The Characterisation and Fragmentation of Peptides by Mass Spectrometry

---

A Thesis submitted for the Degree of

Doctor of Philosophy

by

Adrian Morley Bradford B.Sc (Hons)

Department of Chemistry

University of Adelaide

1996

CONTENTS

STATEMENT

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

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

Signed:

Adrian Morley Bradford

Date: 22/8/96

ACKNOWLEDGMENTS

I am grateful for the help which I have received from a great number of people over the past three years, not only in producing this thesis, but for those who have encouraged and inspired me. Giving recognition to every individual would inevitably double the thickness of this thesis. However, there are those who deserve distinct acknowledgment. In particular, I would like to thank my supervisor Professor John Bowie for his advice, expertise and encouragement throughout my postgraduate studies, and for the opportunity to branch into the exciting and challenging field of biological mass spectrometry.

I would also like to extend my gratitude to the members of the research group who have given assistance and ideas. Some of whom have parted company in the preceding years and include: Dr Mark Raftery, whose assistance at the start of my studies was invaluable; and Dr Russel Waugh, who provided many tips and gave sound advice. Thanks must also go to Tom Blumenthal, Lee Paltridge and Simon Steinborner who have kept the ZAB in running order, especially after it had been moved.

This thesis would not have been possible without the expertise given by Associate Professor Michael Tyler and Dr Margaret Davies from the Department of Zoology at the University of Adelaide. Their ability to correctly identify the Uperoleia species and collect the secretion is duly recognised.

I have been fortunate to share a laboratory with individuals who, when necessary, have formed a great team in keeping the laboratory running and clean. It has been a pleasure to work with the past and current members of Lab

2, particularly with Natalie Williamson and Timothy Bubner who have also provided great company during the preceding three years.

The compilation of this thesis has been an enormous task. It remains for me to thank: Anne Martin, Simon Steinborner and Dr Suresh Dua who proof read the thesis; the Education and Technology Unit, University of Adelaide for allowing me to use their equipment to produce the high quality spectra; and Dr Simon Pyke and Oscar Ferriero for their assistance in scanning and formatting documents. I would also like to thank the Australian Research Council for providing a scholarship which has enabled me to pursue this research.

Thanks must also go to my flatmates Sherri, Christine, Andrew and Mary-Ann who have, over the years, made me feel at home away from home. At times it can be difficult being a country student but I have been privileged to have excellent flatmates who have been patient, understanding and have provided excellent company to live with.

Lastly, I would like to thank my family to whom I dedicate this thesis. Despite the many miles of deserts which separate us, they have always been close at heart. I especially would like to thank my mother and father who provided the necessary support of love and understanding during my education, and for providing me the opportunity to travel to England so that I could visit my late Grandmother.

LEGEND TO FIGURES

LEGEND TO TABLES

ABSTRACT

The fragmentation behaviour of peptides in the gas phase is investigated in this thesis, using the technique of fast atom bombardment (FAB) mass spectrometry (MS) in both the positive and negative modes of ionisation. The established knowledge of the fragmentation behaviour of protonated gas phase peptides is applied in the determination of the sequences of biologically active peptides. The behaviour of deprotonated gas phase peptides is being realised; this thesis reports the fragmentations of deprotonated synthetic tetrapeptides as part of a continuing study of anionic peptide chemistry in the gas phase. Observations have been rationalised by mechanisms.

The structure determination of the amino acid sequences of twenty peptides from the Australian frogs, *Uperoleia inundata* and *Uperoleia mjobergi* are reported. These peptides have been named uperins and were extracted from the dermal glands using the ecologically sound method of surface electrical stimulation. Purification by reverse-phase high performance liquid chromatography (HPLC) afforded workable amounts of material. The sequences were elucidated using FAB MS and CA MS/MS in conjunction with manual Edman degradations, enzymatic hydrolysis (using site-specific enzymes) and esterification procedures. The proposed structures were confirmed by automated solid-phase protein sequencing. The biological activity of a number of uperins have been investigated using synthetic analogues; the details of which are reported. Some of the uperins show remarkable antimicrobial activity. A neurotransmitter of the tachykinin category of peptides has been isolated and the sequence has been determined; this peptide has been named uperin 1.1 and is structurally similar to other amphibian tachykinins.

The anionic investigation into peptide sequencing by mass spectrometry has been extended from di- and tripeptides to tetrapeptides. The fragmentations of a number of tetrapeptides have been studied by CA MS/MS. Of notable significance is that the $[M - H]^-$ ions derived from tetrapeptides generally show two different collision-induced backbone cleavages which permit the determination of the amino acid sequence of the peptide. The first of these involves the formation of the carboxylate anions of either constituent amino acids or fragment peptides. In the second, amino acids or fragment peptides are eliminated as neutrals. A nomenclature has been introduced to distinguish these two processes. There are a number of residues which undergo characteristic side-chain fragmentations irrespective of their location in the peptide sequence. Also, there are some residues which promote pronounced fragmentation at the C-terminal position: if situated at this site, such fragmentations may occur to the exclusion of the backbone cleavages.

The data obtained from these negative ion cleavages are analytically useful and provide, at the very least, complementary information to that given by positive ions.

To my family.....

Chapter 1 INTRODUCTION TO MASS SPECTROMETRY

"I feel sure that there are many problems in Chemistry which could be solved with far greater ease by this than by any other method. The method is surprisingly sensitive - more so even than that of spectrum analysis - requires an infinitesimal amount of material and does not require this to be specially purified...."

J. J. Thomson[1]

1.1 Historical Aspects of Mass Spectrometry

A mass spectrometer is an instrument which produces a beam of ions from a substance being investigated, sorts these ions into a spectrum according to their mass to charge ratios, and records the relative abundance of each species of ion present. The basic principle of mass spectrometry, the separation and registration of atomic masses, was demonstrated many years ago. The origins of mass spectrometry can be traced back to 1886 when Goldstein detected positive rays in a glow discharge tube.[2] In 1898, Wein demonstrated that these rays were beams of ions by deflecting them in electric and magnetic fields.[3] Later, Thomson was able to show the existence of neon isotopes by the dissociation of the ions in a flight tube; a discovery that permitted the separation of ions according to their mass to charge ratio.[4] The first instruments of more sophisticated design were built by Dempster and Aston in 1918.[5, 6] Aston's instrument was particularly useful for accurate mass measurements because all the ions were focussed in one plane, which was occupied by a photographic plate; whereas Dempster's instrument, although less suitable for this purpose, permitted a more accurate measurement of

abundances. Refined over the years, both designs have been used to measure the masses accurately and the relative abundances of a great number of isotopes of the elements. Important advances in the art of electronics led to the construction of reliable, accurate, and user-friendly mass spectrometers. The reproducibility of the mass spectra of gases and volatile hydrocarbons was such that mass spectrometry gained recognition as a tool for quantitative analysis.[7] The petroleum industry was particularly interested, since fast and accurate analyses of complex hydrocarbon mixtures could be facilitated at a time when there was a strong demand for high-quality fuels in the aviation industry; a situation aggravated by World War II. This potential market influenced the commercial production of mass spectrometers and post-war years have seen extensive developments in the design and application of mass spectrometers.

1.2 Mass Spectrometers

Single sector (or single focussing) mass spectrometers[8] were replaced by double sector mass spectrometers and tandem double focussing mass spectrometers;[9, 10] while the quest for greater resolution and enhanced sensitivity resulted in the design and production of exotic mass spectrometers, namely the three[11] and four sector mass spectrometers of varying configuration and geometry.[12] Expansion in design, engineering and technology has seen the name 'mass spectrometer' being applied to instruments such as magnetic / electric sector instruments,[13] quadrupole[14] and triple quadrupole,[15] flowing afterglow (FA),[16, 17] time of flight (TOF),[18] ion trap,[19] ion cyclotron resonance (ICR),[20-22] and Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers.[23] Each of

these instruments can be used to study gas-phase chemical reactions; the choice of instrument being dependent on the specificity of the project objectives.

There are many diverse branches of science which use the powers of mass spectrometry. These range from geology to experimental physics, and from organometallic, inorganic and chemical analyses to protein structure and medical research. The application of mass spectrometry to the study of biomolecules has led to a rapid integration between the fields of biochemistry, organic chemistry and medical research.

In its simplest form, the mass spectrometer is designed to perform three basic functions:

(i) to vaporise compounds of widely varying volatility;

(ii) to produce ions from the resulting gas-phase molecules;*

(iii) to separate these ions according to their mass to charge ratios (m/z), and subsequently detect and record them.

The general design of a mass spectrometer incorporates the basic features illustrated in Figure **1.1**.

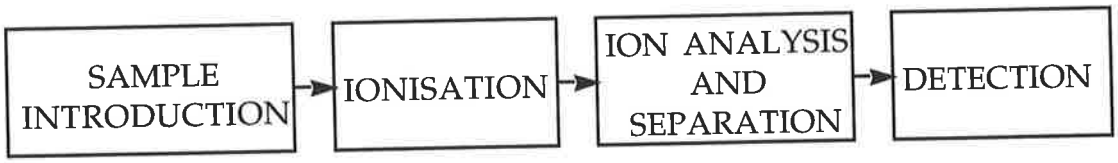


Figure **1.1**: Block diagram illustrating the basic features of a mass spectrometer.

* Except where the volatilisation process directly produces ions rather than neutrals.

1.3 Sample Introduction

All compounds, large or small, can be analysed by mass spectrometry. Samples are introduced into the instrument in different ways so they can be vaporised and ionised, *viz*: gases being leaked in; liquids injected through a septum; and solids introduced using a direct probe. There is little need to use excessive heat to accomplish phase transfer as the pressure differential is such that the samples being analysed are easily vaporised. However, there are some samples that are involatile under these conditions and which also decompose when heated; such compounds must be introduced by a different method.

Coupling mass spectrometry with other analytical instruments illustrates the versatility of mass spectrometric analyses. Gas chromatography is recognised as a useful analytical method for the separation of molecules in the gas-phase based upon, *e.g.* volatility, polarity.[24] Since the sample analyte already exists in the gas-phase, it is convenient and also appropriate to then analyse the sample by mass spectrometry. The value of gas chromatography / mass spectrometry (or the acronym GC/MS) is demonstrated by the analyses of thermally stable, 'unknown' sample mixtures; if they can be separated by GC, then their structures can be determined by mass spectrometry.[24]

Compounds which are not thermally stable may be separated by liquid chromatography methods.[25] The ability of high performance liquid chromatography (HPLC) to separate compounds of similar structure, polarity or molecular weight has been well documented. The nature of the substance(s) being investigated determines the choice of stationary and mobile phases, such that maximum resolution is gained with minimal retention time. As well as being an analytical tool, HPLC can also be used to separate large amounts of material in a preparative manner. This method requires

larger diameter columns containing the stationary phase and uses faster flow rates of the mobile phase in comparison to analytical experiments. The pure compound is isolated by collecting and evaporating the mobile phase.

Attempting to couple liquid chromatography with mass spectrometers proved to be challenging, since the sample analyte is dispersed in liquid. A solution to the problem led to the development of electrospray ionisation, which is rapidly being accepted as the primary method of analysing very large, thermally unstable compounds such as peptides and proteins.[26, 27] The technique of ionisation is not relevant to this thesis, however it should be mentioned that electrospray ionisation generates multiply charged ions of the analyte.[26, 27] All ions discussed in the core of this thesis are singly charged.

Other methods recognised for potential interfacing with mass spectrometers include: supercritical-fluid chromatography (SFC) which uses liquid carbon dioxide as the mobile phase (for reviews see *e.g.* ref 28, 29); and capillary electrophoresis (CE) which can generate high resolution chromatograms in a fraction of the time required by HPLC and does so using micro volumes of sample.[29]

1.4 Ionisation Methods

There are many ionisation techniques used today to produce ions of interest from the sample under investigation. Ionisation methods in common use are: electron impact (EI),[8, 30, 31] chemical ionisation (CI),[32, 33] field desorption (FD),[34] desorption ionisation [including secondary ion mass spectrometry (SIMS),[35-37] fast atom bombardment (FAB),[38, 39] liquid SIMS

(LSIMS),[40] continuous flow FAB,[41] laser desorption (LD),[42] matrix assisted LD ionisation (MALDI),[43] and ^{252}Cf plasma desorption[44, 45]], and several nebulisation techniques such as thermospray[46] and electrospray.[26, 27]

1.5 Electron Impact

The first used and easiest method of ionisation is the process known as electron impact (EI), where a sample vapour at a reduced pressure flows through a region traversed by an electron beam. When electrons pass through or very near a molecule ionisation occurs, as described in scheme (**1.1**).

$$M: + e^- \longrightarrow M^{+\cdot} + 2e^- \qquad (\underline{1.1})$$

Usually this process results in the formation of positive ions, but electron capture is possible, resulting in the formation of negative ions.[47] Ionisation is dependent on the energy of the electron beam, the ionisation potential of the molecule and the nature of that molecule.[48, 49] The energy of the electron beam is typically between 5 and 100 eV, with efficient ionisation generally occurring with an energy of 50 eV. An energy of 70 eV is usually chosen, since sensitivity is close to a maximum and fragmentation is uneffected by small changes in electron energy about this value. This creates reproducible spectra that contain odd-electron ions formed by an ejection of one electron from the molecule.[50, 51] The minimum energy required to remove one electron from the highest occupied molecular orbital of a molecule is called the ionisation potential.[52] A second or even a third electron can be removed, albeit with

decreasing probability due to the increase in energy required. Subsequent fragmentation of the excited molecular ion produces the spectrum of fragment ions that is interpreted to deduce the structure of the original molecule. Only about 1% of the molecules in the source are ionised; and about 20% of EI spectra of small organic molecules do not contain a significant relative abundance of the molecular ion. This can complicate the assignment of the molecular mass of the compound being investigated. Despite this, approximately 90% of all organic molecules can be ionised by this technique.

Electron impact can produce either positively or negatively charged ions in the ion source. The formation of negatively charged species derived by EI are relevant to this thesis (see Chapter 5) and are discussed in detail. There are three main processes for the formation of negative ions by EI.[53-56] The ionisation process occurring is determined by the energy of the electron. These processes are classified as:

(i) Resonance Capture

Providing that the molecule (AB) has a positive electron affinity, then low energy electrons (~0.1 eV) can be captured [scheme (**1.2**)]. In the presence of a non-reactive gas[32] (*e.g.* N_2) or at high sample pressures, the process is more important since the excited species may be stabilised by collision with a neutral molecule or by radiation, otherwise the captured electron is expelled. If the captured electron has enough energy to produce an electronic transition to a higher energy level than that of the required dissociation energy, then the species $AB^{-\cdot}$ will dissociate into $A^{\cdot}$ and B^{-}.

$$AB + e^{-} \longrightarrow AB^{-\cdot} \qquad (\mathbf{1.2})$$

(ii) Dissociative Resonance Capture

Dissociative resonance capture is an important process if the electron beam energy is less than 15 eV and provided that the molecular anion ($AB^{-\cdot}$) readily fragments, then when AB captures an electron it will undergo an electronic transition to give A^- and the radical $B^\cdot$, (or B^- and the radical $A^\cdot$) [scheme (**1.3**)].

$$AB + e^- \longrightarrow [AB^{-\cdot}] \longrightarrow A^- + B^\cdot \quad (\text{or } A^\cdot + B^-) \qquad (\mathbf{1.3})$$

(iii) Ion Pair Formation

Ion pair formation is a non-resonance process that occurs when an electron with large kinetic energy (greater than 1 eV) is captured by a molecule. The translational energy of the electron is converted into vibrational energy within the molecular ion. The ion readily dissociates resulting in the formation of the ion pairs A^- and B^+, (or A^+ and B^-), both of which may be in an excited state; negatively charged molecular ions are rarely observed [scheme (**1.4**)].

$$AB + e^- \longrightarrow A^- + B^+ + e^- \quad (\text{or } A^+ + B^- + e^-) \qquad (\mathbf{1.4})$$

Using typical operating conditions (*i.e.* electron beam energy of 70 eV) all three processes may occur in an EI source. Since the processes are energy and pressure dependent, the intensity of the ion signal will change with sample pressure. Although dependent on the nature of the sample molecule, the probability of electron capture is about 1000 times less than that of electron removal, hence the intensity of negative ions is generally a factor of 10^3 lower than the intensity of positive ions under similar conditions. However,

electron energies in the range 40 - 70 eV have also been used to generate negative ions in electronic ground states, formed by processes (i) and (ii). This energy range often gives more abundant negative ions because in this range, anions are formed by the capture of secondary (or thermal) electrons produced by gas-phase positive ionisation processes, or alternatively from electrode surfaces.[54, 56]

EI ionisation is the most widely used method of ionisation in mass spectrometry. For most molecules it produces both molecular and fragment ions, thus allowing determination of both relative molecular mass and molecular structure for such molecules. However, EI ionisation does have some significant drawbacks:

(i) it may be difficult to measure relative molecular masses for some molecules, especially if almost all the molecular ions formed fragment before they leave the ion source;
(ii) differentiating between isomers can be difficult, especially enantiomers;
(iii) some compounds may undergo thermal decomposition prior to ionisation, or be very prone to fragmentation after ionisation because of the temperature required for vaporisation;
(iv) others may simply be too involatile to give a spectrum.

When such problems are encountered, recourse must be made to alternative methods of ionisation.

1.6 Chemical Ionisation

Munson and Field reported the technique of ionising a sample of molecules by gas-phase ion-molecule reactions.[32] The initial work in chemical ionisation (CI) mass spectrometry used positive ion-molecule reactions[57] and the last 20 years has seen an increased interest in negative ion-molecule reactions by using negative ion chemical ionisation (NICI).[58] The NICI technique requires a high pressure (typically 1 Torr) of a reagent gas such as H_2O or NH_3 which is ionised by electron impact.[59-61] The hydroxide ion and the amide ion formed in the gas-phase are particularly strong bases ($\Delta H_{acid\ H_2O}$ = 390.7 kcal.mol^{-1}, $\Delta H°_{acid\ NH_3}$ = 403.6 kcal.mol^{-1}).[62] The primary ion produced, HO^- or NH_2^-, then reacts with the sample by way of ion-molecule reactions to produce negative ions. The formation of the reactant ions NH_2^- and HO^- from their respective neutrals NH_3[63] and H_2O[64] are shown in schemes (**1.5**)-(**1.7**).*

$$NH_3 + e^- \longrightarrow NH_2^- + H^\cdot \qquad (\mathbf{1.5})$$

$$H_2O + e^- \longrightarrow H^- + HO^\cdot \qquad (\mathbf{1.6})$$

$$H^- + H_2O \longrightarrow H_2 + HO^- \qquad (\mathbf{1.7})$$

Hydroxide ions are also formed in the gas-phase, by electron bombardment of equal amounts of nitrous oxide and a hydrocarbon such as methane, isobutane or cyclohexane [schemes (**1.8**) and (**1.9**)].[65, 66]

$$N_2O + e^- \longrightarrow O^{-\cdot} + N_2 \qquad (\mathbf{1.8})$$

$$O^{-\cdot} + CH_4 \longrightarrow HO^- + CH_3^\cdot \qquad (\mathbf{1.9})$$

* It should be noted that H^- and $O^{-\cdot}$ (from O_2) are also present and these ions also contribute to the formation of $[M - H]^-$ ions.

There are many available reagent gases that can produce reagent ions that act as Brönsted bases including: MeO^- (from MeONO),[67] $O^{-\cdot}$ (from N_2O)[68], and F^- (from NF_3).[69] The generation of parent ions produced by NICI formed from the reactions that occur between a reagent ion and the sample molecule can be categorised into four ion-molecule reactions.

(i) Proton Transfer (Deprotonation)[57]

NICI Brönsted bases react primarily by the process shown in scheme (**1.10**).

$$R^- + MH \rightleftharpoons M^- + RH \qquad (\mathbf{1.10})$$

If the gas-phase proton affinity of R^{-*} is greater than the gas-phase proton affinity of M^-, then the proton transfer reaction will be exothermic and occur efficiently.[70] Equilibrium constant measurements of the reaction can provide values for the relative proton affinities of the negative ions R^- and M^-.[71] Similarly, a relative gas-phase scale for MH and RH can be constructed.[62] Many organic molecules contain acidic hydrogens; thus proton transfer processes are an important class of reactions, since these acidic hydrogens are easily deprotonated. Analyte anions produced by proton transfer have very little excess internal energy. As can be seen from scheme (**1.10**), no new bonds are involved in the formation of the analyte anion; thus most of the excess energy is contained in the neutral derived from the reagent gas ion. As a result, abundant $[M - H]^-$ ions are formed, thus providing molecular weight information,[58] and structural information if collisionally activated (see sections **1.10** and **1.11**).[72]

---

* R^- is the reactant gas phase ion formed from electron impact ionisation of the reagent gas, *i.e.* HO^- or NH_2^-.

(ii) Charge Exchange Reactions[73]

Charge exchange reactions will occur if the electron affinity of M is greater than the affinity of R.[57] Values for the relative electron affinities of M and R can be provided from the measurements of electron transfer equilibria using scheme (**1.11**).[73]

$$R^- + M \longrightarrow M^- + R \qquad (\mathbf{1.11})$$

(iii) Nucleophilic Addition[74]

Reactant ions R^-, which are not strong Brönsted bases generally undergo nucleophilic attack onto the molecule M to form stable addition complexes MR^- [scheme (**1.12**)], often in the presence of a non reactive gas. Complexes such as MR^-, are occasionally formed in a CI source. The addition complex may either be a covalently bound adduct or a solvated ion pair.[75]

$$R^- + M \longrightarrow MR^- \qquad (\mathbf{1.12})$$

(iv) Nucleophilic Displacement[57, 76, 77]

S_N2 nucleophilic displacement reactions are an important method of ion formation, especially to generate specific gas-phase anions. Such a reaction may occur if it is exothermic (or thermoneutral) and if there is no possibility of a competing proton transfer from MX to R [scheme (**1.13**)].[57]

$$R^- + MX \longrightarrow M^- + RX \qquad (\mathbf{1.13})$$

One example of an S_N2 displacement reaction[78] is the formation of the novel anion prepared as described in scheme (**1.14**).[79]

$$MeO^- + Me_3SiCH_2CH_2CHO \longrightarrow {}^-CH_2CH_2CHO + Me_3SiOMe \quad (\mathbf{1.14})$$

1.7 Fast Atom Bombardment (FAB)

The basic requirement of an ionisation technique is to present the sample for analysis in the gas-phase. This limits the EI and NICI techniques to samples that have a relatively low molecular weight, are volatile and are also thermally stable. Consequently, thermally labile molecules, high molecular weight compounds and polar molecules cannot be ionised by these techniques. Therefore a different method of ionisation is required to analyse such samples by mass spectrometry.

Desorption ionisation methods overcome this problem by giving a large pulse of energy to the sample. The effect of this is to put a relatively large amount of energy into translational modes involving sample molecules. Thus, intermolecular bonds involving the sample, such as hydrogen bonds, are broken in preference to covalent bonds and the sample is desorbed from its environment into the gas-phase. Thermal decomposition is reduced by avoiding an equilibrium distribution of the large amount of available energy. This is because the sample molecule leaves its environment within a time magnitude of 10^{-12} seconds. Examples of desorption ionisation processes include: fast atom bombardment (FAB), where the bombarding particle is an energetic (keV) neutral particle;[38, 39, 80] secondary ion mass spectrometry

(SIMS), where the bombarding particle is an energetic (keV) ion;[36, 81] ^{252}Cf plasma desorption, in which the bombarding particle is a very energetic (MeV) ion derived from spontaneous fission of a radionuclide or extracted from a particle accelerator;[82, 83] and laser desorption to the extent that sputtering phenomena may occur in organic samples bombarded by an intense laser beam.[84] FAB was used to generate the positive and negative ions studied and reported in this thesis.

