, M. J.; Brinkworth, C. S.; Bilusich, D.; Bowie, J.H.; Doyle, J. R.; Llewellyn, L. E.; Tyler, M. J. *Toxicon* **2006**, *47*, 664-675.

"Host-defence skin peptides of the Australian Streambank Froglet *Crinia riparia*: isolation and sequence determination by positive and negative ion electrospray mass spectrometry."

Maselli, V. M.; Bilusich, D.; Bowie, J. H.; Tyler, M. J. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 797-803.

"Direct identification of intramolecular disulfide links in peptides using negative ion electrospray mass spectra of underivatised peptides. A joint experimental and theoretical study."

Bilusich, D.; Maselli, V. M.; Brinkworth, C. S.; Samguina, T.; Lebedev, A. T.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 3063-3074.

"The rothein peptides from the skin secretion of Roth's tree frog *Litoria rothii*. Sequence determination using positive and negative ion electrospray mass spectrometry."

Brinkworth, C. S.; Bowie, J. H.; Bilusich, D.; Tyler, M. J. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2716-2724.

"The gas phase cyclisation of deprotonated *N*-aryl-2-cyano-2-diazoacetamides."

Lobodin, V. V.; Morzherin, Y. Y.; Blumenthal, T.; Bilusich, D.; Ovcharenko, V. V.; Bowie, J. H.; Lebedev, A. T. *Arkivoc* **2005**, 189-198.

"The unusual loss of an internal Val residue from the $(M-H)^-$ parent anions of the antimicrobial peptide citropin 1.1 and synthetically modified analogues. Fragmentations which require a specific conformation of the decomposing anion."

Brinkworth, C. S.; Bilusich, D.; Bowie, J. H. *Int. J. Mass Spectrom.* **2004**, *236*, 43-53.

"Negative ion mass spectra of Cys-containing peptides. The characteristic Cys γ backbone cleavage: a joint experimental and theoretical study."

Bilusich, D.; Brinkworth, C. S.; Bowie, J. H. *Rapid Commun. Mass Spectrom*. **2004**, *18*, 544-552.

"The fragmentations of $[M-H]^-$ anions derived from underivatised peptides. The side-chain loss of H_2S from Cys. A joint experimental and theoretical study."

Bilusich, D.; Brinkworth, C. S.; McAnoy, A. M.; Bowie, J. H. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2488-2494.

Maclean, M.J., Brinkworth, C.S., Bilusich, D., Bowie , J.H., Doyle, J.R., Llewellyn, L.E. and Tyler, M.J. (2006) New caerin antibiotic peptides from the skin secretion of the Dainty Green Tree Frog *Litoria gracilenta*. Identification using positive and negative ion electrospray mass spectrometry.
Toxicon, v.47 (6) pp. 664-675, May 2006

NOTE: This publication is included in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://dx.doi.org/10.1016/j.toxicon.2006.01.019

Maselli, V.M., Bilusich, D., Bowie, J.H. and Tyler, M.J. (2006) Host-defence skin peptides of the Australian Streambank Froglet *Crinia riparia*: isolation and sequence determination by positive and negative ion electrospray mass spectrometry.
Rapid Communications in Mass Spectrometry, v.20 (5) pp. 797-803, March 2006

NOTE: This publication is included in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://dx.doi.org/10.1002/rcm.2360

Bilusich, D., Maselli, V.M., Brinkworth, C.S., Samguina, T., Lebedev, A.T. and Bowie, J.H. (2005) Direct identification of intramolecular disulfide links in peptides using negative ion electrospray mass spectra of underivatised peptides. A joint experimental theoretical study.
Rapid Communications in Mass Spectrometry, v. 19 (21) pp. 3063-3074, November 2005

NOTE: This publication is included in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://dx.doi.org/10.1002/rcm.2149

Brinkworth, C.S., Bowie, J.H., Bilusich, D. and Tyler, M.J. (2005) The rothein peptides from the skin secretion of Roth's tree frog *Litoria rothii.* Sequence determination using positive and negative ion electrospray mass spectrometry. *Rapid Communications in Mass Spectrometry, v. 19 (18) pp. 2716-2724, September 2005*

NOTE: This publication is included in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://dx.doi.org/10.1002/rcm.2098

Lobodin, V.V., Morzherin, Y.Y., Blumenthal, T., Bilusich, D., Ovcharenko, V.V., Bowie, J.H. and Lebedev, A.T. (2005) The gas phase cyclization of deprotonated N-aryl-2-cyano-2-diazoacetamides.
ARKIKOV, 2005 (iv) pp. 189-198

NOTE: This publication is included in the print copy of the thesis held in the University of Adelaide Library.

The complete conference proceedings are available free at:

http://www.arkat-usa.org/home.aspx?VIEW=BROWSE&MENU=ARKIVOC

Brinkworth, C.S., Bilusich, D., Bowie , J.H. (2004) The unusual loss of an internal Val residue from the (M – H)- parent anions of the antimicrobial peptide citropin 1.1 and synthetically modified analogues. Fragmentations which require a specific conformation of the decomposing anion.
International Journal of Mass Spectrometry, v.236 (1/3) pp. 43-53, August 2004

NOTE: This publication is included in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://dx.doi.org/10.1016/j.ijms.2004.05.011

Bilusich, D., Brinkworth, C.S. and Bowie, J.H. (2004) Negative ion mass spectra of Cys-containing peptides. The characteristic Cys γ backbone cleavage: a joint experimental and theoretical study.
Rapid Communications in Mass Spectrometry, v. 18 (5) pp. 544-552, March 2004

NOTE: This publication is included in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://dx.doi.org/10.1002/rcm.1360

Bilusich, D., Brinkworth, C.S., McAnoy, A.M. and Bowie, J.H. (2003) The fragmentations of [M – H]- anions derived from underivatised peptides. The side-chain loss of H2S from Cys. A joint experimental and theoretical study.
Rapid Communications in Mass Spectrometry, v. 17 (22) pp. 2488-2494, November 2003

NOTE: This publication is included in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

http://dx.doi.org/10.1002/rcm.1224