FAB is a technique where a neutral atom beam of large translational energy, typically 6 - 8 keV, bombards the sample that is dissolved in a liquid matrix of low volatility. An atom gun is used to generate the beam of fast atoms.* The bombarding atoms must be chemically inert in order to avoid the possibility of unwanted chemical reactions. Argon (or xenon) is often employed for this purpose. In a low pressure environment the FAB gun ionises argon atoms generating a plasma of argon ions, which are then accelerated to a known kinetic energy and focussed into a beam of high intensity. They are then neutralised by exchange of resonant gas-phase charge as they pass their almost stationary neutral counterparts. This process is described by schemes (**1.15**) and (**1.16**). Any remaining argon ions with high translational energies are removed by simple electrostatic deflection, leaving a beam of fast argon atoms which are focussed to strike the sample on the probe tip and is illustrated in Figure **1.2**.

$$Ar + e^- \longrightarrow Ar^+ + 2e^- \qquad (1.15)$$

$$\overrightarrow{Ar^+} + Ar \longrightarrow \overrightarrow{Ar} + Ar^+ \qquad (1.16)$$

* In many cases a caesium ion gun is alternatively used.[85]

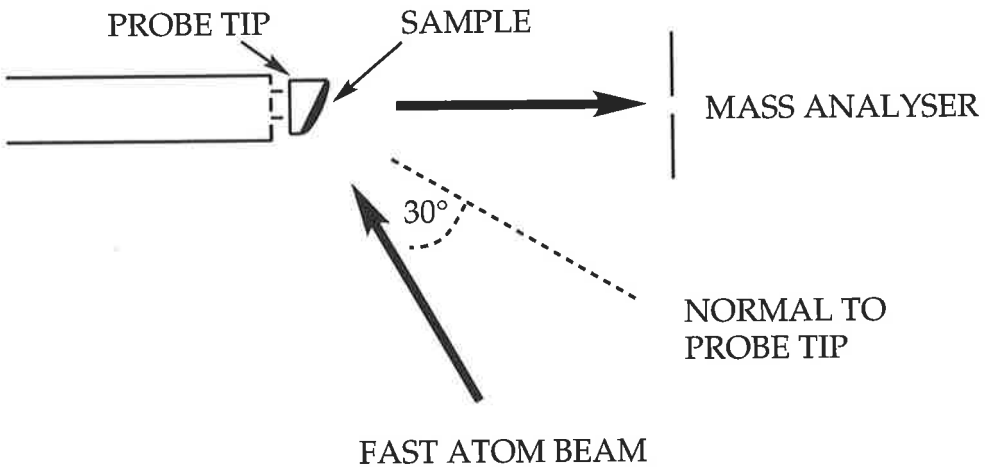


Figure **1.2**: Schematic diagram of the FAB ionisation process.

The probe tip is orientated such that the strongest ion current is produced. This is known as the optimum angle of incidence and is defined as the measured angle between the bombarding atom beam and the normal to the surface of the probe tip and is usually 30°.[86] The sample to be analysed is mixed with a low vapour pressure liquid, called a matrix, and uniformly spread over the probe tip in an attempt to simulate a monolayer;[80, 87] when the fast argon atoms impact into the solution, momentum transfer results and the sample is desorbed, often as an ion. The beam of sample ions is then analysed in the mass spectrometer in the usual way. Often ionised oligomers of the matrix are desorbed along with the sample ions, which may serve to usefully mark the mass spectrum {*e.g.* in positive ion analysis $[(\text{glycerol})_n + H]^+$, m/z (92n + 1), where n is an integer}. The molecular ions formed are most often even electron ions; protonated and deprotonated molecular ions are formed in equal abundance,[88] although only one charged type is admitted into the mass spectrometer for further analysis. Under most conditions the degree of fragmentation is an increasing function of the energy density of the primary ion beam and of the chemical and physical nature of the

sample and its support, since the ions are ejected from the condensed phase. FAB is classed as a 'soft' ionisation method, thus suiting compounds that are thermally unstable, have high molecular weights and are polar.

It is important that the sample should dissolve in the matrix, preferably being marginally more hydrophobic than the matrix so that it will occupy the matrix / 'vacuum' interface. This is desirable because the atom beam only penetrates about 10 nm into the matrix. The most surface-active[89, 90] or hydrophobic components[91] are said to populate the surface and suppress ionisation of the remaining components. Matrix additives may be used to increase surface activity and thereby increase ion currents.[92, 93] Glycerol is a commonly used matrix.[80, 87] Alternative matrices are thioglycerol or a eutectic mixture of dithioerythritol and dithiothreitol (5:1),[94] known as 'magic bullet'.[95]

Positive and negative ion FAB mass spectra are recorded with similar facility, with molecular weights normally being given by abundant $[M + H]^+$ and $[M - H]^-$ ions, respectively. It is true that the neutral sample molecule [M], will also be desorbed, but since the molecules being analysed usually have polar sites (*e.g.* $-CO_2H$, $-NH_2$), then the corresponding ions ($-CO_2^-$, $-NH_3^+$) are desorbed in comparable amounts. The pH of the analyte sample in the matrix can be altered on the probe tip; the addition of a small amount of acid or base can significantly enhance the $[M + H]^+$ and $[M - H]^-$ ion signals respectively.[96] FAB spectra frequently contain structurally useful fragment ions, but in the positive ion spectra of large molecules (*e.g.* peptides) obtained from a matrix, $[M + H]^+$ ions are often the most abundant and dominant ions; making molecular weight determination extremely easy by this technique. Also, the applicability and ease of coupling to existing mass spectrometers have shown that FAB is a useful and valuable technique for producing ions from such compounds.

The actual mechanism of ion formation by FAB processes is still debated, with a number of mechanisms proposed to occur.[70] Such proposals range from: the direct sputtering of ions preformed in solution, with evaporation of these preformed ions from splash droplets;[97] to the formation of gas-phase ion molecule reactions in the "selvedge region" just above the matrix surface after desorption.[98, 99] Overall the desorption process may be generally defined as *"the direct formation of stable molecular and fragment ions from non-volatile samples in the condensed phase or, alternatively as a process in which vaporisation and ionisation appear as a single event"*.[70]

Since FAB is a 'soft' ionisation process, the ions have little excess internal energy, resulting in the formation of intense parent ions with little or no accompanying fragment ions. The main strength of the FAB MS technique lies in the formation of these intense, long lasting parent ions, enabling the use of MS/MS techniques (See sections **1.9 - 1.11**).[100] Fragmentation can be induced by collisional activation of these parent ions, with analysis by MS/MS processes to yield structural information.[101]

1.8 Principles of Mass Spectrometry

Sector Instruments

Sector mass spectrometers analyse charged particles using magnetic and / or electric fields.[5] Prototype mass spectrometers used only one magnetic sector, which restricted the application of ion analysis due to low resolution and poor sensitivity. A lack of adequate vacuum systems was also a contributing factor: spectra frequently contained peaks associated with singly and multiply charged ions, as well as product ions formed from ion-molecule interactions occurring

in the ion source. To be able to study gas-phase molecules it is necessary to have the instrument under high vacuum throughout the experiment, thus decreasing the probability of ion-molecule interactions that would otherwise effect the energy and momentum of the ions under investigation.

Following ionisation the ions are accelerated through a potential difference (V) out of the ion source and into the mass analyser. If the ions are formed with negligible kinetic energy, then the kinetic energy of any ion after it has been fully accelerated out of the ion source is equal to the product of the charge and voltage. [Equation (**1.17**)].[102]

$$mv^2 / 2 \quad = \quad zV \qquad (\mathbf{1.17})$$

Where: m is the mass of the ion; v is the velocity of the ion; z is the charge of one electron; and V is the accelerating voltage. Magnetic and electric fields can both influence the trajectory of charged particles.

Magnetic Sector Analysis

Ions with a high translational energy traverse the first field free region and enter the magnetic sector, where they will experience a centrifugal force (zvB) which is perpendicular to the magnetic field and to their direction of travel, but equal to the centripetal force (mv^2/r). Consequently, the ions will be deflected in a circular path of radius r, in a plane normal to the direction of the magnetic field of strength (B). As a result, ions with different mass to charge ratios (m/z) are separated according to their momentum while traversing the

magnetic sector according to equation (**1.18**).[102] Thus the magnetic sector can be viewed as a momentum analyser.

$$r = mv / zB \qquad (\mathbf{1.18})$$

Ions with equal kinetic energy but different mass will have different trajectories through a fixed magnetic field. The ion separation is dependent on the radius of curvature in the magnetic field as can be seen by equation (**1.19**).[102]

$$m/z = B^2r^2 / 2V \qquad (\mathbf{1.19})$$

Since mass spectrometers are built with a fixed magnetic sector radius, the separation of ions becomes a function of both the accelerating voltage (V) and the magnetic field strength (B). Mass spectra can therefore be recorded by scanning either the accelerating voltage or the magnetic field. Variation in the accelerating voltage tends to cause defocussing of the ions and a reduction in sensitivity; mass spectra are usually obtained by varying the magnetic field strength of the magnetic sector.

Electric Sector Analysis

The electric sector consists of two charged parallel plates, where a uniform electric field exerts a force on the ions that is perpendicular to their direction of motion and in the direction of the applied electric field. As a result, the ions

are deflected in a circular path of radius r, as given by equation (**1.20**), where ***E*** is the electric field strength.[102]

$$mv^2 / r \quad = \quad zE \qquad (\underline{\mathbf{1.20}})$$

The above relationship illustrates that ions of the same translational energy follow the same radial trajectory and will be brought to a common focus. Similarly, ions possessing different kinetic energies follow paths of different radii and may be detected by varying the electric field strength. Thus the electric sector can be viewed as a kinetic energy analyser.

1.9 Tandem Mass Spectrometry

Theoretically, ions which have the same kinetic energy and mass to charge ratio can be focussed to a common point after leaving the magnetic sector. Due to slightly different potentials experienced after acceleration, these ions have a spread in kinetic energies, which allows the ions to diverge. Any divergence by ions that have the same mass to charge ratio, will drastically effect the resolving potential of the instrument. This problem can be averted by incorporating an electric sector either side of the magnetic sector, which will focus the ion beam according to kinetic energy and improve the resolution. The term 'tandem mass spectrometry' is used to describe mass spectrometers that have two or more analysing sectors.[103] In 1934, Mattauch and Herzog began a new era in mass spectrometric research by designing an instrument that incorporated an electric sector followed by a magnetic sector.[9, 10] Many

other instruments were developed,[104] including the reverse geometry variety by Nier and Johnson in 1953.[105] For reverse geometry tandem mass spectrometers, the term 'mass spectrometry / mass spectrometry' (or the acronym MS/MS) is used to describe an instrument that uses mass spectrometry for both mass separation and identification.[102] In such an MS/MS experiment, the fragmentation of a parent ion generates daughter ions each with a different momentum and kinetic energy from that of the parent and it is the mass to charge ratio of the various daughter ions that is physically measured (more specifically, it is the energy of the daughter ion that is measured, from which the mass of the ion is obtained). A tandem double focussing mass spectrometer focusses ions according to both momentum and kinetic energy, which increases both the intensity and the resolution of the ion signal.[106]

The ability of a mass spectrometer to separate ions of differing masses is known as the resolving power, or resolution. For peaks to be resolved, the height of the valley between adjacent peaks separated by Δm must be 10% of the height of the peaks at m and $m + \Delta m$.[107] The resolution of an instrument may be determined by the ratio of $m/\Delta m$. This is illustrated in Figure **1.3**.

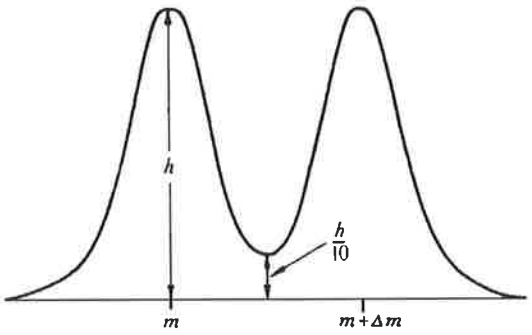


Figure **1.3**: Illustration of the 10% valley between the resolved peaks m and Δm, which are used to determine the resolution of a mass spectrometer.[108]

1.10 Collisional Activation[109]

The methods used to generate ions in the gas-phase are classified as 'hard' or 'soft' ionisation techniques. 'Hard' ionisation methods, such as electron impact, give excess internal energy to the sample being investigated and results in fragmentation in the source. Analysis of the these fragmentations allows the structure of the sample to be elucidated. 'Soft' ionisation methods such as CI and FAB produce ions that have relatively little excess internal energy, resulting in the formation and detection of abundant molecular ions. This is particularly the case when peptides are ionised by FAB, as stable even electron ions $[M + H]^+$ or $[M - H]^-$, are produced which have little tendency to fragment and little structural information is obtained. Mass spectrometers that use soft ionisation methods are designed with the incorporation of collision cells located at focal points in field free regions. Ions of high translational energy (*e.g.* 8 keV) enter a collision cell and impact with an inert gas such as helium or argon. The concentration of collision gas in the collision cell is directly proportional to the pressure and this is measured by a gauge situated just outside the collision cell.* A pressure of approximately 5×10^{-7} Torr reduces the intensity of the ion beam by 10%, which has been reported to correspond essentially to single collision conditions.[110] Kinetic energy is the only available ground state energy possessed by an ion after being fully accelerated from the ion source. This process is known as collisional activation and involves converting some portion of the ion kinetic energy into internal electronic excitation energy from the interactions between the ion and the electrons of the collision gas. The excess internal excitation energy generated by the collision process is rapidly converted into vibrational energy which:

---

* Since the gauge is outside the cell, the measured pressure is slightly lower than that inside the cell due to the gas dissipating from the cell.

(i) causes the ions to fragment, thereby producing daughter ions and releasing energy through exothermic fragmentations;

(ii) intensifies existing fragmentations;

(iii) allows reactions to occur for pathways which normally have large activation barriers.

Many positive and negative ion spectra have been generated using collisional activation, yielding analytical and structural information or giving mechanistic information related to gas-phase physical organic chemistry. In cases where collisional activation results in the dissociation of that ion in the same reaction region, the process is termed collision induced dissociation (CID). In a reverse geometry instrument this leads to the acquisition of collisionally activated mass analysed ion kinetic energy spectra (CA MIKES or CA MS/MS).

1.11 Collision Activated Mass Analysed Ion Kinetic Energy Spectroscopy

Reverse sector instruments have the magnet preceding the electric sector. Therefore, it is possible to focus the magnet on an ion of a particular mass to charge ratio before analysis by the electric sector. Ions of different mass to charge ratios will have differing trajectories and consequently collide with the walls of the mass spectrometer. This has powerful analytical applications if there is a collision cell situated in the second field free region. Collisional activation of the mass selected ion is therefore possible; and by scanning the electric sector, the resulting fragment ions are analysed according to their kinetic energy.

If a parent ion of mass M_p^-, decomposes on collisional activation to daughter ions of mass M_d^- and neutrals of mass M_n [equation (**1.21**)], the initial translational energy of M_p^- will be shared among the daughter ions M_d^- and the neutral molecules M_n.

$$M_p^- \xrightarrow{\text{CA}} M_d^- + M_n \qquad (\underline{\mathbf{1.21}})$$

Thus a daughter ion M_d^-, will possess a lower kinetic energy E_d, than that of the parent ion M_p^-. A reduction in the electric sector voltage E_p, by the value M_d/M_p, will enable daughter ions of mass M_d, to traverse the electric sector and produce a daughter ion spectrum. The masses of the daughter ions, M_d, are given by equation (**1.22**).

$$M_d = M_p \times E_d / E_p \qquad (\underline{\mathbf{1.22}})$$

where M_p is the mass of the parent ion beam set by the magnetic sector; E_p is the electric sector voltage required to transmit the parent ion beam; and E_d is the reduced electric sector voltage required to transmit a daughter ion of mass M_d as measured from the spectrum.

The peaks detected in a CA MIKES experiment are generally broad with the width of the peaks corresponding to the amount of kinetic energy released in the cleavage process to form that particular ion. A large amount of kinetic energy release is generally indicated by the occurrence of a disc-shaped peak and is indicative of a fragmentation that has a significant reverse activation energy.[111, 112]

Interference Peaks

Occasionally anomalous peaks are observed in MIKES experiments, which can be formed from higher mass ions that fragment prior to the magnet and form daughter ions that have the same momentum as the selected parent ion. These ions therefore traverse the magnetic sector along with the parent ion beam. Since the electric sector will focus ions into a narrow energy range, any ions formed prior to the magnetic sector are highly resolved. Therefore, artefact or interference peaks can be identified by their highly resolved, narrow appearance in comparison to the daughter ion peaks. However, if the artefact ions fragment in the second field free region, they then become artefact daughter ions which are indistinguishable from true daughter ions.

1.12 Quasi Equilibrium Theory[113]

The theory of unimolecular reactions is often referred to as the Quasi equilibrium theory.[114] The Quasi equilibrium theory of mass spectrometry attempts to (i) rationalise how parent ions (which are formed with an internal energy in excess of the ground state) can decompose to form fragment ions and neutral species; and (ii) how a parent ion may fragment by more than one pathway. The assumption that energy is randomised throughout the ion prior to fragmentation has important implications. The daughter ion spectrum of a mass selected parent ion is determined by: the energy dependent rate constants of possible reaction pathways; the time required for the reaction to occur; and the internal energy distribution of the ions.

A mass spectrum is a recording of a series of competing consecutive unimolecular or collision induced decompositions of a parent ion. Ions are

initially produced in various electronic, vibrational and rotational energy levels, but rapid transitions (without energy release) lead to randomisation of the energy and the formation of vibrationally excited ground state ions which undergo fragmentation. Mass spectra therefore depend upon the initial transfer of energy and not upon the method by which the energy is transferred.

The work presented in this thesis is composed of two sections investigating peptide chemistry in the gas-phase. Chapters 3 and 4 apply positive ion analysis to determine the primary sequences of amphibian peptides. Chapter 5 investigates the negative ion fragmentations of tetrapeptides as part of an ongoing research into the understanding of deprotonated peptide chemistry in the gas-phase. Due to constant improvements in design and technology, there are many excellent mass spectrometers available today which are suitable for these objectives: one such instrument is the VG ZAB 2HF mass spectrometer.

1.13 Vacuum Generators ZAB 2 Sector High Field Mass Spectrometer[115, 116]

The VG ZAB 2HF instrument is housed in the Department of Chemistry, University of Adelaide and was used to obtain all mass spectra presented in this thesis. It is a sector instrument consisting of a magnetic sector* followed by an electric sector† arranged in the reverse Nier-Johnson configuration. A schematic diagram of the instrument is in Figure **1.4**.

Positive and negative ions may be generated in either the combined EI/CI ion source, or using the FAB ion source and operation is possible in either the

---

* The VG ZAB 2HF instrument has a magnetic sector of 30 cm radius that subtends an arc of 55°.
† The VG ZAB 2HF instrument has an electric sector of 38 cm radius that subtends an arc of 81°.

positive or negative mode of analysis. Both modes have been used in the work described in this thesis.

The instrument's reversed geometry configuration and double focussing capability allow high mass and energy resolution with the ions separated and focussed according to their momentum vectors in the magnetic sector and analysed in terms of kinetic energy in the electric sector. This arrangement therefore permits mass analysed ion kinetic energy spectra (MIKES) to be recorded.[117] Charge reversal and linked scanning of the magnetic and electric sectors are also possible. There are collision cells incorporated in the first and second field free regions, between the ionisation chamber and the magnet, and the magnet and electric sector, respectively. These collision cells enable collisional activation and hence collision induced decompositions to be observed. Ion detection is enhanced by an electron multiplier situated in the second field free region for single-focussing and a photon multiplier located after the electric sector for double-focussing operation.

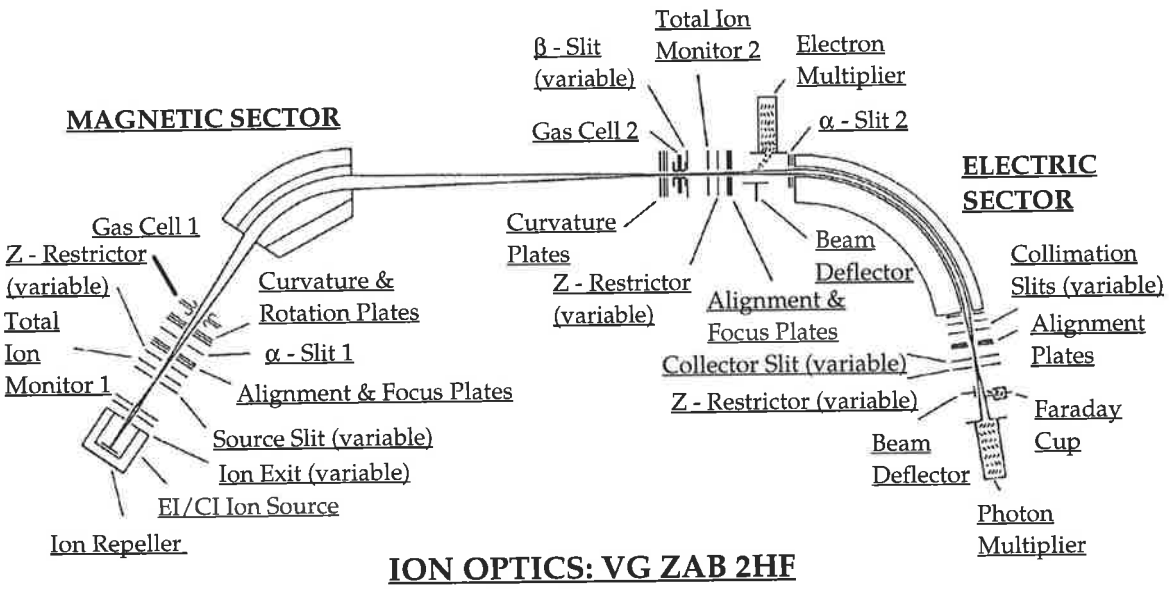


Figure 1.4: Schematic diagram of the VG ZAB 2HF mass spectrometer.

Chapter 2 AMPHIBIANS AND PEPTIDES

"An observation of substantial value is that all, or nearly all, amines and peptides found in amphibian skin have their counterpart in mammalian tissues where they usually occur in much lesser variety and concentration. Hence, results obtained in the study of amphibian skin are of interest transcending comparative pharmacology and biochemistry, as they may substantially contribute to the understanding and interpretation of facts assessed in mammals and may offer the basis for new research trends in higher vertebrates."

Vittorio Erspamer[118]

2.1 Introduction to Amphibians

The word "amphibian" is derived from the Greek words *amphi* and *bios*, meaning "living a double life", which accurately defines an animal that is capable of living in water as well as on land.

There are three orders of amphibia:[119]

(i) Anura (frogs and toads);

(ii) Caudata (salamanders and newts);

(iii) Gymnophiona (wormlike burrowers),

of which, the anurans have evolved into 24 families and more than 4000 distinct species.[120] Anurans are the amphibians that are the focus of discussion in this thesis. At the close of the Paleozoic Period, around 300 - 350

million years ago, amphibians made the most remarkable evolutionary changes out of all the vertebrates, by emerging from the Devonian seas to begin living on land.[121, 122] The structural changes that took place for amphibians to adapt to terrestrial habitation included the development of: legs, as a method of transportation between water sources; lungs, to permit breathing when water conditions deteriorated;[123] a mechanism for maintaining eggs under moist conditions;[124] and dermal glands in the skin that produced mucus secretions to prevent desiccation. It is the development of the skin secretory mechanism that has remained a major contributor which influenced survival in new habitats and the stability of the populations throughout the ages.[123]

2.2 Glandular Characteristics

The environments that amphibians inhabit contain a plethora of animal predators and micro-organisms. Consequently the dermal glands contain various host defence chemicals, including alkaloid toxins,[125, 126] biogenic amines,[127-131] proteins,[130] mucins,[125] bioactive peptides,[130] and peptide fragments associated with prohormone processing events;[132] all of which have contributed to the welfare of the animal in hostile environments. The dermal layer of the skin contains a non-homogenous distribution of three histologically distinct types of cutaneous glands: the mucous,[125] lipid[133] and the granular[125, 134-137]; which are primarily distributed throughout the dorsal surface of the skin[135, 136, 138] and are linked to the exterior surface by means of secretory ducts.[125, 135, 139, 140] The granular glands are of current interest, as they are filled with circular vesicles with diameters of only a few microns and this is where the components for the secretion are synthesised

and stored.[135, 141] They are occasionally aggregated in hypertrophied regions,[142] which is a feature peculiar to amphibians. These regions have been classified into eight forms,[142] which are listed together with their corresponding anatomical position in Table **2.1**. An example of one of these dermal glands is depicted in Figure **2.1**.

Expulsion of the glandular secretions is under neuronal regulation[143] and occurs in response to stress, such as in a defensive behavioural disposition against a predator (including bacteria); for which 80 - 90% of the glandular contents are rapidly depleted.[135, 141, 144, 145] The secretion can also be released by stimulating the amphibian's peripheral nervous system by methods such as: handling;[125, 138, 146] mild electric current;[137, 147] surface electrical stimulation;[148] and injections of specific neuromodulators.[135, 141, 143, 149]

Table 2.1

Hypertrophied Regions and their Corresponding Anatomical Position in Frogs

(1)	Coccygeal	on the flanks of each side of the body;
(2)	Femoral	on the posterior side of each femur;
(3)	Inguinal	on each side of the body in the groin;
(4)	Parotoid	on the shoulders;
(5)	Rostral	on the head;
(6)	Submandibular	adjacent to the lower jaw;
(7)	Supralabial	on the upper lip extending posteriorly beyond the angle of the jaw;
(8)	Tibial	on the calf.

Figure **2.1**: *Litoria splendida* has an enormous rostral gland. Photograph courtesy of M. J. Tyler, Department of Zoology, University of Adelaide, 5005.

2.3 Human Interest

Human association with anurans occurs in various forms in many cultures and societies, with the earliest recorded inter-relationship between the two occurring some four thousand years ago.(150)* Anurans have always intrigued people especially in areas such as: mythology; witchcraft; folktales; literature; medical science; food; or poisons for hunting and warfare (for a review see ref

* Undoubtedly the association predates written history.

151). This curiosity has led researchers to scrutinise frogs and seek scientific explanations for their characteristic physiological and biochemical traits; and also try and understand the significance of why particular frogs are used by ancient and current tribal cultures for certain rituals, medical cures, hunting and tribal warfare.

Amphibians still continue to arouse, on occasions, a sense of stigmatism, even though there is renewed interest in the therapeutic properties of amphibian skins - properties which have been recognised for more than two thousand years.[142] For example, a typical medicinal application of powdered frog skin was used as a heart stimulant and a diuretic.[152] In parts of South America the treatment of warts and fungal infections was simply to strap a frog or toad to the infected area.[120] Furthermore, Amazonian natives would mix saliva with the skin secretion from a live frog, and then administer the concoction to a line of fresh burns. The purpose was to achieve 'hunting magic' by reaching a state of euphoria, after recuperating from profound malaise and then listlessness.[153] Oddly enough there exists no evidence of aborigines using frogs in any facet of their culture,[120, 154] except the Water Holding Frog *Cyclorana platycephalus* , which has been used as a source of water in difficult times.[120] However, not all frogs, and in particular certain parts of their anatomy, could not be used for therapeutic gain; there exist numerous accounts in medical journals of people succumbing to poisons from touching or eating frogs.[155] One recent example exemplifies the dangers of frog toxins, where an Australian student who took up a friendly bet to eat the ovaries of *Bufo marinus*, consequently won $20 and a cardiac arrest.[156]

The quest to chemically identify the bioactive agents in amphibian skin began in the early 1960's, and has since resulted in them being classified[126] into four major categories as: (i) peptides;[130] (ii) biogenic amines;[127-131]

(iii) bufogenins;[157] and (iv) alkaloids.[125, 126] Vittorio Erspamer and colleagues pioneered the work into the discovery and isolation of amphibian skin peptides and has since established that there is a wealth of biologically active peptides awaiting discovery.[118, 128, 130, 131, 158-175] In certain frogs, some of these peptides are present in the stomach,[176] brain,[177, 178] central nervous system,[179] and blood.[180, 181]

Characteristic features of frog skin peptides are:

(i) they are often present in large quantities in the skin, especially when compared to their counterparts in mammalian tissues;[174]

(ii) different classes of pharmacologically active molecules are stored within the skins of different species.[182-184]

Erspamer has predicted that every frog peptide will have a mammalian counterpart[118] and cumulative evidence is supporting this prediction. A multitude of species have been investigated, and has yielded an unexpected diversity in the catalogue of previously undiscovered peptides. With over 4000 species of frogs known to date, only a comparatively small number have been investigated. With the amphibian's tremendous diversity of readily accessible, abundant, biologically active peptides, the frog is a vital source in the discovery of hormones, neuropeptides and other substances that may not be easily found within mammalian model systems.[174]

There is accumulating evidence that global frog populations are in serious decline.[185, 186] Unfortunately, the traditional method to isolate these bioactive peptides have involved killing the frog. In many studies more than one thousand specimens of a single species would be slaughtered and their skins removed and dehydrated before processing.[159, 167, 187] It is therefore of little wonder that methods of extraction had to be developed whereby the

secretions could be obtained without the animal being sacrificed. There are many techniques that have recently been developed to meet this requirement.[135, 137, 141, 143, 144, 147-149] The latest method is termed 'surface electrical stimulation' (SES).[148]

2.4 Extraction and Isolation

For the work presented in this thesis, SES was the method used to extract the secretions from the frog's dermal glands. It is a benign method where a typical extraction involves holding the frog by the back legs, moistening the skin with distilled water, and rubbing the back of the animal gently in circular motions with a 21 G platinum bipolar electrode. A milky white, viscous, and occasionally odoriferous secretion soon appears on the skin surface, which is washed from the frog with deionised water into a collecting vessel. This procedure is continued for 30 seconds per frog and appears to have no adverse effect toward the animal; the same specimen can be 'milked' on monthly intervals, which is sufficient time for the glands to replenish.* The results obtained are similar, if not better, than if hundreds of animals were sacrificed.

Previous methods[159, 167, 187] for the isolation of the individual peptides required the dried frog skins to be powdered, extracted with copious volumes of alcohol before separation using alumina column chromatography, with the structure determined by various bio-assays. Advances in technology have seen this method being surpassed by newer methods of chromatography, namely reversed phase high performance liquid chromatography (HPLC).

---

* Previous studies have shown that the glands are replenished within 7 days.[141]

2.5 Reported Peptides from Frogs

The peptides that have so far been isolated from frogs have been grouped into categories by structure and biological activity. Many peptides that have similar structure tend to show similar biological activity, although the extent of activity varies from peptide to peptide. Below is a list of peptide families, together with the name and sequence of an example peptide, and the frog from which it was isolated. Residues which are highlighted are common to the family; those which begin with a capital letter are the L enantiomer; residues with lower case beginnings are the D enantiomer.

Angiotensins [crinia angiotensin II, *Crinia georgiana*[188]]

Ala Pro Gly Asp Arg Ile Tyr Val His Pro Phe (OH) (1)

Bombesins [litorin, *Litoria aurea*[189]]

pGlu Gln **Trp Ala Val Gly** His Phe **Met (NH_2)** (2)

Bombinins [bombinin, *Bombina variegata*[190]]

Gly Ile Gly Ala Ser Ile Leu **Ser Ala** Gly **Lys** Ser Ala **Leu Lys Gly Leu Ala Lys Gly Leu Ala Glu His Phe** Ala Asn (NH_2) (3)

Bradykinins [ranakinin-R, *Rana rugosa*[191]]

Arg **Pro** Pro **Gly Phe** Ser **Pro Phe Arg** Ile Ala Pro Glu Ile Val (OH) (4)

Caerins [caerin 1.1, *Litoria splendida*[192]]

Gly Leu Leu **Ser Val Leu Gly Ser Val Ala Lys His Val Leu Pro His Val Val Pro Val Ile Ala Glu His Leu (NH_2)** (5)

Caerulins [caerulein, *Litoria caerulea*[193]]

pGlu Gln Asp **Tyr(SO_3H)** Thr **Gly Trp Met Asp Phe (NH_2)** (6)

Dermorphins [deltorphin II, *Phyllomedusa bicolour*[173]]

Tyr ala Phe Glu Val Val Gly (NH_2) **(7)**

Dynastins [dynastin 1, *Limnodynastes interioris*[194]]

Gly Leu Leu Ser Gly **Leu** Gly Leu (OH) **(8)**

Magainins [magainin II, *Xenopus laevis*[141, 180]]

Gly Ile Gly Lys Phe Leu His Ser Ala Lys **Lys Phe Gly Lys Ala Phe Val Gly Glu Ile Met Lys Ser (OH)** **(9)**

Pipinins [pipinin I, *Rana pipiens*[195]]

Phe Leu Pro Ile Ile Ala Gly Val **Ala Ala Lys Val Phe** Pro **Lys Ile Phe Cys Ala Ile Ser Lys Lys Cys (OH)** **(10)**

Tachykinins [physalaemin, *Physalaemus fascumaculatus*[196]]

pGlu Ala Asp Pro Asn Lys **Phe** Tyr **Gly Leu Met (NH_2)** **(11)**

Tryptophyllins [tryptophyllin L1, *Litoria Rubella*[197]]

Phe **Pro Trp** Leu **(NH_2)** **(12)**

Xenopsins [xenopsin, *Xenopus laevis*[198]]

pGlu Gly **Lys Arg Pro** Trp **Ile Leu (OH)** **(13)**

Most amphibian peptides have molecular weights below 2000 Daltons. However, some like the magainins and bombinins have molecular weights in the range of 2000 to 3000 Daltons; these particular peptides show potent antimicrobial activity and are thought to exist as simple amphipathic α-helices that associate with lipid membranes and disrupt normal membrane functions.[165, 174, 199]

2.6 Amphiphilicity and Mechanisms of Action

It has been suggested that the functional environment of any peptide acting at a biological surface, such as a membrane or cell surface, will often be amphiphilic.[200, 201] Thus, the expression of activity will usually involve binding at the surface between the hydrophobic core of a structure and the aqueous surrounding. This type of anisotropic environment is likely to induce the formation of discrete segments of secondary structure in the peptide. This will occur if these structures result in the segregation of hydrophobic and hydrophilic amino acid residues within the peptide sequence, creating separate domains of complementary amphiphilicity. The formation of amphiphilic secondary structures, such as α-helices and β-strands and the properties that they exhibit, have been studied in numerous peptide model systems.

A linear sequence composed of alternating hydrophobic and hydrophilic amino acid residues, results in the formation of an amphiphilic β-strand.[202] Peptides of six or more residues with high β-strand content have a pronounced tendency to self-associate* forming amphiphilic β-sheets,[202-206] burying the hydrophobic regions. Alternatively they can bind strongly at amphiphilic interfaces such as the surfaces of phospholipid vesicles, serum lipoproteins, or even at the air-water interface, where extremely stable monolayers are formed.[207]

Peptides can form amphiphilic α-helical structures where the distribution of the hydrophobic and hydrophilic amino acid residues in the linear sequence depends on: the size and shape of the hydrophobic domain formed; the

---

* This self-association is evident in longer sequences since they are seldom soluble in aqueous solutions.[203, 204]

repetition and frequency of such domains; and the availability for hydrogen bond formation within the helix. Peptides that are about 20 residues long and can form α-helices with a hydrophobic domain lying parallel to the helix axis along one side of the helix also self-associate.[208, 209] Discrete aggregated forms of tetramers are observed and they possess a high α-helical content, although the monomeric peptides have very little ordered structure in aqueous solution. These peptides will bind to phospholipid surfaces and form stable monolayers at the air-water interface; and they behave as monomers with an α-helical structure in both situations. Increasing the fraction of the surface of the helical structure which is hydrophobic from one third to two thirds causes an increase in the strength of these interactions, since the surface tension of the peptide has now increased.[209] The incorporation of a basic residue (*e.g.* lysine or arginine) into the sequence such that it occupies a position in the centre of the hydrophobic domain of the helical structure prevents self-association, but does not effect binding to phospholipid vesicles or monolayer stability.[210] Studies of peptides in the helix promoting solvent 1, 1, 1- trifluoroethanol (TFE) show, that with suitable stabilisation by the environment, the transition from a predominantly random coil to an α-helix will occur in peptides that are about 10 to 15 residues long,[203] which is consistent with data available for the stabilisation of amphiphilic α-helices at hydrophobic - hydrophilic interfaces.[206, 208, 211, 212] 600 MHz NMR studies of caerin 1.1 (**5**) in TFE have shown that it forms an α-helix from residues Gly (1) to Val (9) and again from His (16) to Leu (25).[213] This is shown in Figure **2.2**. The central portion from Ala (10) to Pro (15) is random in orientation. Caerin 1.1 serves to illustrate that peptides may not necessarily, nor entirely orientate themselves in a complete α-helix.

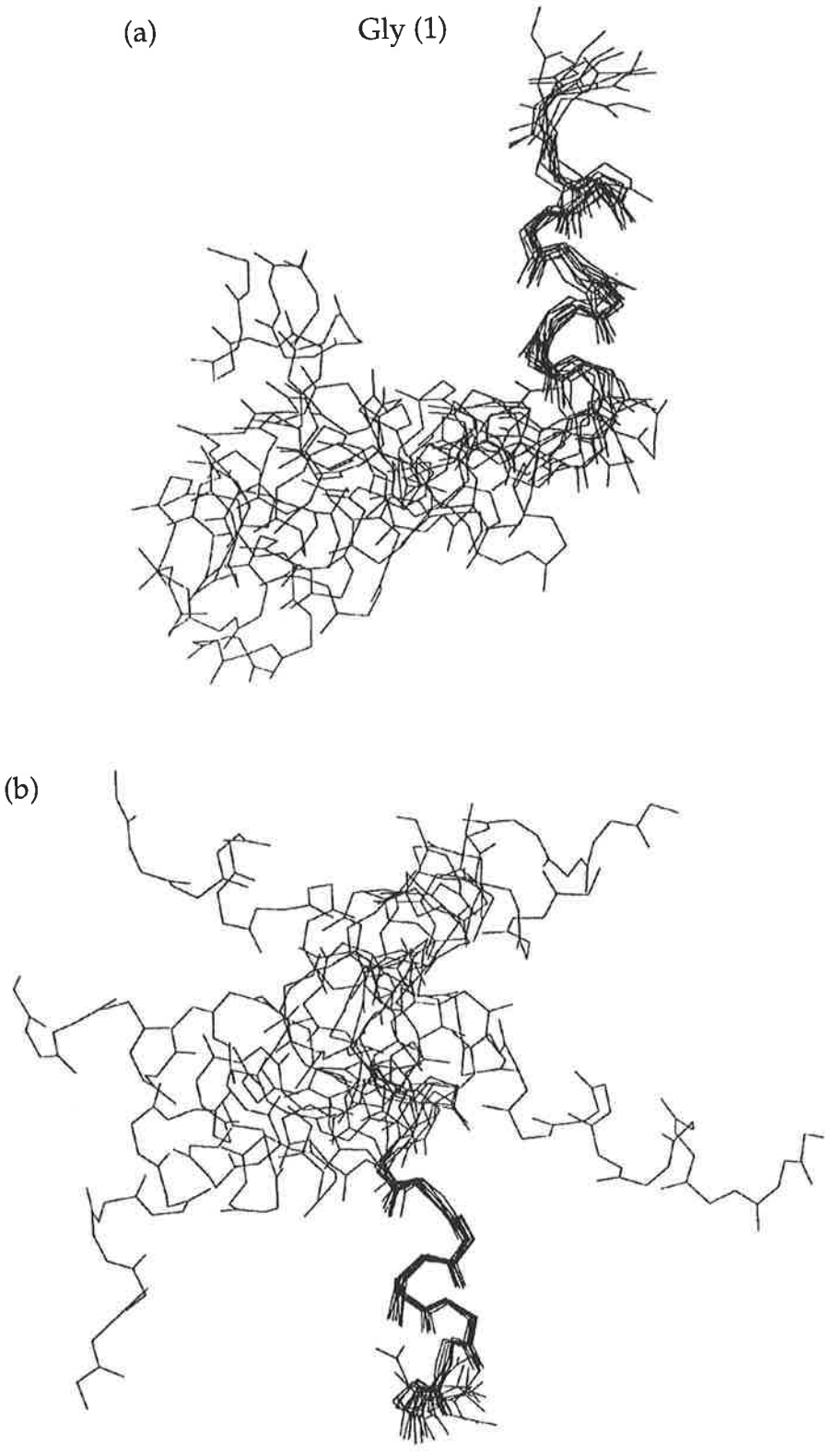


Figure **2.2**: 10 overlayed scans of caerin 1.1 using a 600 MHz NMR instrument, show that an α-helical secondary structure is attained by the amphiphilic residues (a) at the N-terminus [(1) - (9)], and (b) at the C-terminus [(16) - (25)].

The identity of the hydrophobic and hydrophilic residues which constitute an amphiphilic secondary structure is a much stronger determinant of the type of secondary structure formed in an amphiphilic environment than the periodicity with which they occur.[206, 214]* However, the amino acid content may make a significant difference to the overall stability of the structure formed. An initial search for the potential regions of amphiphilic secondary structures in peptides can be made by:

(a) scanning the linear sequences of residues for regions of alternating hydrophobic and hydrophilic residues that might form β sheets; and

(b) projecting the sequences on a helical net,[215]† or as an Edmundson wheel[216]‡ diagram to identify areas where the hydrophobic and hydrophilic residues are segregated in separate domains on the α-helix surface; many potential amphiphilic helical structures of peptides have been identified by this method. Magainin II (**9**) has been projected as the helical net and Edmundson wheel diagram in Figure **2.3**.

---

* This is true for all amino acids except for proline, which cannot participate in hydrogen bonding and also causes kinks in peptide structures.

† A helical net diagram is a two-dimensional representation of the three-dimensional structure of an α-helix and can be crudely thought of as a "cylinder sliced open and flattened". Thus the residues in the polypeptide chain can be viewed in the same plane allowing the distinction between hyrophobic and hydrophilic residues. The N-terminal end of the peptide is always located at the bottom of the page and the the point of contact is represented by the α-carbon atoms of the residues in the peptide. The parallel lines represent where complete revolutions in the helix start and finish, which equates to the width of the lines being 3.6 amino acids per revolution. The net is repeated every five revolutions.

‡ A 'helical wheel' diagram is a two-dimensional representation of the three-dimensional structure of an α-helix. The perimeter of the wheel corresponds to the backbone of the polypeptide chain. The α-side chains are projected onto the plane of the page, which is perpendicular to the axis of the helix. Adjacent residues are spaced 100° apart on the circumference since there are 3.6 residues per turn in the helix which corresponds to 18 residues per entire revolution of a helix. The α-carbon atoms of the residues in the peptide are numbered from 1 to 18 for a complete helix starting at the N-terminal end.

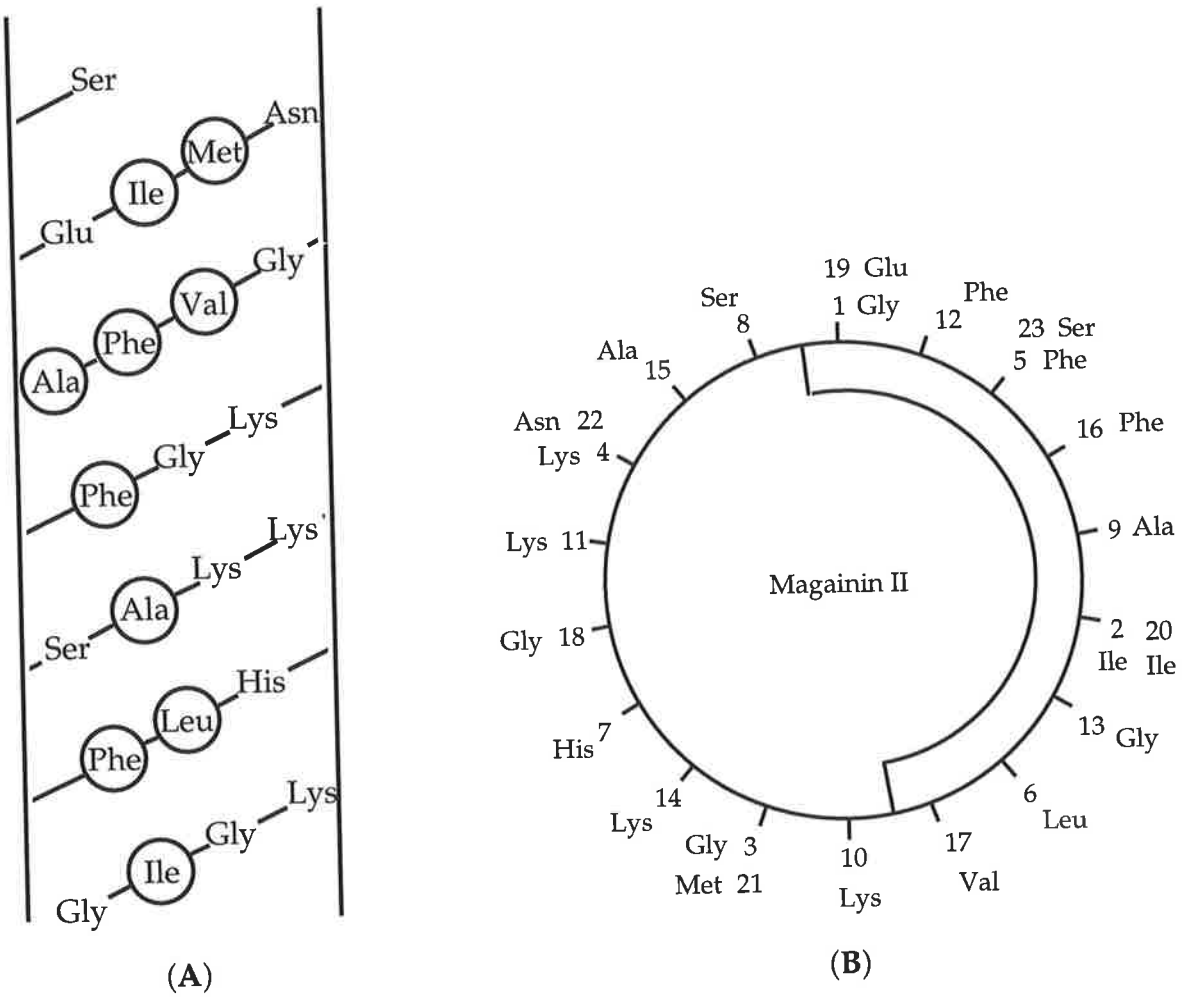


Figure **2.3**: Magainin II (**9**) illustrated in (**A**) a helical net diagram, with the hydrophobic residues circled, and (**B**) Edmundson wheel projection showing the hydrophobic region enclosed by the semi-circle.

The following postulates[207] reflect upon the properties of peptides with membrane affinity:

(a) Membrane affinity is an intrinsic property of the amino acid sequence of the peptide.

(b) Membrane affinity is determined by the ability of the peptide to assume an amphiphilic structure.

(c) The amphiphilicity of the peptide is defined by its secondary structure.

(d) The secondary structure of the peptide monomer in aqueous solution differs from that of the peptide interacting with the membrane.

(e) Peptides with membrane affinity do form micelles in aqueous solution and the conformation of the peptide in these micelles approximates that assumed in the membrane.

(f) The extent of penetration of the peptide into the lipid bilayer, ranging from weak adsorption to the disruption and fusion of the lipid bilayer, depends on the relative hydrophobic - lipophilic balance of the amphiphilic peptide.

(g) The orientation of the amphiphilic peptide with respect to the plane of the membrane lipid bilayer depends on the geometry of the hydrophobic and hydrophilic regions with respect to the secondary structural axis of the peptide.

(h) The amphiphilic and functional domains of peptides with membrane affinity are frequently independent of each other. Consequently, the structure of the amino acids in the amphiphilic region can be widely altered without noticeable loss of activity, provided that the substitutions conserve the hydrophobicity and hydrophilicity of the residues.

Hypotheses have developed regarding which segments of the sequence penetrate the membrane, resulting in a central need to be able to measure the relative affinities of the interacting hydrophobic phases. A number of hydrophobicity scales have been produced,[217-227] with each method constituting a separate operational definition; hence there is variation between scales due to the interaction of a particular amino acid residue with water. The Bull and Breese scale[218] is the hydrophobicity scale referred to when hydrophobicity is discussed in this thesis, and is reproduced as Appendix A.

Peptides with α-helical secondary structures and a characteristic balance of hydrophobicity are able to associate with lipid membranes;[228] aggregation of such peptides results in the formation of ion-channels.[229] The host cell dies as a result of an influx (or efflux) of selective ions through the newly formed opening, or as a result of a change in cell voltage.[230] The mechanism of ion-channel formation is independent of peptide chirality. Recent studies of a synthetic all D-enantiomer of magainin II compared to the natural product provides evidence to support this proposal.[231] More recently, a synthetic all D-enantiomer of caerin 1.1 has also confirmed this.[232] Results from biological testing of caerin 1.1 (**5**) show that it possesses remarkable activity against a large number of viruses and bacteria. A different mode of action may be occurring in this particular case as a modified version of caerin 1.1 (residues 3 - 25) has been found to be inactive. Loss of activity may be attributed to the peptide being unable to span the lipid bilayer; similar observations with cecropins have been reported.[233]

To be able to investigate the intimate bioscience of such a peptide, the amino acid sequence of that peptide must be determined. From a historical perspective, amphibian peptides were discovered by means of their antimicrobial activity from bioassays,[234] with the abundant, highly active peptides being identified easily. Nowadays the process of cataloguing peptides of significant abundance uses HPLC and mass spectrometric techniques; the frog peptides reported in this thesis were catalogued by the same process, where the sequence determination was predominantly by mass spectrometric investigations.

2.7 Peptide Sequencing

The three major methods* used in the sequencing of peptides are: (i) DNA sequencing;[238, 239] (ii) automated Edman degradations;[240] (iii) and mass spectrometry.

(i) DNA Sequencing

DNA sequencing is a fast and efficient method for indirectly deducing amino acid sequences and is particularly useful for large peptides and proteins. Such sequencing involves using the genetic code to represent the amino acid sequence that is derived from the translation of the ribosome messenger RNA (mRNA). The codon of each amino acid corresponds to three specific base pairs in the mRNA chain. Correct elucidation of nucleotide base sequence is vital for the determination of the protein sequence and must be interpreted without one error. The vitality of which can be illustrated by the fact that every DNA sequence has three reading frames, each coding for a different string of amino acids, giving one correct and two meaningless sequences. Consequently, some protein sequence data must be required to define the correct reading frame. Post-translational modifications are common and cannot be predicted by DNA sequencing. Such processes generate further complications in structure determination.[241] There are many known covalent post-translational modification processes,[242-245] which include carboxylation, decarboxylation, disulphide bond formation, glycosylation, hydroxylation, methylation, N-terminal acetylation, phosphorylation, proteolytic processing by specific proteases,[246] and sulphonation as examples.

---

* Crystallography (for a review of peptide crystallography see ref 235) and NMR[236, 237] can also be used for structure elucidation. These two techniques are not used as frequently as they require significantly more sample. Crystallography requires the peptide to be crystallized; NMR studies generally require a powerful instrument, such as a 600 MHz machine which is very expensive. The great advantage of these two methods is that they both can determine the stereochemistry of each individual amino acid in the peptide, thus determining whether they are D or L amino acids.

(ii) Automated Edman Degradations

Automated Edman degradations allow for the direct determination of the primary structure of a peptide or protein where the amino acids from the N-terminal are sequentially removed for analysis by HPLC. This procedure is sensitive and allows for small samples to be loaded for analysis.

Unfortunately, automated sequencing has some disadvantages such as:

(a) being unable to process peptides that are blocked at the N-terminus.[247] (These include N-terminally modified peptides[248] and secondary amides such as pyroglutamic acid, where the lone pair of electrons on the nitrogen are delocalised into the amide carbonyl, causing the initial cleavage reaction with the isothiocyanate not to occur). Secondary amines, such as proline and its derivatives, are removed much slower than amino acids which contain primary amines;

(b) it cannot correctly identify rarer amino acids or amino acids which have been modified,[244] because the corresponding phenylthiohydantoin (PTH) derivative either has an unknown retention time or is insoluble in organic solvents;

(c) retain the peptide on the membrane as it sequentially gets smaller, since a decrease in size can lead to an increase in solubility, hence sensitivity and vital information are lost.

(iii) Mass Spectrometry

Whereas DNA sequencing and automated Edman degradations are respectively viewed as biological and chemical methods of peptide sequencing, mass spectrometry is viewed as a physical method and has a number of unique capabilities. Mass spectrometry can:

(a) directly measure the molecular weight of a peptide with high accuracy

and resolution. If necessary, this can be achieved using amounts of material at the picomole or femtomole level.

(b) obtain results from complex mixtures, including those that result from enzymic digests, Edman degradations, "ladder" sequencing, and modified peptides, such as methyl esters.

(c) provide excellent sequencing information as a result of tandem mass spectrometric methods, such as CA MIKES (discussed in sections **1.9 - 1.11**), even if mixtures are used.

(d) distinguish the identity and location of post-translational modifications.

2.8 Mass Spectrometry and Peptides

Mass spectrometry is gaining in reputation as the choice method to obtain peptide and protein sequencing information. The advent of FAB heralded a new era in the way mass spectrometry could analyse biological samples, particularly peptides and proteins. In recent times, electrospray mass spectrometry has developed into a powerful method for the characterisation of peptides and proteins, as it couples HPLC with mass spectrometry.

It is possible to determine the primary structure of a peptide using only mass spectrometry, but for efficiency and clarity it is used in combination with other methods, such as chemical derivation, Edman degradations, and enzymic digests.

Typical chemical derivatives made of peptides are: methyl esters, which determine the presence and number of free carboxyl and primary amide groups in the structure; and acylated peptides where the number of primary

amine groups is determined by acylation. The detection of an increase in molecular weight determines (i) that the reaction has worked; (ii) that the specific functionality is present; and (iii) the number of specific groups modified.

Edman degradations remove one amino acid from the N-terminus of a peptide per cycle. Therefore, Edman / FAB MS[249] determines the molecular mass of the truncated peptide. The difference in masses determines which specific amino acid has been removed. Of all the common amino acids, there are only two pairs that cannot be clarified by this method. These pairs are the isomeric residues leucine and isoleucine,* and the isobaric residues lysine and glutamine. The latter of course, can be identified by chemical modification of the side chain, where lysine can be acylated, and glutamine can be converted to the methyl ester. Removal of amino acids from the C-terminus is also possible.[251-253] There are recent reports of improved methods of removing amino acids from the N-terminal end, which are based on the Edman degradation method, but differ in methodology by generating a mixture of truncated peptides, which when analysed give a series of molecular ions corresponding to sequential losses of amino acids, and ultimately the peptide sequence is read directly from the mass spectrum.[254, 255] These processes are known as peptide ladder sequencing.

Enzyme digests are extremely useful in peptide sequencing by mass spectrometry. The use of specific enzymes is indicative of the presence of specific amino acids. For example, Lys-C will only cleave a peptide at the carboxyl side of a lysine residue, and Asp-N will cleave a peptide only at the

---

* Although Edman / FAB MS cannot distinguish between leucine and isoleucine, Ramsay *et al*[250] have reported a method which distinguishes between them by MIKES analysis of the deprotonated ion of the PTH derivative following the cleavage step in the Edman degradation.

N-terminal side of an aspartic acid residue. There are many enzymes that can be used. They are classed as either endo-proteases, which are site-specific, and exo-peptidases, which sequentially degrade the peptide from either the amino- or carboxy-termini; FAB mass spectrometry is used to determine either the residue removed, the mass of the remaining peptide, or the product masses of enzyme digests.[256] This process is known as FAB mapping.[257-259]

2.9 Fragmentations of Protonated Peptides in the Gas-Phase

Tandem mass spectrometry has been discussed previously in sections **1.9 - 1.11**, and is a powerful method of obtaining peptide sequencing information.[260] Positive ionisation results in the formation of protonated gas-phase peptides which, when subjected to CA MIKES, generate a series of fragment ions. The mass difference between adjacent ion pairs reveals the identity of consecutive amino acids. There are specific sites where fragmentation may take place and consequently there are many ion series that can be generated, depending upon the type and position of certain amino acids in the peptide.[261, 262] These fragment ion series have been comprehensively studied, resulting in the fragmentation behaviour of protonated peptides being understood.[263-265] This behaviour can be rationalised using a nomenclature that identifies the processes which lead to the formation of these fragmentations.

Roepstorff and Fohlman[266] proposed a nomenclature where the fragment ions generated from any of the three different bonds along the peptide backbone were denoted by the upper-case letters A, B and C for cleavages with charge retention on the N-terminal side of the cleavage site. Similarly, the letters X, Y and Z denoted fragmentations where charge retention was on the

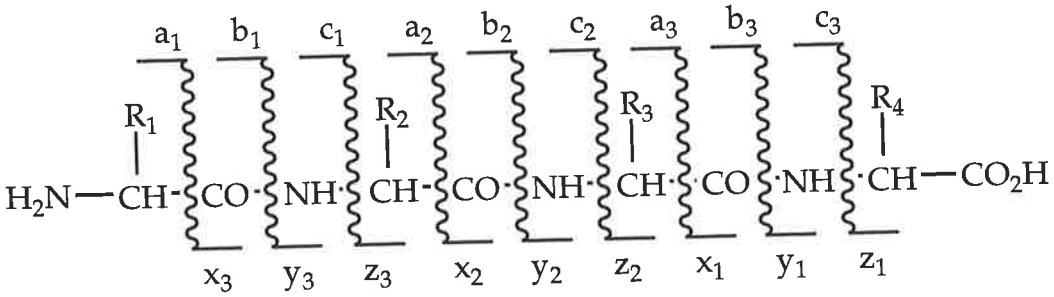
C-terminal side of the cleavage site. The addendum of prime and double prime (or + 1 and + 2) to the letters indicated the addition of the respective number of hydrogens. Biemann[261, 267] modified this nomenclature by denoting the particular fragmentations* as lower-case letters: the ions, '$Y_n + 2$' and '$C_n + 2$', are simply denoted as 'y_n' and 'c_n' as they frequently involve the addition of two hydrogens. The numerical number appearing as a subscript denotes the number of amide bonds from the terminus where the fragmentation is taking place. Side chain fragmentations occasionally observed in high energy CA spectra are 'w' and 'd' ions, which can identify specific fragmentations of the isomeric residues leucine and isoleucine[268]†. Internal fragment ions have also been reported.[271] These processes are summarised[263, 264, 272] in Figure **2.4**.

An efficient method for the complete elucidation of peptide and protein sequences integrates mass spectrometric techniques with the classical methods.[261, 273] In the past decade the Bowie research group has characterised over eighty peptide sequences from the secretions of about fifteen species of Australian frogs, mainly of the *Litoria*[192, 197, 280-288] and *Limnodynastes*[194, 289] genera; the approach involved using mass spectrometry and confirming the sequences by automated peptide sequencing: the two techniques giving complementary information.

---

* Lower case letters are used to avoid confusion with the single letter nomenclature of amino acids.

† Such ions which distinguish between the two isomers are not always observed in our spectra. Consequently we rely upon automated sequencing to identify these residues. However, it should be noted that two independent research groups have recently reported methods of differentiating between the them. Squire[269] *et al.* have reported a method using neutralisation-reionisation mass spectrometry (NRMS - for a review of NRMS see ref 270). Ramsay[250] *et al.* have analysed the negative ion MS/MS of the PTH derivative following the cleavage step in the Edman degradation.



N-terminal Ions

a_n: H-(NH-CHR-CO)$_{n-1}$-$\overset{+}{N}$H=CH(R_n)

H-(NH-CHR-CO)$_{n-1}$-NH-CH(=$CR_n'R_n''$) [H^+]

a_n+1: H-(NH-CHR-CO)$_{n-1}$-NH-CH(R_n)• [H^+]

b_n: H-(NH-CHR-CO)$_{n-1}$-NH-CH(R_n)-CO^+

c_n: H-(NH-CHR-CO)$_{n-1}$-NH_2 [H^+]

d_n: H-(NH-CHR-CO)$_{n-1}$-NH-CH(=CR_n') [H^+]

C-terminal Ions

x_n: $^+$OC-(NH-CHR-CO)$_n$-OH

O=C=N-CHR$_n$-CO-(NH-CHR-CO)$_{n-1}$-OH [H^+]

y_n: H-(NH-CHR-CO)$_{n-1}$-OH [H^+]

y_n-2: H-N=C(R_n)-CO-(NH-CHR-CO)$_{n-1}$-OH [H^+]

z_n: CH(=$CR_n'R_n''$)-CO-(NH-CHR-CO)$_{n-1}$-OH [H^+]

z_n+1: •CH(R_n)-CO-(NH-CHR-CO)$_{n-1}$-OH [H^+]

v_n: HN=CH-CO-(NH-CHR-CO)$_{n-1}$-OH [H^+]

w_n: CH(=CR_n')-CO-(NH-CHR-CO)$_{n-1}$-OH [H^+]

Internal Fragment Ions

Amino-acylium ions:	H_2N-CHR$_{n-1}$-CO-NH-CHR$_n$-CO^+
Amino-immonium ions:	H_2N-CHR$_{n-1}$-CO-$\overset{+}{N}$H=CHR$_n$
Amino acid immonium ions:	$\overset{+}{H_2}$N=CHR

Figure **2.4**: Peptide fragment ions observed in positive ion FAB CA MS/MS spectra, where 'R' represents the side chains of the amino acids and 'R_n' ' and 'R_n" ' are substituents, if any, at the β-carbon of the n^{th} amino acid.

2.10 Reported Peptides from Australian Frogs

Erspamer has investigated frog skin peptides from different continents.[274] There are suggestions that amphibian peptides are of interest from a taxonomical and evolutionary point of view, as restricted groups of amphibians living in far distant geographic areas may contain closely related or identical peptides. There are more than 208 species of known frogs in Australia.[279] The investigation into Australian frog skin peptides led to the isolation of the neuropeptides caerulein (**6**),[193] from *Litoria caerulea* and uperolein (**14**),[275] from *Uperoleia rugosa*.

The biological and pharmacological activity of caerulein shows that there is wide spectrum activity, which resembles that of the gastrointestinal hormones gastrin and cholecystokinin, and has clinical applications.[160, 276, 277] Caerulein has also been isolated from the skins of a South American frog *Leptodactylus pentadactylus labyrinthicus*, and the African clawed frog *Xenopus laevis*.[278]

Uperolein is a member of the tachykinin family of peptides. Tachykinins are fast acting hypotensive peptides and are widely found in mammals, amphibians and octopi. Since the tachykinins are of particular relevance to the subject matter of this thesis, the structures of a select few are listed in Table **2.2**.

Table **2.2**: The Structures of a selection of Tachykinins.

Tachykinin	Sequence [a]
(**11**) Physalaemin	pGlu Ala Asp Pro Asn Lys **Phe** Tyr **Gly Leu Met** (NH_2)
(**14**) Substance P	Arg Pro Lys Pro Gln Gln **Phe** Phe **Gly Leu Met** (NH_2)
(**15**) Uperolein	pGlu Pro Asp Pro Asn Ala **Phe** Tyr **Gly Leu Met** (NH_2)
(**16**) Eledoisin	pGlu Pro Ser Lys Asp Ala **Phe** Ile **Gly Leu Met** (NH_2)
(**17**) Phyllomedusin	pGlu Asn Pro Asn Arg **Phe** Ile **Gly Leu Met** (NH_2)
(**18**) [Lys^5, Thr^6]-(**11**)	pGlu Ala Asp Pro Lys Thr **Phe** Tyr **Gly Leu Met** (NH_2)

[a] The general sequence is highlighted.

2.11 Tachykinins

Erspamer's investigation into the skins of *Uperoleia* species led to the discovery of uperolein (**15**), [Lys^5, Thr^6] physalaemin (**18**) and the isolation of litorin (**2**) and [Glu (OEt)] litorin (**19**).[187]

pGlu Gln Trp Ala Val Gly His Phe Met (NH_2) (**2**)

pGlu Glu(OEt) Trp Ala Val Gly His Phe Met (NH_2) (**19**)

Litorin and [Glu(OEt)] litorin are members of the bombesin family of peptides; whereas uperolein (or [Pro^2, Ala^6] physalaemin) and [Lys^5, Thr^6] physalaemin (**18**) are members of the tachykinin family of peptides, and are structurally related to physalaemin (**11**) isolated from the skin of the South American *leptodactylidae* frog *Physalaemus fascumaculatus*. Bioassays show that uperolein has activity similar to physalaemin and exhibits potent vasodilation and hypotensive action together with intense spasmogenic activity of smooth muscle.[159]

A large number of tachykinin analogues have been synthesised and examined pharmacologically by several workers[290-298] in an attempt to elucidate the structure - activity relationships within this group. The observations and conclusions from these studies are pertinent to the content of this thesis and are briefly discussed below (see ref 160 for a more detailed discussion of these observations).

(i) Shortening of the tachykinins by successive elimination of the N-terminal amino residues down to the C-terminal hexapeptide, left the hypotensive effect essentially unchanged. A minimum of five amino acid residues was necessary for a detectable, but very low, activity.

(ii) The C-terminal amide group is apparently not essential for biological activity. In the C-terminal hexapeptide, methionine-nitrile could replace the methioninamide residue with no loss in hypotensive activity.

(iii) Changes in amino acid residues gave rise to erratic results. In general, any change in the C-terminal tripeptide or in the residue at position 5 (from the C-terminus) caused an enormous decline in activity. Conversely, replacing the methioninamide with ethioninamide caused a considerable enhancement of most activities. Replacing alanine at position 6 (from the C-terminal) with lysine enhanced the biological activity by increasing the affinity of the new peptide for vascular smooth muscle.

(iv) The entirely D-enantiomer of the hexapeptide was inactive of itself and did not antagonise the L-enantiomer. Replacement of an L-amino acid residue with the corresponding D-enantiomer in the hexapeptide caused a striking decrease in activity.

(v) Substitutions of various amino acids in various positions yielded unpredictable results: the biological activity either diminished or disappeared, or a definite potentiation of activity was observed.

The comparative potencies of the tachykinins vary for different *in vivo* and *in vitro* assays, with physalaemin being the most potent, uperolein being most closely chemically and biologically related to it, and substance P the most dissimilar.[159] All of the amphibian tachykinins behaved qualitatively in the same manner, the order of potency being: physalaemin > phyllomedusin > uperolein. The vascular effects induced by tachykinins can be obtained with amounts as low as 10^{-12} to 10^{-13} g/Kg.[159] The physiological effects of tachykinins have been extensively studied,[290-298] and a brief synopsis is given here.

Experiments performed in human subjects showed that tachykinins, given at a rate of 0.6 μg/min, caused a transient fall in mean arterial pressure (tachycardia). Higher doses (2 to 5 μg/min) caused similar but more pronounced changes, with side effects including generalised intense erythema, burning of the eyes and dizziness. They are also a powerful dilator of vessels in skin and skeletal muscle. Higher doses produced an erratic hypertension which could be completely blocked (or reversed) by sympatholytic drugs.[299]

Tachykinins are completely devoid of toxicity in acute experiments and doses as large as one million times the hypotensive threshold dose could be given with full recovery by the animals: this appears to be a unique example of tolerance. They are also highly active in modifying brain circulation. Even though the most prominent activity of the tachykinins is the vasodilating, hypotensive action, a practical application in the treatment of hypertensive conditions has not been found because of the brief duration of the effect and the possible changes in vascular permeability. Tachykinins produce a strong

stimulation of secretions from the salivary and lachrymal glands; as a consequence, clinical applications have been tested for the treatment of Sjöegren's syndrome and other types of *keratoconjunctivitis sicca*.[300-302]

The biological importance of amphibian tachykinins, which are present in high concentration in amphibian skin, is poorly understood. It has been proposed that tachykinins are major transmitters of the amine precursor uptake system,[303] and that these peptides could be expressed in tissues derived from the neuroectoderm, such as the cutaneous skin glands. It is proposed that the tachykinins synthesised in the skin may act as neuromodulators. Alternatively, they may act locally in the regulation of water and electrolyte balance as well as being involved in defense against infection of predatory attacks.[174] Much of the current research into tachykinins investigates the action of the mammalian analogue, substance P (for a review of mammalian tachykinins see ref 179). There is evidence to suggest that the presence of substance P in the frog spinal cord indicates that tachykinins may play a neurotransmitter role in the primary sensory afferents, as has been established in mammals. An increasing body of evidence indicates that, in mammals, substance P is implicated in nociception.[304-308] In amphibia, electrophysiological studies support the hypothesis that substance P is a putative transmitter involved in segment reflex activity[308] and has a direct depolarising action on motoneurons.[309]

Apart from tachykinins, there are a number of amphibian peptides which have been, or are currently being assessed for their viability for clinical applications. There are too many to mention here, however, the determination for finding clinical uses of amphibian peptides is gaining momentum and world wide attention. For example, the majority of the highly acclaimed magainins are already into the next phase of drug assessment.[155]

Chapter 3 *UPEROLEIA INUNDATA*

"Sweet are the uses of adversity,
Which, like the toad, ugly and venomous,
Wears yet a precious jewel in his head...."
W. Shakespeare, *As You Like It*, (II, i).(310)

3.1 Historical Aspects of the Genus *Uperoleia*

Prior to 1980, two geographically widespread genera of small, glandular-skinned, fossorial frogs were recognised in Australia: *Uperoleia* and *Glauertia*. Together they formed one of the most confusing and inadequately defined components of the Australian anuran fauna. For example, there were two principal problems concerning the genus *Uperoleia*: (i) that the concepts of *Uperoleia* hinged largely on morphological data derived from 'abundant' specimens of *U. marmorata* collected in Western Australia;(311) and (ii) that the locality descriptions of where these specific specimens were collected was imprecisely recorded: a fact which is reflected by this species being reported only twice from Western Australia between 1841 and 1980.(312, 313) Therefore, the concepts of the genus may well have been used to describe similar frogs found several thousand kilometres away in eastern Australia.(312, 314)

The genus *Glauertia* was described by Loveridge in 1933(315) with minimal attention to morphological detail. The taking of measurements as well as the anatomical examination for generic characters was rendered difficult by the shrivelled condition of the series, which apparently has been attributed to the direct result of the specimens being immersed in a concentrated alcoholic solution.

Many herpetologists have attempted to redefine the two genera,[312, 314, 316-319] culminating in the following events:

(i) In 1933, *Glauertia russelli* was classified as a distinct species based on the presence of partly webbed toes, which distinguished it from the *Uperoleia* species that were then known to lack interdigital webbing;[315]

(ii) In 1940, *Glauertia* was redefined to include *Pseudophryne mjobergi** and a new species *G. orientalis*: both having webbed toes and extensively exposed *frontoparietal fontanelles* (forehead);[312]

(iii) In 1971, *G. mjobergi* was transferred to *Uperoleia* by Lynch, who commented that the genera were closely related;[318]

(iv) By 1980 it had been established that there was no clear dichotomy between the genera, and no single character on which *Glauertia* could be maintained. Both the extensive webbing and the exposed *frontoparietal fontanelles* of *Glauertia* are extremes of trends apparent in *Uperoleia*. Since it was no longer possible to maintain recognition of *Glauertia*, all species in the genus were transferred to *Uperoleia*.[321]

3.2 The Genus *Uperoleia*

The genus *Uperoleia* belongs to the family of anurans known as *Leptodactylidae* and was first reported by Gray in 1841.[311] *Uperoleia* is a very conservative genus, with all species having similar body proportions. All members of this genus are burrowers, generally have warty skin and attain a maximum length of 3.5 cm, with limbs being very short. They have cryptic

---

* *Psuedophryne* and *Uperoleia* genera are very closely related.[120] This is evident from the isolation of five tachykinins from *Psuedophryne guntheri* which are structurally similar to uperolein.[320]

markings and the differences between most species are extremely subtle. Since 1980 the number of reported species in this genus has risen from 6 to 24 following intensive research[120, 321] and it is therefore understandable that past researches involving this genus may well have been complicated by misidentification. For example, Erspamer's investigation into *Uperoleia* led to the extraction of peptides from *U. rugosa* (collected from Brisbane and New South Wales) and *U. marmorata* (collected from Brisbane).[159, 187, 276] However, *U. marmorata* is now known to be confined to the Kimberley region on the other side of the continent.[322] Clearly the material examined by Erspamer and his colleagues was incorrectly identified and compounded several distinct species.*

The distribution of *Uperoleia* extends from the Pilbara across the north of Australia and down the eastern side of Australia[322] (see Figure **3.1**). It does not reach Tasmania.

The 24 species of this genus include: *U. arenicola, U. altissima, U. aspera, U. borealis, U. capitulata, U. crassa, U. fusca, U. glandulosa, U. inundata, U. laevigata, U. lithomoda, U. littlejohni, U. marmorata, U. martini, U. micromeles, U. mimula, U. minima, U. mjobergi, U. orientalis, U. rugosa, U. russelli, U. talpa, U. trachyderma* and *U. tyleri.*

We have commenced a project to identify the major peptide components of the secretions of several *Uperoleia* species, and to determine the biological activities of the peptides. Raftery[323] undertook a preliminary investigation into the secretions of *U. inundata, U. lithomoda, U. littlejohni, U. micromeles, U. mimula* and *U. rugosa*. The initial results showed that (i) these species

---

* *U. Rugosa* shares it's habitat with *U. capitulata, U. fusca,* and *U. laevigata*. All have similar markings.

Figure **3.1**: Distribution of the genus *Uperoleia* in Australia, adapted from refs.[120, 321]

have a significant number of peptides in their secretions (as determined by HPLC and MS), and (ii) *U. micromeles* contains uperolein, whereas *U. rugosa* does not! This thesis reports the results obtained from detailed studies of two *Uperoleia* species: *U. inundata* and *U. mjobergi*.

3.3 *Uperoleia inundata*

Figure 3.2: Photograph of ***Uperoleia inundata***, courtesy of M. J. Tyler, Department of Zoology, The University of Adelaide, South Australia, 5005.

The Floodplain Toadlet *U. inundata* was first described by Tyler *et al* in 1981.(321) It is a small frog attaining a length of between 1.8 and 3.2 cm and has glands distributed throughout the dorsal surface as can be seen in Figure 3.2. Figure 3.3 shows that it is confined to a region extending from the coastal areas of the Northern Territory (through Arnhem Land and Kakadu), to the Gulf country of Queensland. The seven specimens studied for this project were collected in the vicinity of Jabiru, and maintained in captivity.

Figure **3.3**: Distribution of *Uperoleia inundata*.

Thirteen peptides have now been isolated and characterised from the dorsal gland secretion of *U. inundata*. The peptides have been named uperins and have been classified into six subclasses. The amino acid sequences of the thirteen peptides are listed in Table **3.1**. The methods by which the primary structures of the uperin peptides were determined are described in section **3.4**.

The benign method of mild, surface electrical stimulation of the skin glands (see section **2.4**) was used to effect release of the glandular secretions. On average, each 'milking' produced about 3.5 mg of peptide material per frog.

Table 3.1: Amino Acid Sequence of the Uperin Peptides from *U. inundata*.

Uperin	[M + H]+ [a]	Amino Acid Sequence
1.1	1208	pGlu Ala Asp Pro Asn Ala Phe Tyr Gly Leu Met (NH_2)
2.1	1926	Gly Ile Val Asp Phe Ala Lys Lys Val Val Gly Gly Ile Arg Asn Ala Leu Gly Ile (OH)
2.2	1926	Gly Phe Val Asp Leu Ala Lys Lys Val Val Gly Gly Ile Arg Asn Ala Leu Gly Ile (OH)
2.3	1974	Gly Phe Phe Asp Leu Ala Lys Lys Val Val Gly Gly Ile Arg Asn Ala Leu Gly Ile (OH)
2.4	1913	Gly Ile Leu Asp Phe Ala Lys Thr Val Val Gly Gly Ile Arg Asn Ala Leu Gly Ile (OH)
2.5	1940	Gly Ile Val Asp Phe Ala Lys Gly Val Leu Gly Lys Ile Lys Asn Val Leu Gly Ile (OH)
3.1	1827	Gly Val Leu Asp Ala Phe Arg Lys Ile Ala Thr Val Val Lys Asn Val Val (NH_2)
3.2	1841	Gly Val Leu Asp Ala Phe Arg Lys Ile Ala Thr Val Val Lys Asn Leu Val (NH_2)
3.3	1813	Gly Val Leu Asp Ala Phe Lys Lys Ile Ala Thr Val Val Lys Asn Leu Val (NH_2)
4.1	1724	Gly Val Gly Ser Phe Ile His Lys Val Val Ser Ala Ile Lys Asn Val Ala (NH_2)
5.1	1456	Phe Gln Phe Val Asn Pro Ser Asp Ile Val Phe Gly Ser (OH)
6.1	3230	Gly Leu Ala Gly Ala Ile Ser Ser Ala Leu Asp Lys Leu Lys Gln Ser Gln Leu Ile Lys Asn Tyr Ala Lys Lys Leu Gly Tyr Pro Arg (OH)
6.2	3258	Gly Leu Ala Gly Ala Ile Ser Ser Val Leu Asp Lys Leu Lys Gln Ser Gln Leu Ile Lys Asn Tyr Ala Lys Lys Leu Gly Tyr Pro Arg (OH)

[a] Nominal mass

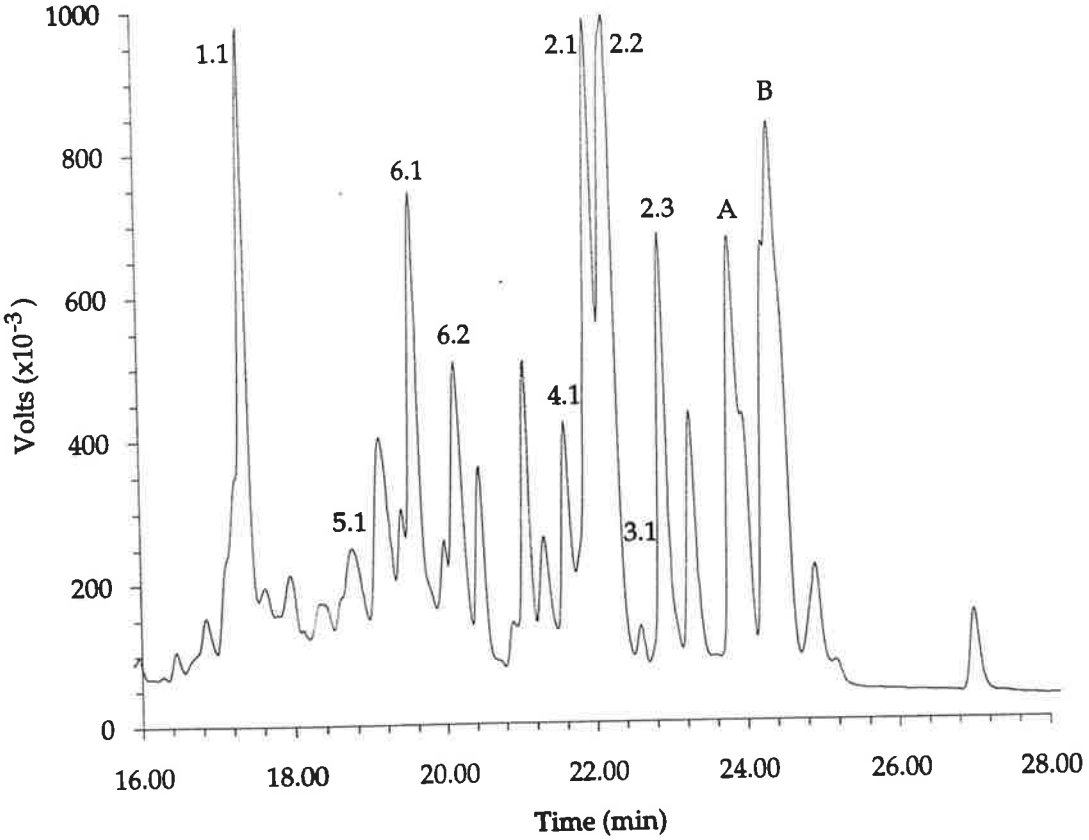


Figure 3.4: Analytical HPLC trace of peptides from the glands of *U. inundata*. For full experimental details see Experimental, section **6.3**. The labelled peaks are unresolved and correspond to the coelution of [**A**] uperins 3.2 and 3.3, and [**B**] uperins 2.4 and 2.5.

A typical HPLC chromatogram of the peptide material is shown in Figure **3.4**. It is a complicated chromatogram, with more than 50 peptides being present (as determined by FAB MS in the 1000 to 3500 Da mass range of individual HPLC peaks). Combined MS and HPLC data indicate that eight of these are major components, another fifteen are minor constituents, while the remainder are trace components. HPLC data also indicate that each 'milking' of a single frog produces 100 - 200 μg of the major components and 25 - 50 μg of each of the minor peptides. This chapter reports the structures of the eight major peptides and five of the minor components. The majority of peptides have molecular weights between 1200 and 2000 Da, with two having molecular weights greater than 3200 Da.

3.4 Sequence Determination

The sequences of the uperins (except uperins 6.1 and 6.2) were determined primarily by FAB MS and associated techniques. Positive ion FAB MS determined the molecular weights of each peptide and modified derivatives (as appropriate). Manual Edman / FAB MS was used to identify a number of amino acids from the N-terminal end of each peptide. Esterification / FAB MS was used to determine the identity of the C-terminal group and to specify how many amide and carboxylate side chains are present. An increase in mass by 14 or 15 Da (or multiples thereof), respectively indicates the presence of CO_2H and $CONH_2$ groups. Enzymic digest experiments provided essential data in most cases. A successful digest indicated the presence of a particular amino acid and cleaved the peptide into smaller segments; whereas an unsuccessful digest suggested that the specific residue under investigation was not present. The above methods in conjunction with CA MS/MS data generally provided the complete sequence of the peptide; with the exception that the isomeric residues leucine and isoleucine were not distinguished by these methods.* Automated sequencing clarified this issue in all cases and also confirmed the proposed sequence.† In the case of uperins 6.1 and 6.2, we decided that it would be inefficient to use the ZAB 2HF instrument to sequence the peptides. Instead, their molecular weights were confirmed using the matrix assisted laser desorption ionisation (MALDI) time of flight (TOF) mass spectrometer[328] at the University of California, San Francisco, and their amino acid sequences determined solely using the automated Edman procedure.

---

* MS/MS data can on occasions differentiate between leucine and isoleucine by the loss of C_3H_6 from the former.[263, 268, 269] Such fragmentations are not always observed in the spectra of the uperins. For brevity, leucine is specified as the residue observed by mass spectrometry, with the clear understanding that this means either. The correct structures are listed in Table **3.1**.

† It has yet to be determined whether the amino acid residues are composed of the 'D' or 'L' configurations. Most amphibian peptides contain only 'L' amino acid residues. Exceptions are the opioid peptides, the dermorphins[324-326] and deltorphins[327] which each contain 'D' Ala and 'D' Met respectively at position 2. Synthetic peptides with the 'L' amino acids have the same retention times to the natural peptides on the HPLC column.

Uperin 1.1

(a) Uperin 1.1 is an endecapeptide forming a parent ion $[M + H]^+ = 1208$.* It has an HPLC retention time of 17.3 min (see Figure **3.4**). Since it coelutes with at least 10 other minor peptides, it was repurified by HPLC before further sequence determinations were carried out.

(b) The conventional FAB mass spectrum (m/z range 1 300 to 500) shows significant fragmentation which allows for almost complete peptide sequencing. This is one of a very few peptides which we have isolated which shows this feature: all other uperin peptides (except uperin 4.1) give poor conventional mass spectra, and are consequently sequenced using tandem MS/MS techniques. The FAB mass spectrum of uperin 1.1 (Figure **3.5**) indicates the sequence to be pGlu Ala Asp [Pro + Asn] Ala Phe Tyr Gly Leu (NH_2), with the relative positions of Pro and Asn undetermined.

(c) Manual Edman / FAB MS experiments proved to be unsuccessful, suggesting that the N-terminal residue of the peptide was blocked by a pyroglutamic acid (pyrrolidone carboxylate) residue.

(d) Esterification of uperin 1.1 increased the mass by 76 Da, which is accounted for by the presence of the pyroglutamate residue together with two amides and one acid group. Since the sequence is known to contain one Asp and one Asn residue, the C-terminus must contain a $CONH_2$ group.

---

* Masses of protonated peptides are given as nominal masses: *i.e.* the mass given by the summation of the integral masses of the individual amino acid residues.

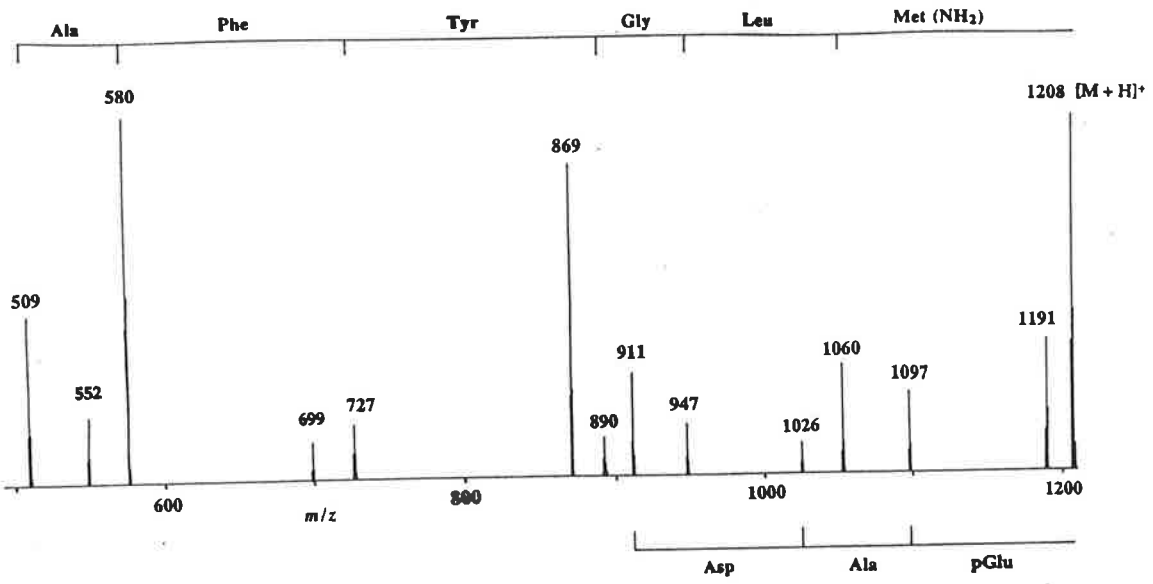


Figure **3.5**: Conventional FAB mass spectrum of uperin 1.1. Relative abundance of peak / magnetic sector plot with the masses of peaks indicated in Daltons. See Experimental (section **6.1**) for details. The schematic arrows shown at the top of the spectrum indicate data for 'b' ions, while those at the bottom provide sequence information from 'y' ions. Peaks at m/z 552 and 699 are respectively 'a_6' and 'a_7' ions.*

(e) Uperin 1.1 is thus likely to have the sequence shown in Table **3.1**, but it is necessary to confirm this structure since there are tachykinins with Asp (3) and Asn (5), and also Asn (3) and Asp (5).[320] To clarify the permutation, uperin 1.1 was digested with endoprotease Asp-N (which cleaves at the N-terminal side of aspartic acid residues), giving peptide **1.1A**

* The intense ion at m/z 869 observed in the FAB mass spectrum of uperin 1.1 is considered to be a 'w_7' ion, formed by cleavage through the proline residue.

($[M + H]^+ = 1026$).* Formation of this peptide involved the removal of the first two residues of uperin 1.1, *viz* pGlu and Ala. This identifies Asp (3), and confirms the sequence of uperin 1.1 (Table **3.1**).

(f) Automated sequencing was required to determine whether residue 10 is Leu or Ile and to confirm the overall sequence. This was achieved using peptide **1.1A** since it was unblocked at the N-terminal residue and both Asp and Asn were unmodified.† The amino acid sequence of uperin 1.1 is listed in Table **3.1**. Uperin 1.1 was synthesised (commercially): the natural and synthetic peptides have identical FAB mass spectra and HPLC retention times.

Uperin 2.1

(a) Uperin 2.1 has a HPLC retention time of 21.9 minutes (see Figure **3.4**). It coelutes with uperin 2.2 and three other trace peptides. Uperin 2.1 was repurified and separated from uperin 2.2 by HPLC (for details see Experimental, section **6.3**). FAB MS gave a parent ion $[M + H]^+ = 1926$.

(b) Uperin 2.1 is cleaved by Lys-C to form two peptides:

2.1A ($[M + H]^+ = 749$); and

2.1B ($[M + H]^+ = 1068$).

---

* The cleaved dipeptide corresponding to pGlu Ala OH ($[M + H]^+ = 201$) was not observed due to masking by the FAB matrix.

† Convertion of uperin 1.1 to the methyl ester, ring opens the pyroglutamate (to methyl glutamate) which is able to be analysed by automated sequencing procedures. However, the Asp and Asn residues would be converted to methyl aspartate and thus the identity of these residues could not be clarified by this method.

(c) The first eight residues from the N-terminal end of uperin 2.1 are Gly Leu Val Asp Phe Ala Lys Lys as shown by the sequential manual Edman / FAB MS method. Thus the Lys-C enzyme cleaved uperin 2.1 at Lys (7) and Lys (8). Manual Edman / FAB MS also shows that the first four residues of peptide **2.1B** are Val Val Gly Gly which leaves a residual peptide $[M + H]^+ = 756$, *i.e.* peptide **2.1C**.

(d) Peptides **2.1A, 2.1B** and **2.1C** were sequenced by CA MS/MS. The data which provides sequence information is listed in Table **3.2**. From these experiments it is clear that peptide **2.1A** corresponds to the first seven residues of uperin 2.1, and that a culmination of the data from peptides **2.1B** and **2.1C** give the sequence of peptide **2.1B** as Val Val Gly Gly Leu Arg Asn Ala Leu Gly Leu (OH). A representative spectrum of peptide **2.1B** is pictured in Figure **3.6**.

Table **3.2**: CA MS/MS Data for Lys-C Products from Uperin 2.1.

Peptide	*m/z*	Ion	Observed fragment ions *m/z* (%)	Peptide Sequence
2.1A	749	'b'	270 (17); 385 (27); 532 (55); 603 (30); 731 (100)	Val Asp Phe Ala Lys
		'y'	692 (37); 579 (26); 480 (26); 365 (27); 218 (20)	Gly Leu Val Asp Phe
Sequence of **2.1A** by CA MS/MS data: **Gly Leu Val Asp Phe Ala Lys (OH)**				
2.1B	1068	'a'	554 (10); 668 (14); 739 (16); 852 (23)	Asn Ala Leu (170)
		'b'	582 (11); 696 (13); 937 (51); 1050 (100)	Asn (241) Leu
		'y'	969 (25); 870 (23); 813 (20); 643 (23)	Val Val Gly (170)
Partial sequence of **2.1B**: **Val Val Gly (170) [Arg] Asn Ala Leu Gly Leu (OH)**				
2.1C	756	'b'	384 (12); 455 (22); 568 (17); 625 (39); 738 (100)	Ala Leu Gly Leu
		'y'	643 (58); 302 (19)	Leu (341)
Partial sequence of **2.1C** by CA MS/MS data: **Leu (341) Ala Leu Gly Leu (OH)**				

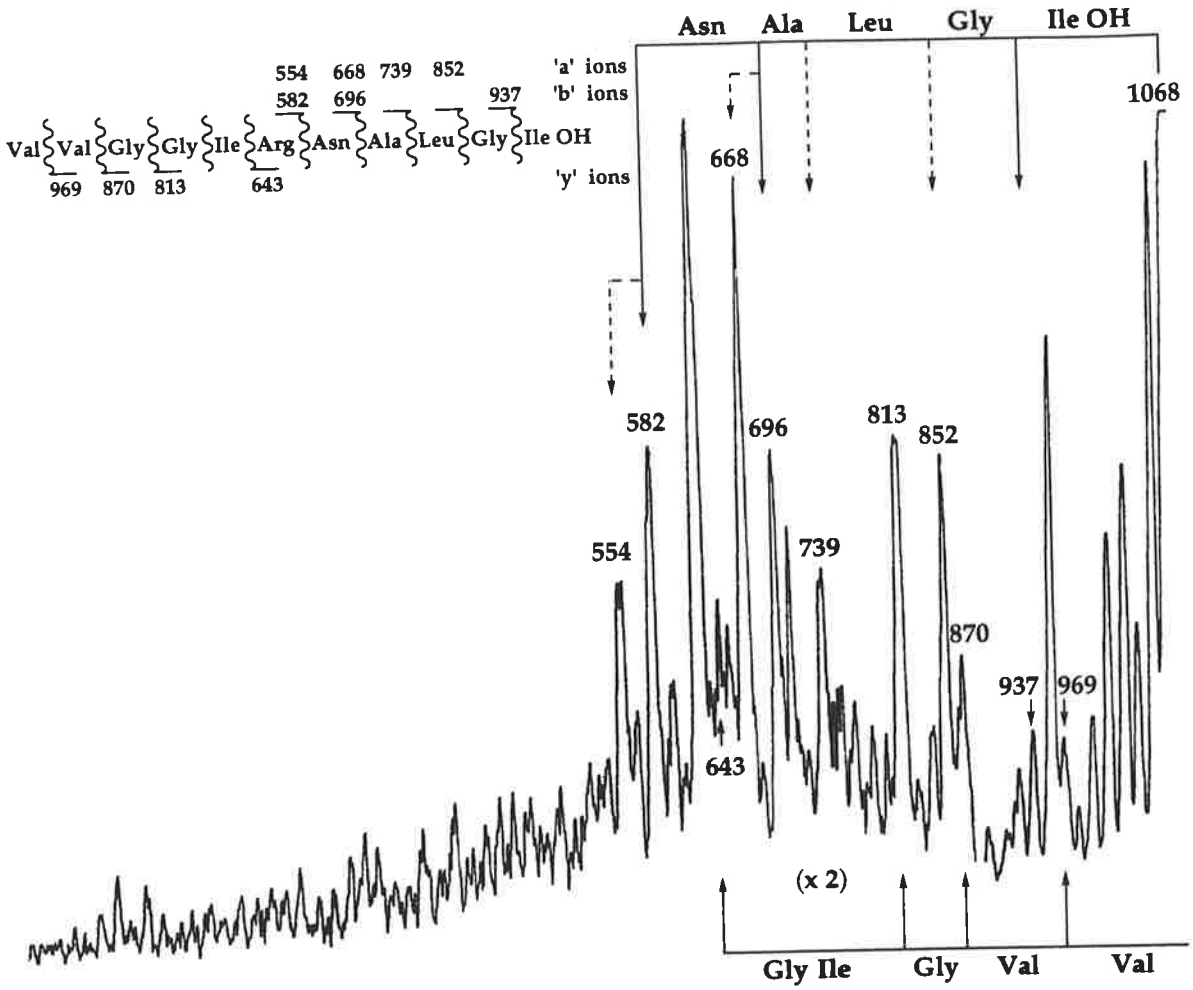


Figure **3.6**: CA FAB MS/MS spectrum of peptide **2.1B** ($[M + H]^+$ = 1068) derived from the Lys-C digest of uperin 2.1. The schematic arrows at the top and bottom represent the 'b' and 'y' cleavage ions respectively. The spectrum is void of information below m/z 500. Characterisation is aided by the use of 'a' ions (see Table **3.2**).

(e) Methylation of uperin 2.1 gives a tris methyl ester ($[M + H]^+$ = 1969) indicating the presence of one amide and two acid groups in the peptide. The sequence data indicate the presence of Asp and Asn residues, thus the C-terminal residue must contain a carboxyl (CO_2H) group.

(f) Combination of all of the above data provide the overall structure of uperin 2.1 except that isomeric Leu and Ile are not differentiated. Automated sequencing indicates Leu (17), Ile (2, 3 and 19). The structure of uperin 2.1 is as shown in Table **3.1**.

Uperin 2.2

(a) Uperin 2.2 forms a parent ion of $[M + H]^+ = 1926$ and is isomeric with uperin 2.1. It has a retention time of 22.2 minutes and coelutes with uperin 2.1 (see above). The structure determination of uperin 2.2 is the same as that for uperin 2.1 with the following exceptions. The only difference between the two structures is at residues 2 and 5.

(b) Manual Edman / FAB MS shows the first eight residues of uperin 2.2 to be Gly Phe Val Asp Leu Ala Lys Lys.

(c) Treatment of uperin 2.2 with Lys-C gives two peptides:

2.2A ($[M + H]^+ = 749$); and

2.2B ($[M + H]^+ = 1068$).

(d) Peptide **2.2A** corresponds to the first seven residues of uperin 2.2 as confirmed by the CA MS/MS data listed in Table **3.3**. Peptide **2.2B** has the same mass spectrum as peptide **2.1B** (see Table **3.2**).

Table **3.3**: CA MS/MS Data for Lys-C Products from Uperin 2.2.

Peptide	***m/z***	**Ion**	**Observed fragment ions *m/z* (%)**	**Peptide Sequence**
2.2A	749	'b'	205 (88); 304 (39); 419 (99); 532 (52); 603 (43); 731 (100)	Val Asp Leu Ala Lys
		'y'	692 (78); 545 (15); 446 (37); 331 (35); 218 (40)	Gly Phe Val Asp Leu
Sequence of **2.2A** by CA MS/MS data:				**Gly Phe Val Asp Leu Ala Lys (OH)**

(e) Methylation of uperin 2.2 gave a tris methyl ester $[M + H]^+ = 1969$, and indicates that there are two CO_2H and one $CONH_2$ groups present.

(f) Uperin 2.2 therefore has the structure shown in Table **3.1.** This structure has been confirmed by automated sequencing and synthesis.

Uperin 2.3

(a) Uperin 2.3 elutes from the HPLC column with a retention time of 22.9 mins (see Figure **3.4**) and forms a parent ion $[M + H]^+ = 1974$.

(b) Manual Edman / FAB MS of uperin 2.3 determined that the first five amino acids from the N-terminal to be Gly Phe Phe Asp Leu and that there are two Lys residues in the molecule.* The mass of the peptide after the first manual Edman degradation increased by 213 Da, which is accounted for by the addition of two PITC units (nominal mass: 135 Da) to two lysine residues, with the loss of one Gly residue (from the N-terminal end of the peptide).

(c) Lys-C digest of uperin 2.3 yielded peptides:

2.3A ($[M + H]^+ = 797$); and

2.3B ($[M + H]^+ = 1068$).

therefore there must be two Lys residues adjacent to each other.†

(d) CA MS/MS studies show that **2.3A** has the sequence Gly Phe Phe Asp Leu Ala Lys (OH), *i.e.* the first seven residues of uperin 2.3; peptide **2.3B** is identical to peptides **2.1B** and **2.2B**. These results are tabulated in Table **3.4**, and Figure **3.7** is an example spectrum of peptide **2.3**.

(e) Esterification of uperin 2.3 gave a product with a mass 43 Da greater than that of uperin 2.3 indicating the presence of one $CONH_2$ and two CO_2H groups.

(f) Combination of the above data together with automated sequencing gives the structure of uperin 2.3 which is listed in Table **3.1**.

---

* The coupling reagent, phenylisothiocyanate (PITC), is specific for primary amines. The α-side chain of lysine terminates with a primary amine, which can couple with PITC. No further reaction occurs, thus an increase in mass is observed.

† A peak in the full scan FAB mass spectrum at m/z 925 (25% intensity of m/z 797) is associated with '797 + Lys', where the Lys-Lys bond was not completely digested.

Table **3.4**: CA MS/MS Data for Lys-C Products from Uperin 2.3.

Peptide	*m/z*	Ion	Observed fragment ions *m/z* (%)	Peptide Sequence
2.3A	797	'b'	205 (18); 352 (25); 467 (100); 580 (49); 651 (44); 779 (81)	Phe Asp Leu Ala Lys
		'y'	740 (20); 593 (33); 446 (44); 331 (44); 218 (16)	Gly Phe Phe Asp Leu
Sequence of **2.3A** by CA MS/MS data:				**Gly Phe Phe Asp Leu Ala Lys (OH)**

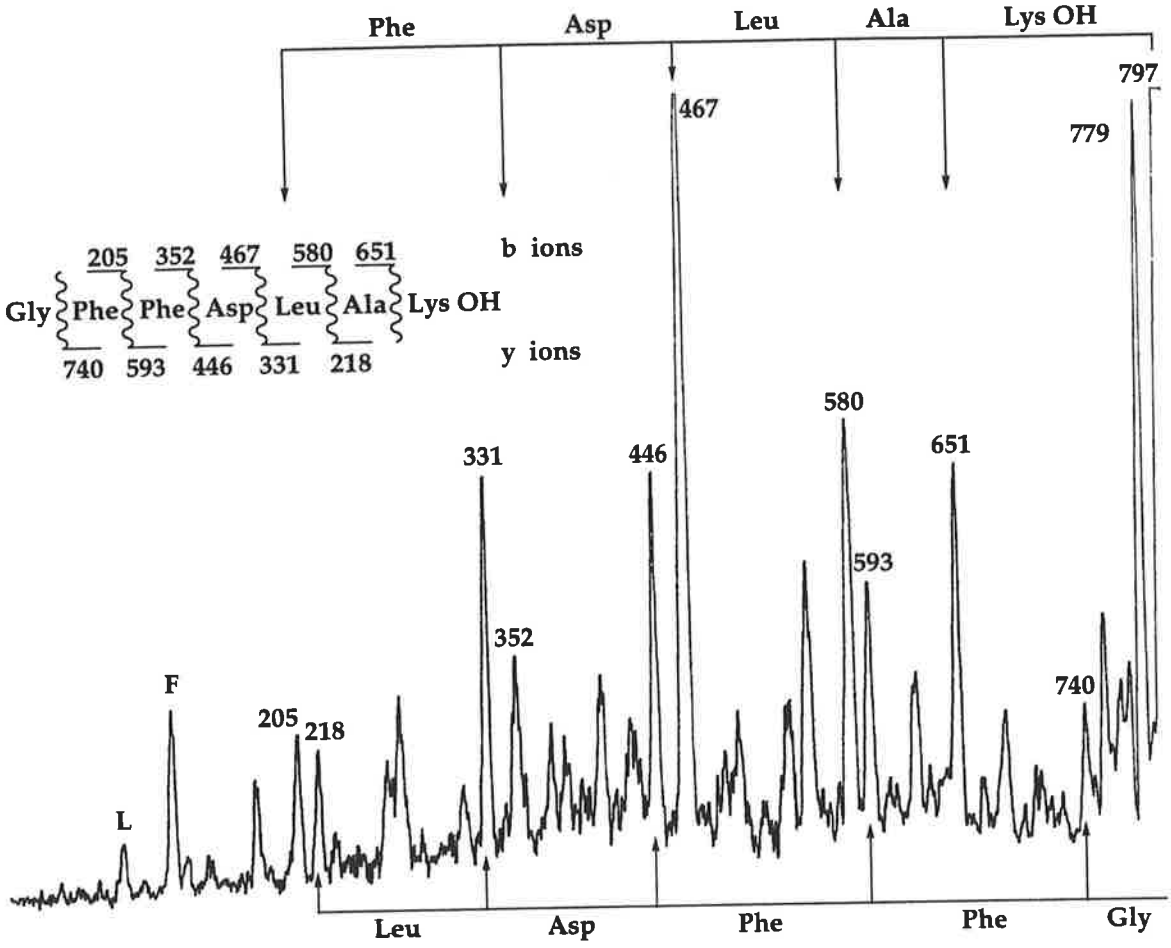


Figure **3.7**: CA FAB MS/MS spectrum of peptide **2.3A** ($[M + H]^+ = 797$), formed from the Lys-C digest of uperin 2.3. The schematic arrows at the top of the spectrum indicate the sequence derived from the 'b' cleavages, while those at the bottom relate to sequence information from 'y' cleavages. Peaks labelled **F** (m/z 120) and **L** (m/z 86) are associated with the respective formation of internal fragment ions of phenylalanine and leucine.

Uperin 2.4

(a) Uperin 2.4 elutes from the HPLC column with a retention time of 24.3 mins (Figure **3.4**), and has a parent ion of $[M + H]^+ = 1913$. It coelutes with uperin 2.5 and a trace peptide. Repurification by HPLC separated these three peptides.

(b) Uperin 2.4 afforded two peptides when treated with Lys-C enzyme. The results of this digest are listed in Table **3.5**.

Table **3.5**: Results of the Enzymic Digests of Uperin 2.4.

Peptide Digested	Enzyme Used	Resultant Peptide	$[M + H]^+$
Uperin 2.4	Lys-C	**2.4A**	763
		2.4B	1169
2.4C	Arg-C	**2.4D**	600
		2.4E	487

(c) Manual Edman / FAB MS indicates the first three residues of uperin 2.4 to be Gly Leu Leu. Manual Edman / FAB MS determined that the first two residues from the N-terminus of **2.4B** are Thr Val; the first degradation resulting in a residual peptide $[M + H]^+ = 1068$, *i.e.* **2.4C**. Peptide **2.4C** was digested with Arg-C yielding two smaller peptides **2.4D** and **2.4E** (Table **3.5**).

(d) Esterification of peptide **2.4A** gave an increase in mass of 28 Da indicating that two CO_2H groups are present. Since **2.4A** is a product of a Lys-C digest, then one of these CO_2H groups must be at the C-terminus of **2.4A**. Esterification of peptide **2.4B** gave an increase in mass of 29 Da, indicating that one amide and one acid group are present.

(e) CA MS/MS studies related to uperin 2.4 are reported in Table **3.6** and show that:

(i) **2.4A** has the sequence Gly Leu Leu Asp Phe Ala Lys (OH);

(ii) **2.4D** proved to be inconclusive;

(iii) **2.4E** is Asn Ala Leu Gly Leu (OH), *i.e.* the last five residues of uperin 2.4.

Table **3.6**: CA MS/MS Data for Lys-C Products from Uperin 2.4.

Peptide	***m/z***	**Ion**	**Observed fragment ions *m/z* (%)**	**Peptide Sequence**
2.4A	763	'b'	171 (5); 284 (7); 399 (20); 546 (40); 617 (57); 745 (100)	Leu Asp Phe Ala Lys
		'y'	706 (39); 593 (65); 480 (40); 365 (26); 218 (10)	Gly Leu Leu Asp Phe
Sequence of **2.4A** by CA MS/MS data:				**Gly Leu Leu Asp Phe Ala Lys (OH)**
2.4E	487	'b'	185 (3); 298 (8); 355 (7); 468 (100)	Leu Gly Leu
		'y'	370 (27); 299 (7); 186 (3)	Asn Ala Leu
Sequence of **2.4E** by CA MS/MS data:				**Asn Ala Leu Gly Leu (OH)**

(f) Automated sequencing of uperin 2.4 indicates the sequence to correspond to that listed in Table 3.1, and shows that residues 9-14 are Val Val Gly Gly Ile Arg (OH). Residues 9-19 are therefore identical in uperins 2.1, 2.2, 2.3 and 2.4.

Uperin 2.5

(a) FAB MS: $[M + H]^+ = 1940$.

(b) HPLC retention time: 24.5 minutes (Figure **3.4**). Uperin 2.5 coelutes with uperin 2.4 (see above) and is the major peptide in this fraction.

(c) Manual Edman / FAB MS: Gly Leu Val.

(d) The results of enzymic digests and esterification are listed in Table **3.7**. Lys-C digest of Uperin 2.5 gave peptides **2.5A, 2.5B** and **2.5C**.

Table **3.7**: Results of the Lys-C Digest and Subsequent Esterification Experiments for Uperin 2.5.

Peptide	$[M + H]^+$	esterification	Δ^*	Acids	Amides
2.5A	749	777	28	2	0
2.5B	473	487	14	1	0
2.5C	515	544	29	1	1

(e) CA MS/MS studies are recorded in Table **3.8** and Figures **3.8 - 3.10** and show that:

(i) **2.5A** is Gly Leu Val Asp Phe Ala Lys (OH), which are the first seven residues of uperin 2.5;

(ii) **2.5B** is Gly Val Leu Gly Lys (OH);

(iii) **2.5C** is Asn Val Leu Gly Leu (OH), which are the last five residues of uperin 2.5 since it contains a C-terminal Leu reside.

---

* The number of acids and amides determined in these experiments pertain to the digest fragments and not the parent peptide; thus C-terminal acids result from a successful digest.

Table 3.8: CA MS/MS Data for Lys-C Products from Uperin 2.5.

Peptide	*m/z*	Ion	Observed fragment ions *m/z* (%)	Peptide Sequence
2.5A	749	'b'	171(18); 270 (14); 385 (29); 532 (55); 603 (83); 731 (100)	Val Asp Phe Ala Lys
		'y'	692 (32); 579 (70); 480 (63); 365 (46); 218 (22);	Gly Leu Val Asp Phe
Sequence of **2.5A** by CA MS/MS data:				**Gly Leu Val Asp Phe Ala Lys (OH)**
2.5B	473	'b'	157 (9); 270 (13); 327 (30); 455 (71)	Leu Gly Lys
		'y'	416 (25); 317 (100); 204 (38)	Gly Val Leu
Sequence of **2.5B** by CA MS/MS data:				**Gly Val Leu Gly Lys (OH)**
2.5C	515	'b'	214 (11); 327 (97); 384 (64); 497 (100)	Leu Gly Leu
		'y'	401 (9); 302 (67), 189 (5)	Asn Val Leu
Sequence of **2.5C** by CA MS/MS data:				**Asn Val Leu Gly Leu (OH)**

(f) The masses of the residues identified so far fall 241 Da lower than the nominal mass of uperin 2.5; there must be an undetected portion from the Lys-C digest. The unidentified portion must contain a C-terminal Lys; this, in turn, means that the 'missing' portion is Leu Lys (OH). Two possible structures for uperin 2.5 follow from the above data, *viz*:

Gly Leu Val Asp Phe Ala Lys Gly Val Leu Gly Lys **Leu Lys** Asn Val Leu Gly Leu **2.5D**

Gly Leu Val Asp Phe Ala Lys **Leu Lys** Gly Val Leu Gly Lys Asn Val Leu Gly Leu **2.5E**

The correct structure is clarified by FAB MS analysis as the Lys-C digest progresses. The basis of this experiment relies on the fact that different lysines will be cleaved at differing rates, depending on the local and chemical environments which they are in, and their respective positions in the peptide

sequence. In such an experiment, both **2.5D** and **2.5E** will eventually be digested to give peptides **2.5A - C**, but will be distinguished by the fragments resulting from partial digest: $[M + H]^+ = 756$ for **2.5D**; $[M + H]^+ = 990$ for **2.5E**. After 100 min of digest, a peptide was identified at m/z 756, verifying that **2.5D** is the proposed structure of uperin 2.5.

(g) Automated sequencing shows Ile (2, 13 and 19), as well as Leu (13) and Lys (14), and confirms the sequence of uperin 2.5.

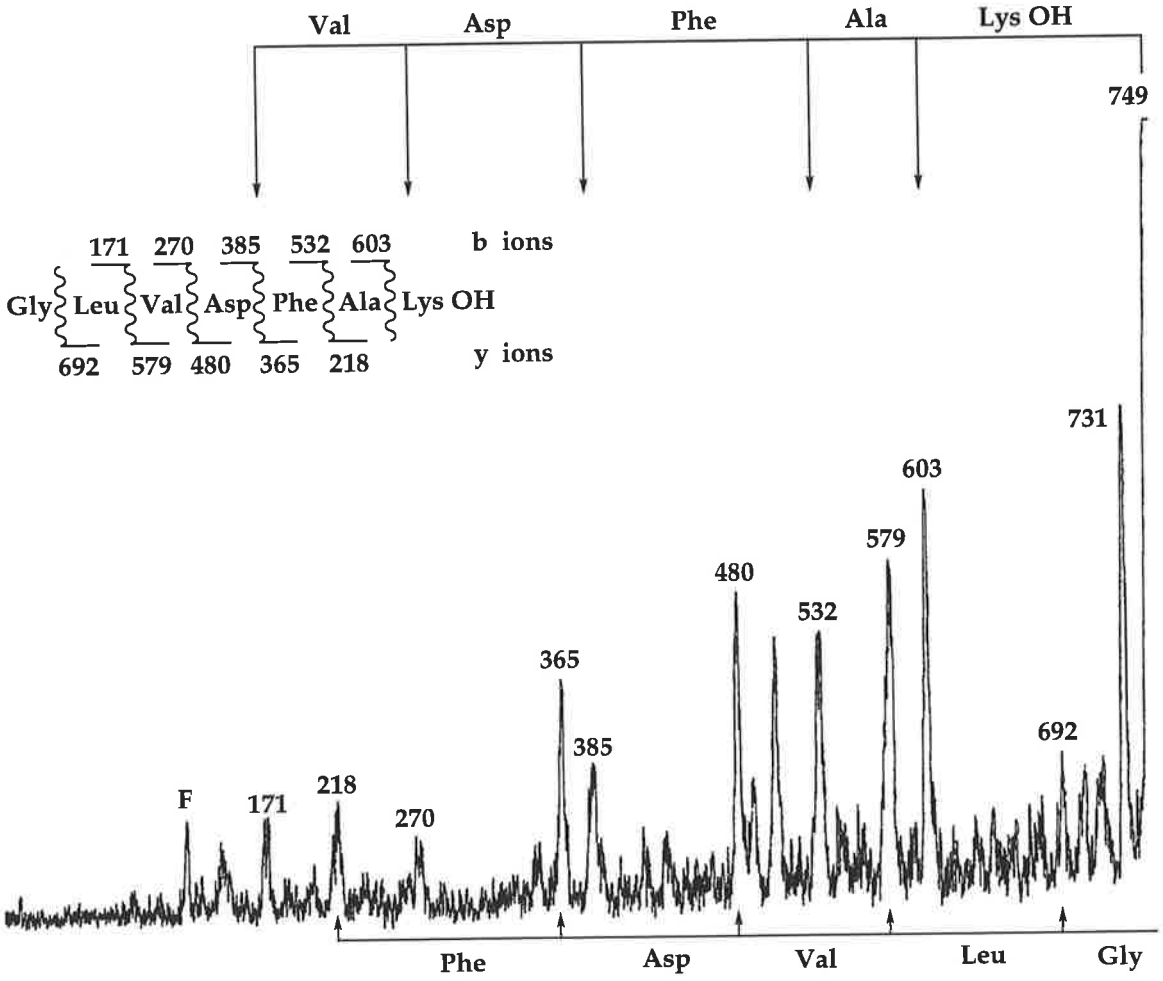


Figure 3.8: CA FAB MS/MS spectrum of peptide **2.5A** ($[M + H]^+ = 749$), formed from the Lys -C digest of uperin 2.5, with the 'b' ions being represented by schematic arrows at the top of the spectrum and 'y' ions at the bottom. The peak labelled **F** (m/z 120) corresponds to an internal fragment ion associated with phenylalanine.

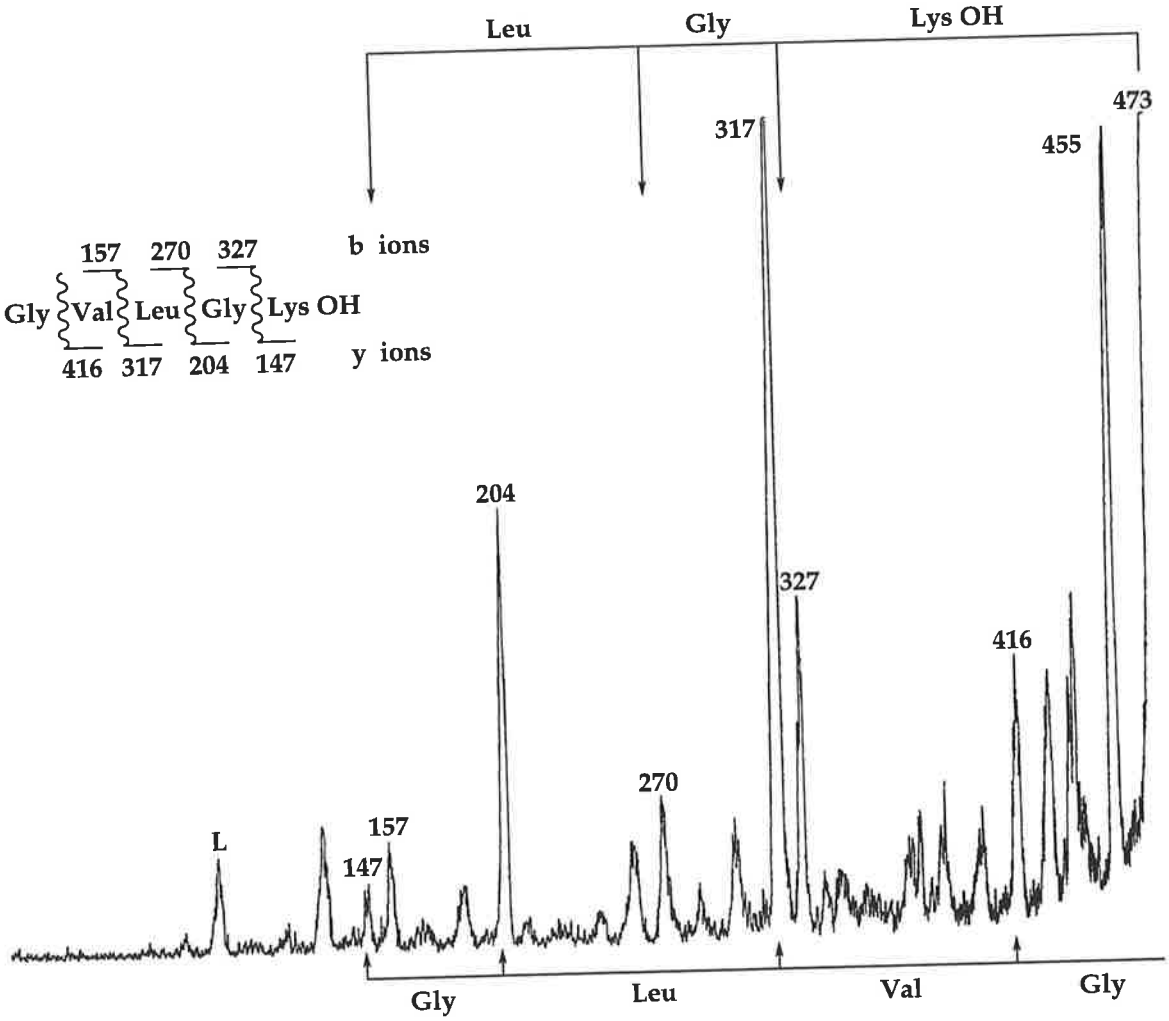


Figure **3.9**: CA FAB MS/MS spectrum of peptide **2.5B** ($[M + H]^+ = 473$), formed from the Lys -C digest of uperin 2.5. The schematic arrows at the top indicate sequence information gained from 'b' ions. The 'y' ions are represented at the bottom. The peak labelled **L** (m/z 86) corresponds to an internal fragment ion associated with leucine.

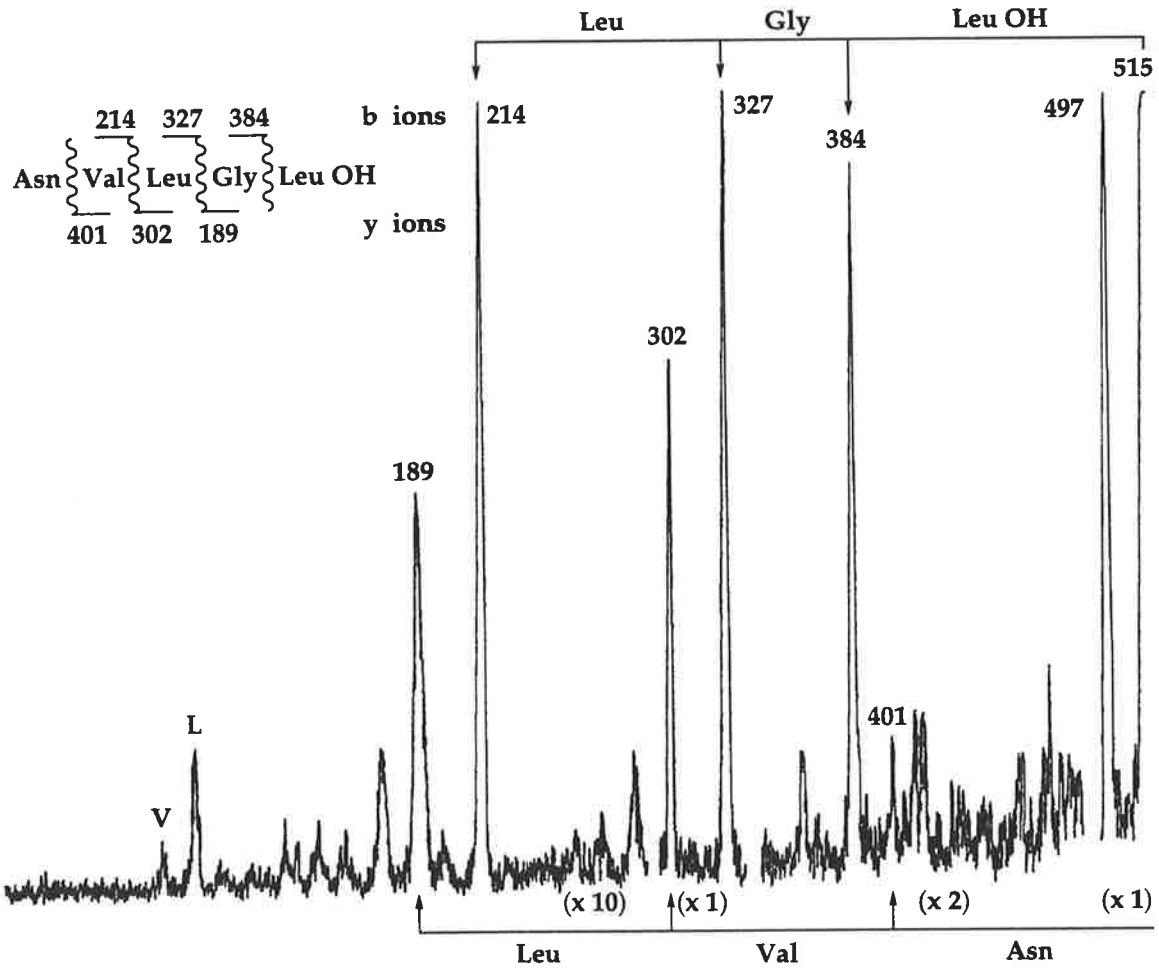


Figure **3.10**: CA FAB MS/MS spectrum of peptide **2.5C** ($[M + H]^+ = 515$), formed from the Lys-C digest of uperin 2.5. The schematic at the top of the spectrum indicate the 'b' ions which provide sequence information; similarly the sequence information provided by 'y' ions are denoted at the bottom. Valine and leucine show internal fragment ions, which are indicated at m/z 72 and 86 respectively by the letters **V** and **L**.

Uperin 3.1

In this case, only sufficient material was available to allow the following experiments: (i) automated Edman sequencing; (ii) one manual Edman; and (iii) two MS experiments.

(a) Uperin 3.1 has a retention time of 22.6 min and is a minor peptide. It coelutes with 4 trace peptides and was repurified by HPLC; it has a $[M + H]^+$ value of 1827.

(b) The manual Edman / FAB MS increased the mass by 213 Da, corresponding to the loss of a Gly residue and the attachment of two PITC units to two Lys residues (see footnote page 71).

(c) Due to the limited amount of material available, the automated sequencing was carried out before the mass spectrometric analyses. This gave the partial sequence as Gly Val Leu Asp Ala Phe Arg Lys Ile Ala Thr Val Val (14) Asn Val (17). Residues 14 and 17 were missed by the sequencer, which is a direct consequence of the limited amounts of material available.

(d) Enzymic digest with Lys-C cleaves uperin 3.1 at positions 8 and 14 to form three peptides:

3.1A	$[M + H]^+ = 630$	(residues 9 - 14);
3.1B	$[M + H]^+ = 905$	(residues 1 - 8); and
3.1C	$[M + H]^+ = 330$	(residues 15 - 17).

Two Lys residues in the structure is consistent with the observation from the Edman degradation.

(e) CA MS/MS (Table **3.9**) gave the sequence of:

(i) **3.1A** as Leu Ala Thr Val Val Lys (OH). A representative spectrum is shown in Figure **3.11**.

(ii) **3.1B** gave the partial sequence Gly (Val + Leu) Asp Ala Phe Arg Lys (OH).

(iii) **3.1C** were inconclusive.

Table **3.9:** CA MS/MS Data for Lys-C Products from Uperin 3.1.

Peptide	***m/z***	**Ion**	**Observed fragment ions *m/z* (%)**	**Peptide Sequence**
3.1A	630	'b'	185 (53); 286 (72); 385 (60); 484 (39); 612 (86)	Thr Val Val Lys
		'y'	517 (62); 446 (74); 345 (27); 246 (47)	Leu Ala Thr Val
Sequence of **3.1A** by CA MS/MS data:				**Leu Ala Thr Val Val Lys (OH)**
3.1B	905	'b'	759 (73); 887 (100)	Lys
		'y'	848 (38); 636 (13); 521 (18); 450 (14); 303 (12)	Gly (212) Asp Ala Phe
Sequence of **3.1B** by CA MS/MS data:				**Gly (212) Asp Ala Phe [Arg] Lys (OH)**

(f) It has been established by CA MS/MS and enzyme digests that residue 14 of uperin 2.5, which was skipped by the sequencer, is lysine. This leaves only residue 17 as being unidentified. The nominal mass of the parent ion ($[M + H]^+ = 1827$) infers that the C-terminal residue has a mass of 115 Da, *i.e.* Val (NH_2). Uperin 3.1 therefore corresponds to the structure given in Table **3.1**, and this has been confirmed by synthetic studies.

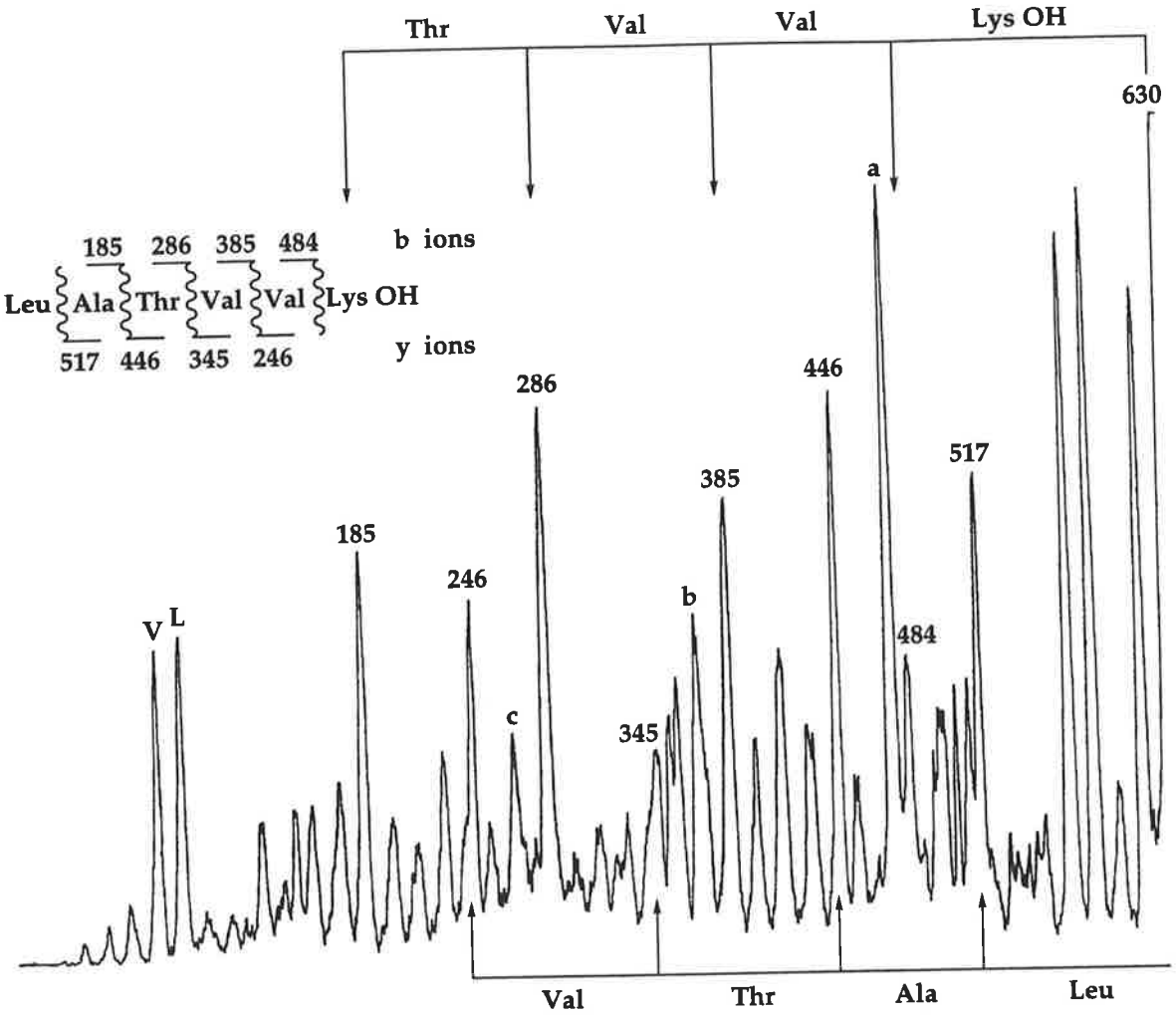


Figure **3.11**: CA FAB MS/MS spectrum of peptide **3.1A** ($[M + H]^+ = 630$), formed from the Lys-C digest of uperin 3.1. Sequence information is provided by 'b' and 'y' ions which are respectively shown as schematics at the top and bottom of the spectrum. Intense internal fragmentation ions for valine and leucine are shown at m/z 72 and 86 respectively. The base peak (**a**) in the spectrum corresponds to m/z 474 which is an 'x_4' fragmentation involving the loss of Leu Ala.[329] Peaks (**b**) and (**c**), at m/z 367 and 268, correspond to loss of H_2O from the respective 'b' ions at m/z 385 and 286: this is characteristic of threonine.

Uperins 3.2 and 3.3

(a) Uperin 3.2 ($[M + H]^+ = 1841$), is a minor peptide that has a retention time of 23.8 minutes and coelutes with a major peptide, uperin 3.3, ($[M + H]^+ = 1813$) (Figure **3.4**). Uperins 3.2 and 3.3 could not be separated by HPLC, therefore their structures were determined primarily by tandem MS/MS as a mixture. The notation of **3.2,3X** denotes peptides which are derived from this mixture.

(b) Enzymic digest with Lys-C produced only four peptides: **3.2,3A**, **3.2,3B**, **3.2,3C** and **3.2,3D**. The masses of these peptides and their esterified products are listed in Table **3.10**

Table **3.10**: Results of the mixed Lys-C digest and the subsequent esterification experiments for uperins 3.2 and 3.3.

Peptide	$[M + H]^+$	esterification	Δ	Acids	Amides
3.2,3A	630	644	14	1	0
3.2,3B	905	933	28	2	0
3.2,3C	749	777	28	2	0
3.2,3D	344	374	30	0	2

(c) CA MS/MS sequence studies (Table **3.11**) show that:

(i) the spectra of peptides **3.2,3A** and **3.2,3B** were identical to the spectra of peptides **3.1A** and **3.1B** respectively;

(ii) **3.2,3C** gave the structure as Gly Val Leu Asp Ala Phe Lys (OH);

(iii) the $[M + H]^+$ species from **3.2,3D** has a C-terminal Val (NH_2).The identity of the remainder of peptide **3.2,3D** could not be determined from this spectrum.

Table **3.11**: CA MS/MS Data for Lys-C Products from a Mixture of Uperins 3.2 and 3.3.

Peptide	*m/z*	Ion	Observed fragment ions *m/z* (%)	Peptide Sequence
3.2,3C	749	'b'	456 (12); 603 (48); 731 (100)	Phe Lys
		'y'	692 (61); 593 (46); 480 (16); 365 (8); 294 (7)	Gly Val Leu Asp Ala
Sequence of **3.2,3C** by CA MS/MS data:				**Gly Val Leu Asp Ala Phe Lys (OH)**
3.2,3D	344	'b'	228 (76); 327 (100)	Val
		'y'	none detected	
Sequence of **3.2,3D** by CA MS/MS data:				**(228) Val (NH_2)**

(d) The data obtained are summarised as follows:

3.2,3A:	Leu Ala Thr Val Val Lys (OH)	$[M + H]^+ = 630$
3.2,3B:	Gly Val Leu Asp Ala Phe Arg Lys (OH)	$[M + H]^+ = 905$
3.2,3C:	Gly Val Leu Asp Ala Phe Lys (OH)	$[M + H]^+ = 749$
3.2,3D:	[(Asn + Leu) or (Gln + Val)] Val (NH_2)	$[M + H]^+ = 344$

The sequences of uperins 3.2 and 3.3 must be composed of a combination of these four peptides. Peptide **3.2,3D** must be the last three residues of both peptides due to the presence of the C-terminal amide, *i.e.* Val (NH_2). The nominal masses of the parent peptides, implies that uperin 3.2 must therefore be composed of peptides **3.2,3A**, **3.2,3B** and **3.2,3D**; whereas uperin 3.3 must be composed of peptides **3.2,3A**, **3.2,3C** and **3.2,3D**. Summation of the nominal masses of the residues which comprise uperin 3.3 fall 128 Da short of the mass of the parent ion: there must be an additional Lys (unobserved in any of the mass spectra) which resides at either position 1 or 8.

(e) Automated sequencing of uperins 3.2 and 3.3 (as a mixture) showed that they differed in structure only at residue 7, which the sequencer interpreted as being Lys and Arg. The structures of uperin 3.2 and 3.3 are shown in Table **3.1**.

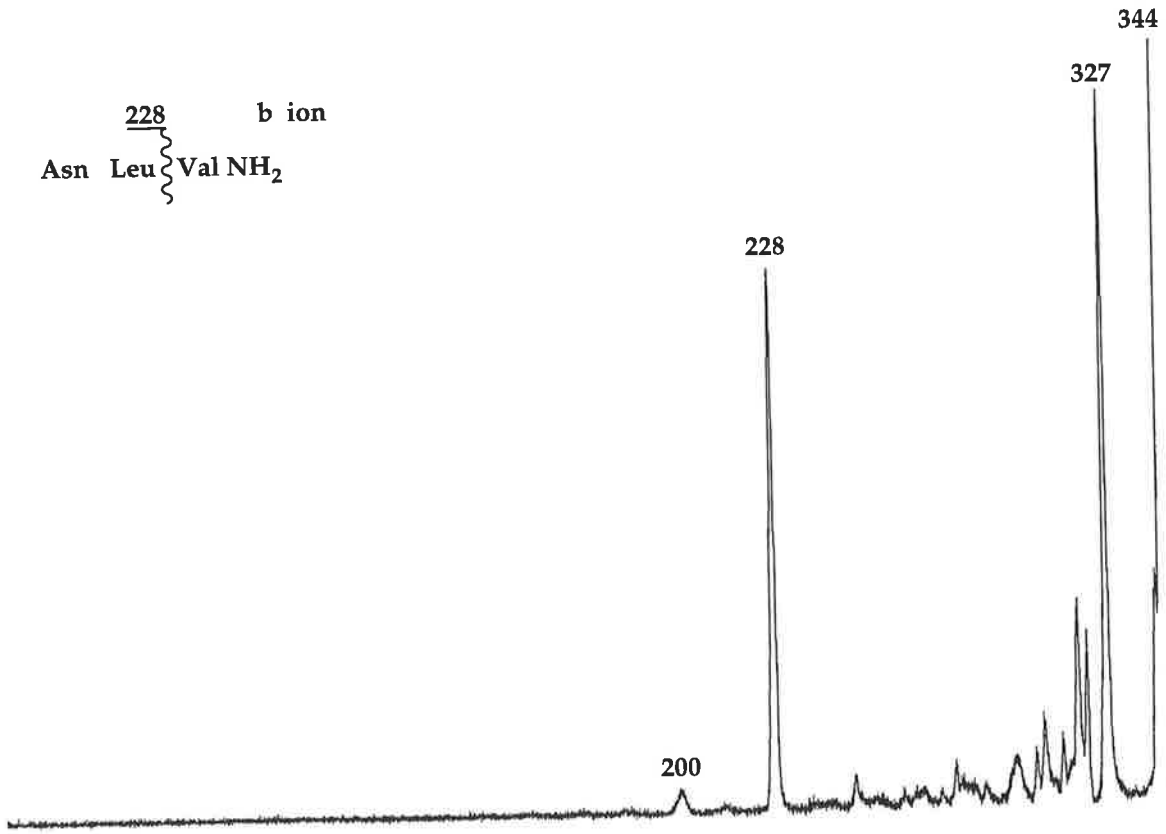


Figure **3.12**: CA FAB MS/MS spectra of peptide **3.2,3D** ($[M + H]^+ = 344$), formed from the Lys-C digest of an inseparable mixture of uperins 3.2 and 3.3. Loss of NH_3 accounts for the base peak at m/z 327. The only sequence information available derives from the 'b_2' and 'a_2' cleavage ions at m/z 228 and 200.

Uperin 4.1

(a) FAB MS of Uperin 4.1 gives $[M + H]^+ = 1724$. It has a retention time of 22.0 minutes and elutes with about seven trace peptides and was repurified by HPLC.

(b) Manual Edman / FAB MS indicated that the first two residues from the N-terminal end of the peptide are Gly Val, and that two Lys residues are present as an increase of 213 Da was observed for the first degradation.

(c) The conventional FAB MS spectrum (m/z range 1800 to 600) of uperin 4.1 shows significant fragmentation (cf. Uperin 1.1). The spectrum is presented as Figure **3.13** and gives the partial sequence:

Gly (156) Ser (260) His Lys Val Val Ser Ala Leu Lys Asn Val Ala (NH_2).

(d) Enzymic digest with Lys-C gave 3 peptides: **4.1A**, **4.1B** and **4.1C**. The masses of these peptides and their esterified products are listed in Table **3.12**.

Table **3.12**: Results of the Lys-C Digest and Subsequent Esterification Experiments for Uperin 4.1.

Peptide	$[M + H]^+$	esterification	Δ	Acids	Amides
4.1A	844	858	14	1	0
4.1B	616	630	14	1	0
4.1C	302	332	30	0	2

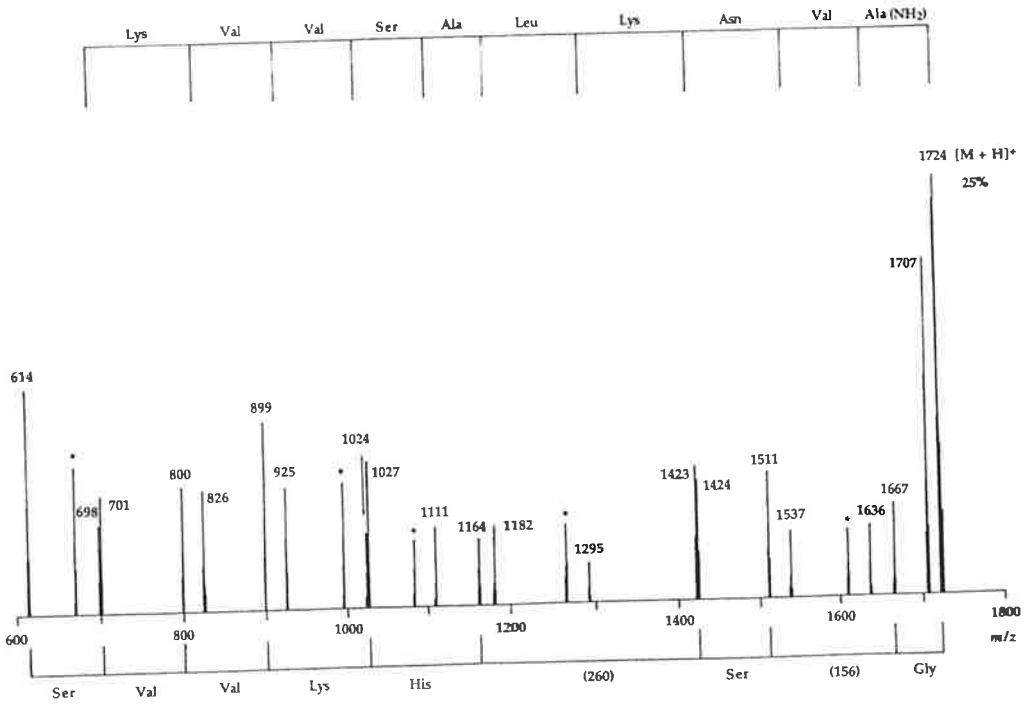


Figure **3.13**: Conventional FAB mass spectrum of uperin 4.1. Relative abundance of peak / magnetic sector plot with masses of peaks indicated in Daltons. See Experimental (section **6.1**) for full details. Sequence information is provided by 'b' and 'y' cleavage ions which are indicated by the schematic arrows at the top and bottom of the spectrum respectively. Peaks marked with an asterisk (*) are the observed 'a' ions, which are 28 Daltons lower than the corresponding 'b' ions.

(e) CA MS/MS data (Table **3.13**) gives the structures for:

(i) **4.1A** as Gly Val Gly Ser Phe Leu His Lys (OH), which are the first eight residues of uperin 4.1;

(ii) **4.1B** as Val Val Ser Ala Leu Lys (OH). This spectrum is shown as Figure **3.14**.

(iii) **4.1C** is inconclusive.

Table **3.13**: CA MS/MS Data for Lys-C Products from Uperin 4.1.

Peptide	***m/z***	**Ion**	**Observed fragment ions *m/z* (%)**	**Peptide Sequence**
4.1A	844	'b'	301 (2); 448 (4); 561 (6); 698 (41); 826 (100)	Phe Leu His Lys
		'y'	787 (18); 688 (15); 631 (4); 544 (5); 397 (6); 284 (9)	Gly Val Gly Ser Phe Leu
Sequence of **4.1A** by CA MS/MS data:				**Gly Val Gly Ser Phe Leu His Lys (OH)**
4.1B	616	'b'	199 (18); 286 (14); 357 (50); 470 (65); 598 (100)	Ser Ala Leu Lys
		'y'	517 (22); 418 (62); 331 (16); 260 (16)	Val Val Ser Ala
Sequence of **4.1B** by CA MS/MS data:				**Val Val Ser Ala Leu Lys (OH)**

(f) From the available information, the proposed sequence of uperin 4.1 is:

Gly Val Gly Ser Phe Leu His Lys Val Val Ser Ala Leu Lys Asn Val Ala (NH_2)

Automated sequencing confirms this structure and identifies Ile (6, 13).

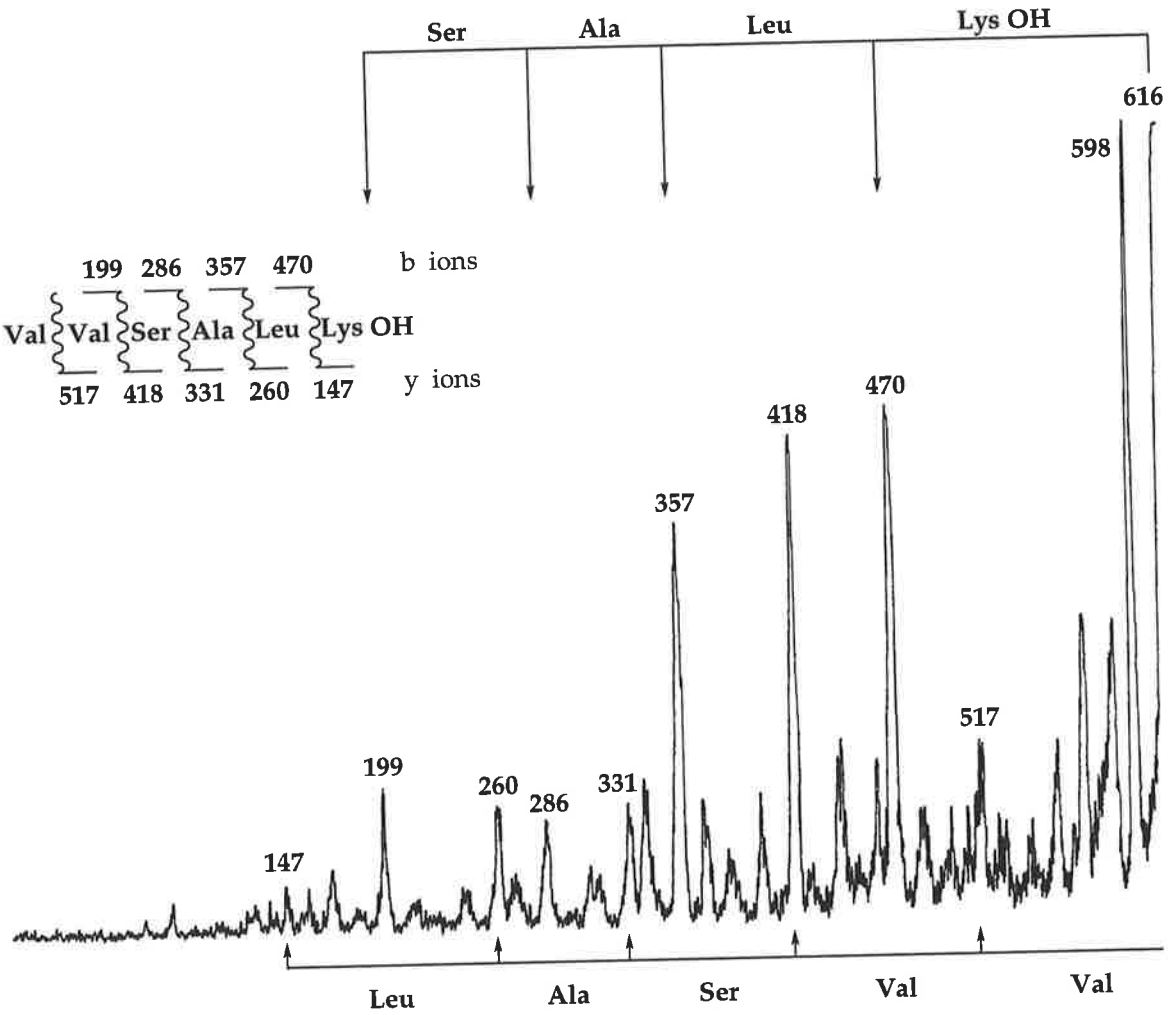


Figure **3.14**: CA FAB MS/MS spectrum of uperin 4.1B ($[M + H]^+ = 616$), formed from the Lys-C digest of uperin 4.1. The top of the spectrum lists the 'b' ions; the bottom shows the 'y' ions.

Uperin 5.1

Uperin 5.1 has a retention time of 18.8 min, coelutes with 8 trace peptides and has a parent ion of $[M + H]^+ = 1456$. In this case there was only sufficient material for an $[M + H]^+$ measurement and an automated Edman determination. The structure is listed in Table **3.1**.

Uperins 6.1 and 6.2

FAB MS showed that uperins 6.1 and 6.2 had molecular weights in excess of 3200 Da: MALDI TOF mass spectrometry yield parent ions of masses 3233.85 and 3261.85 Da respectively. This together with automated Edman sequencing gives respective structures as shown in Table **3.1**.

3.5 Primary Structural Features of the Uperin Peptides from *U. inundata*

Based on the primary sequences, the thirteen novel peptides have been classified into six sub-groups. The basis of this classification involved grouping peptides that have the same number of residues in the sequence as well as the same identity at the C-terminal position, *i.e.* primary amide or free acid. Uperin 4.1 was not classed as a uperin 3 type peptide for the following reasons. It does not have an aspartic acid residue at position 4 which is characteristic of all uperin 3 peptides; and it shares only 30% homology for identical residues in the same position as other uperin 3 type peptides: uperins 3.1 - 3.3 share 90% homology. Also, uperin 2.5 is the most dissimilar uperin 2 peptide (55% homology, cf. uperins 2.1 - 2.4, 80% homology). Uperins 6.1 and 6.2 which differ only at residue 9, have greater than 95% homology.

3.6 Amphiphilic Secondary Structural Features of the Uperin Peptides

Although the primary sequences of the uperins have been determined, the secondary structures must also be considered so as to give plausible explanations for the biological activity that each of the peptides may possess. The potential regions of amphiphilic secondary structure of uperins 2.3 and 3.3 are shown as both the Edmundson wheel and helical net diagrams in Figures **3.15** and **3.16** respectively. Uperins 2.1 - 2.5, 3.1 - 3.3 and 4.1 show well defined hydrophobic and hydrophilic zones. Such a property suggests the possibility of antimicrobial activity, as the peptide may form an α-helix and bind to the surface of biological membranes, such as lipids in the outer membrane of the bacterial cell. The helical peptide ultimately penetrates the membrane, forms ion channels and causes an influx (or efflux) of ions, resulting in cell death.(330)

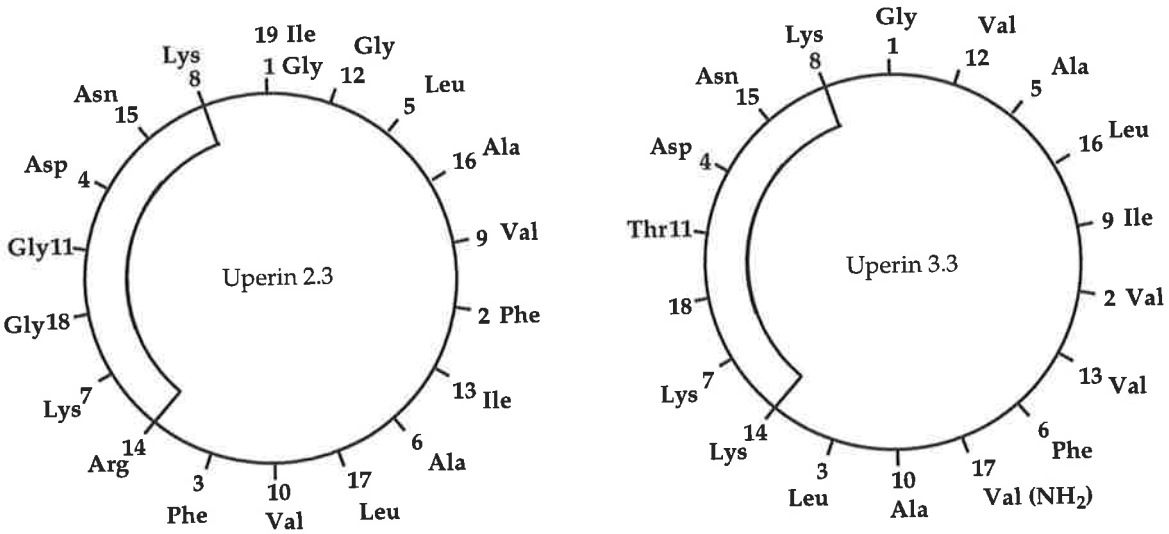


Figure **3.15**: Edmundson wheel projections for uperins 2.3 and 3.3.

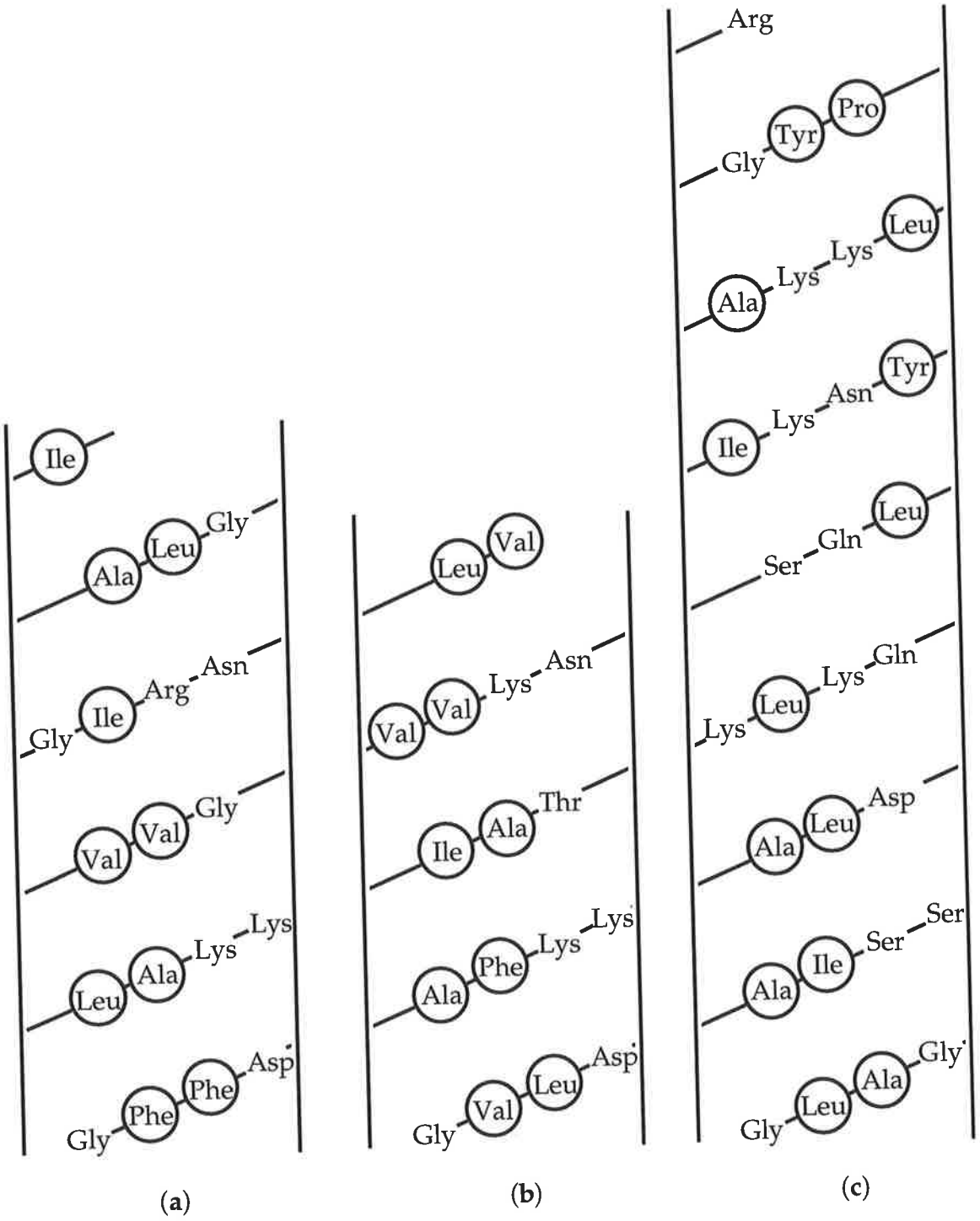


Figure **3.16**: Helical net diagrams for uperins 2.3 (**a**); 3.3 (**b**); and 6.1 (**c**). Hydrophobic residues are circled.

The Edmundson wheel projections for 5.1, 6.1 and 6.2 are unusual in that they show no discrete zones. The helical net diagrams for uperins 6.1 [Figure **3.16(c)**] and 6.2 are more informative as they suggest that there are distinct hydrophilic and hydrophobic zones. The hydrophilic zone is clearly visible as a continuous span of "uncircled residues" from Gly (4) to Gly (27). The hydrophobic zone spans the "circled residues" from Leu (2) to Pro (29), but is not clearly seen since the residues Leu (18) to Pro (29) are split by the net edges. The helical net and Edmundson wheel diagrams only suggest the secondary structure of a peptide. Detailed studies of the secondary and tertiary structures of a peptide can be obtained by using very powerful NMR instruments (*e.g.* 600 MHz ^{1}H NMR).*

Since amphibian peptides have been proposed to have mammalian analogues,[118] a data bank search† for uperins 2 - 5 was investigated. This has shown no statistically significant correlation with any peptides or proteins from other organisms.

---

* The tertiary structure of a peptide in solution may be determined, in part, by a combination of techniques such as circular dicroism, Fourier transform infrared and Raman spectroscopies and nuclear magnetic resonance.[237] NMR is arguably the most important of these techniques because of the abundance of information it provides. Many of the NMR techniques developed for the determination of large protein structures are also applicable to determination of smaller peptide structures.[238] The nuclear Overhauser effect (NOE) is the most useful parameter measured from either one dimensional NOE difference spectra or from two dimensional NOESY (or ROESY) spectra. The NOE arises from the cross relaxation between two nuclei which are close in space, allowing the determination of a number of geometric constraints for each residue of the peptide. Coupling constants, particularly the vicinal coupling from the amide proton to the α-proton may also be used to provide information about the tertiary structure.

† Australian National Genomic Information Service[331] using GENINFO (R) BLAST Network Service (Cruncher).

3.7 Biological Activity of the Uperins

Due to the small quantities of natural peptides isolated in comparison to the amounts required for biological testing, synthetic analogues composed of all L-amino acids were submitted for biological testing.* The testing program for the uperins has so far only involved screening for antibiotic and antimicrobial activity. The activity results of a number of synthetic uperins are listed in Table **3.14**. What is evident from these results is that there is no uperin which shows pronounced wide-spectrum antibiotic character (cf. caerin 1.1 from the *Litoria* genus).(193) Instead, the uperins show at least weak antibiotic activities against all of the listed micro-organisms. The minimum inhibitory concentration values (MIC) are in the range 100 - 250 μg/mL. There are some uperins which show medium activity (1 - 50 μg/mL) against individual pathogens: these are indicated by an asterisk in Table **3.14**. An explanation for these observations is that *Uperoleia inundata* may use uperins 2.1 - 2.5, 3.1 - 3.3 and 4.1 collectively to constitute an antibiotic 'cocktail'. Perhaps this is one reason why there is such a variety of structurally related peptides in the skin of *Uperoleia inundata*.†

---

* For financial reasons, not all uperins could be synthesised.

† This proposal may account for a similar scenario for the recently reported gaerurin peptides from *Rana rugosa*.(332)

Table 3.14: General Antibiotic Activity of the Uperin Peptides.

Organism	2.1	2.2	2.3	2.4	2.5	3.1	4.1	6.1
Bacillus cereus			*					
Escherichia coli								*
Leuconostoc lactis	*		*		*	*	*	
Leuconostoc mesenteriodes								
Listeria innocua	*				*			
Micrococcus luteus						*		
Pasteurella haemolytica								
Staphylococcus aureus	*		*		*			
Streptococcus uberis								

* indicates activity < 50 μg/ml

What is surprising about these results is that uperins 2.2 and 2.3 (which differ only at residue 3) have vastly different activities. This behaviour is not understood as it would have been reasonable to expect these two peptides to show relatively similar activity. Further investigations into these two peptides need to be conducted to rationalise these observations. Not listed in Table **3.14** is uperin 5.1 which showed no antibiotic activity in the testing program. Uperins 6.1 and 6.2 have a higher proportion of hydrophilic residues than any reported amphibian peptide (*i.e.* 16 out of a total of 30 residues). The lack of intense antibiotic activity for these peptides does not necessarily imply that they are poorly active: perhaps these peptides have a certain and specific function. At this stage the role of these particular peptides in the amphibian integument is not known.

3.8 Uperin 1.1 - A New Tachykinin

Uperin 1.1 is a member of the tachykinin family of peptides since it contains the distinct C-terminal sequence of Phe X Gly Leu Met (NH_2), where X can be any amino acid. It differs from uperolein and physalaemin by having alanine at positions 2 and 6 respectively. The activity of uperin 1.1 has been studied in the Rome laboratories(333) and experimental details are recorded in section **6.10**. In smooth muscle testing, it is only marginally less active than uperolein: *e.g.*, in guinea pig ileum smooth muscle testing, it shows formidable activity (at 0.4 nanograms per mL). It also elicits a response (reduction) in rabbit blood pressure at a concentration as low as 5 nanograms per kilogram (of body weight). Uperin 1.1 is thus one of the more active of the tachykinin type neuropeptides.(cf: 159) The results are shown in Table **3.15**. The activity of physalaemin was considered equal to 100, that of the other two tachykinins was expressed as a percentage of this activity and the number of experiments are given in parentheses. In the case of the guinea pig ileum tests, the percentages correlate to the following minimum concentration of peptides: physalaemin (0.2 ng/mL); uperolein (0.25 ng/mL); and uperin 1.1 (0.4 ng/mL). The minimum concentration of the standard, physalaemin, required to initiate a reduction in rabbit blood pressure = 2 ng/kg.(159)

Table 3.15: Bioassay of Uperin 1.1.*

Test Preparation	Physalaemin	Uperolein	Uperin 1.1
Guinea-pig ileum	100	80 ± 15 [8]	57 ± 13 [3]
Rabbit terminal colon	100	80 ± 20 [10]	64 ± 14 [3]
Rabbit blood pressure	100	46 ± 15 [7]	37 ± 3 [2]

* Figures are with respect to physalaemin (100%)

Chapter 4 *UPEROLEIA MJOBERGI*

Figure **4.1**: Photograph of ***Uperoleia mjobergi***, courtesy of Dr. M. Davies, Department of Zoology, The University of Adelaide, 5005.

4.1 *Uperoleia mjobergi*

Mjöberg's toadlet, *Uperoleia mjobergi*, was first described by Andersson in 1913 as *Psuedophryne mjobergi*,(334) but was transferred to the genus *Glauertia* in 1940,(312) before being classified in the current genus in 1971.(318) It is a small

frog attaining a length of between 1.9 - 2.5 cm. The skin secretion is contained in granular glands which can be seen distinctly as warty protuberances on the dorsal surface of the animal. A picture of *U. mjobergi* is shown in Figure **4.1**. It is confined to northern Western Australia in the area around Broome, Derby and Fitzroy River as illustrated in Figure **4.2** and is very similar to *U. inundata*, which is found in the 'Top End' of Australia. The nine specimens studied for this project were collected in the vicinity of Derby and maintained in captivity.

Figure **4.2**: Distribution of *Uperoleia mjobergi*.

Currently, seven peptides have been isolated and characterised from the dorsal glands of *U. mjobergi*. Since the seven peptides are similar in structure to either the uperin 2 or 3 type peptides (from *U. inundata*), the same nomenclature and numbering systems are continued in this study. The seven peptides are named uperins 2.6 - 2.8 and 3.4 - 3.7. The amino acid sequences of

the seven peptides are listed in Table **4.1**. Although *U. inundata* and *U. mjobergi* have structurally related peptides, there are no peptides common to both species. This chapter details how the primary structures of the uperins isolated from *U. mjobergi* were determined: the methods used are the same as for *U. inundata*, except that enzyme digests were widely used (*i.e.* Lys-C, Arg-C,* Asp-N* and chymotrypsin*), and manual Edman degradations were deemed unnecessary.

As in chapter 3, the benign method of surface electrical stimulation was used to effect release of the glandular secretion (see section **2.4**). The 'milking' of the nine specimens produced, after work-up, a total of 67.5 mg of lyophilised peptide material (an average of 7.5 mg per frog).

A typical analytical HPLC chromatogram showing the separation of the lyophilised material from *U. mjobergi* is shown in Figure **4.3**. A combination of data from FAB MS (*m/z* 3500 to 500 Da mass range) of the individual HPLC peaks and the HPLC data determined that: (i) the peptides eluted in the 22.0 - 26.0 minute range; (ii) that uperins 3.4, 3.5 and 3.7 are present in high concentrations (*ca.* 1 mg each per frog); and (iii) that uperins 2.6 - 2.8 and 3.6 are present in low concentrations (*ca.* 0.2 - 0.3 mg per frog). The identity of the components eluting before 22.0 minutes has not yet been determined as they do not show masses > 1000 Da by FAB MS. The peptides have molecular weights between 1700 - 2000 Da.

---

* The uperins from *U. inundata* had been characterised by the time these enzymes arrived and thus the enzymes were only used in the studies involving the uperins from *U. mjobergi*.

Table 4.1: Amino Acid Sequence of the Uperins from *U. mjobergi*.

Uperin	$[M + H]^+$ [a]	Amino Acid Sequence
2.6	1948	Gly Ile Leu Asp Ile Ala Lys Lys Leu Val Gly Gly Ile Arg Asn Val Leu Gly Ile (OH)
2.7	1948	Gly Ile Ile Asp Ile Ala Lys Lys Leu Val Gly Gly Ile Arg Asn Val Leu Gly Ile (OH)
2.8	1978	Gly Ile Leu Asp Val Ala Lys Thr Leu Val Gly Lys Leu Arg Asn Val Leu Gly Ile (OH)
3.4	1735	Gly Val Gly Asp Leu Ile Arg Lys Ala Val Ala Ala Ile Lys Asn Ile Val (NH_2)
3.5	1779	Gly Val Gly Asp Leu Ile Arg Lys Ala Val Ser Val Ile Lys Asn Ile Val (NH_2)
3.6	1826	Gly Val Ile Asp Ala Ala Lys Lys Val Val Asn Val Leu Lys Asn Leu Phe (NH_2)
3.7	1843	Gly Val Gly Asp Ile Phe Arg Lys Ile Val Ser Thr Ile Lys Asn Val Val (NH_2)

[a] Nominal mass

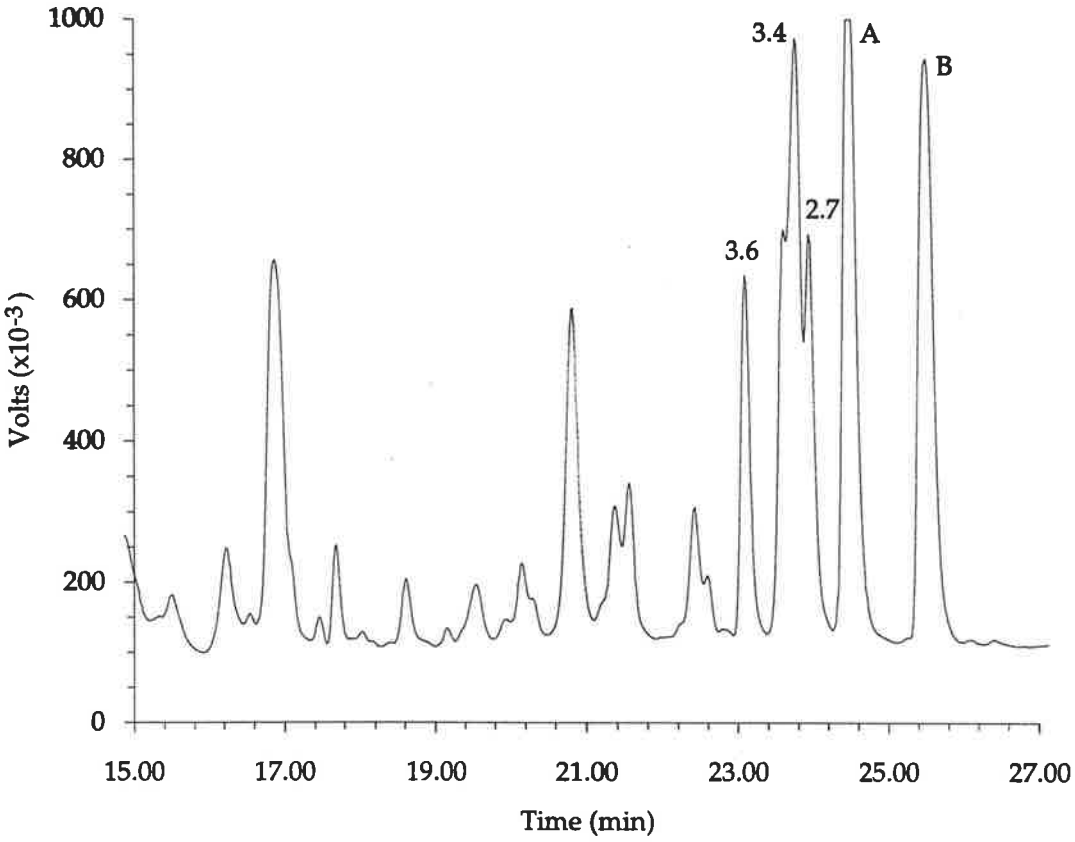


Figure **4.3**: Analytical HPLC trace of peptides from the glands of *U. mjobergi*. For full experimental details see section **6.3**. The labelled peaks are unresolved and correspond to the coelution of (**A**) uperins 2.6 and 3.7, and (**B**) uperins 2.8 and 3.5.

4.2 Structure Determination of Peptides from *Uperoleia mjobergi*

The structure determination used in this chapter is essentially the same as that described in chapter three: mass spectrometric techniques were primarily used. Firstly, positive ion FAB MS determined the molecular weight of the peptide. Enzymic digest using Lys-C, Arg-C, Asp-N and chymotrypsin (as appropriate) cleaved the peptide into at least two smaller peptides. The masses of the digest fragments were analysed by FAB MS. Conversion of the parent peptide to the methyl esters determined the quantity of acids and primary amides within the sequence: more importantly it indicated if the C-terminus was post translationally modified (*i.e.* $CO_2H \rightarrow CONH_2$). An increment of 14 Da indicated the presence of an acid (*i.e.* $CO_2H \rightarrow CO_2Me$); an increment of 15 Da indicated the presence of a primary amide (*i.e.* $CONH_2 \rightarrow CO_2Me$). The resultant peptides were sequenced by CA MS/MS which provided the full amino acid sequence of the majority of peptides. Finally, automated sequencing was used to distinguish between the isomers Leu and Ile, and confirm the sequence of each peptide.

For clarity, the data for the structure determination of the uperins are presented in the following concise format where the mass spectrometric data and the results of enzyme digestion can be clearly seen. There is a lot of sequencing information generated in these two processes, and only the essential data are presented to avoid the tables becoming cluttered. The molecular mass, essential HPLC data and the esterification results are also listed. All masses listed in the tables are expressed in Daltons.

Uperin 2.6

(a) FAB MS: $[M + H]^+ = 1948$.

(b) HPLC retention time: 24.4 minutes (Figure **4.3**). Uperin 2.6 coelutes with uperin 3.7 and two unidentified peptides. Repurification by HPLC separated uperins 2.6 and 3.7.

(c) Esterification of uperin 2.6 with methanol increased the mass by 43 Da, which corresponds to the presence of two acid groups and one amide.

(d) The results of enzymic digests and CA MS/MS experiments are listed in Tables **4.2** and **4.3**, respectively.

Table **4.2**: Results of the Enzyme Digests of Uperin 2.6.

Peptide Digested	Enzyme Used	Resultant Peptide	$[M + H]^+$
Uperin 2.6	Lys-C	**2.6A**	729[a]
		2.6B	857
		2.6C	1110
Uperin 2.6	Asp-N	**2.6D**	302[a]
		2.6E	1665
Uperin 2.6	Arg-C	**2.6F**	515[a]
		2.6G	1452[b]
2.6G	Lys-C	**2.6A**	729[a]
		2.6B	857
		2.6H	614[a]
		2.6I	742[b]

[a] CA MS/MS studies performed on this ion.

[b] CA MS/MS studies performed on this ion were inconclusive.

Table **4.3**: CA MS/MS Data for the Enzyme Products from Uperin 2.6.

Peptide	***m/z***	**Ion**	**Observed fragment ions [*m/z* (%)]**	**Peptide sequence**
2.6A	729	'b'	171 (4); 284 (12); 399 (21); 512 (41); 583 (48); 711 (100)	Leu Asp Leu Ala Lys
		'y'	672 (46); 559 (36); 446 (15); 331 (15); 218 (13); 147 (7)	Leu Asp Leu Ala
Sequence of **2.6A** by CA MS/MS data:				**Gly Leu Leu Asp Leu Ala Lys (OH)**
2.6D	302	'b'	171 (100); 284 (41)	Leu
		'y'	245 (24); 132 (9)	Gly Leu
Sequence of **2.6D** by CA MS/MS data:				**Gly Leu Leu (OH)**
2.6F	515	'b'	214 (9); 327 (98); 384 (97); 497 (100)	Leu Gly Leu
		'y'	401 (9); 302 (42); 189 (4)	Asn Val Leu
Sequence of **2.6F** by CA MS/MS data:				**Asn Val Leu Gly Leu (OH)**
2.6H	614	'b'	213 (9); 270 (21); 327 (72); 440 (15); 596 (100)	Gly Gly Leu Arg
		'y'	501 (86); 402 (37); 345 (13); 175 (19)	Leu Val Gly Gly Leu
Sequence of **2.6H** by CA MS/MS data:				**Leu Val Gly Gly Leu Arg (OH)**

From these experiments it can be seen that **2.6D** must correspond to the first three residues from the N-terminal end of uperin 2.6. It is a product of the Asp-N digest, and does not contain the Asp residue. Consequently, **2.6A** must be the first seven residues of uperin 2.6. Similarly, **2.6F** must be the last five residues of uperin 2.6 since it is the product of the Arg-C digest, and **2.6H** must be the central portion of uperin 2.6. Both Lys-C digests give products at *m/z* 729 (**2.6A**) and 857 (**2.6B**); **2.6B** corresponds to **2.6A** + Lys, resulting from incomplete digestion of uperin 2.6.

(f) The sequence from MS data gives Gly Leu Leu Asp Leu Ala Lys Lys Leu Val Gly Gly Leu Arg Asn Val Leu Gly Leu (OH). Automated sequencing confirms the sequence and identifies Ile (2, 5, 13 and 19).

Uperin 2.7

(a) FAB MS: $[M + H]^+ = 1948$.

(b) HPLC retention time: 23.6 minutes (Figure **4.3**). Uperin 2.7 coelutes with uperin 3.4. Repurification by HPLC separated uperins 2.7 and 3.4.

(c) Uperin 2.7 is isomeric with uperin 2.6. Enzymic and MS results are identical for uperins 2.6 and 2.7. Automated sequencing indicated that they differ at only one residue, *i.e.* Leu (3) for uperin 2.6, and Ile (3) for uperin 2.7.

Uperin 2.8

(a) FAB MS: $[M + H]^+ = 1978$.

(b) HPLC retention time: 25.8 minutes (Figure **4.3**). Uperin 2.8 coelutes with uperin 3.5, which were separated by HPLC upon repurification.

(c) Esterification of uperin 2.8 with methanol increased the mass by 43 Da, indicating the presence of two CO_2H groups and one $CONH_2$.

(d) The results of enzymic digests are listed in Table **4.4**, and the results of CA MS/MS experiments are recorded in Table **4.5**.

Table **4.4**: Results of the Enzyme Digests of Uperin 2.8.

Peptide Digested	Enzyme Used	Resultant Peptide	$[M + H]^+$
Uperin 2.8	Lys-C	**2.8A**	517[a]
		2.8B	715[a]
		2.8C	784[a]
Uperin 2.8	Asp-N	**2.8D**	302[a]
		2.8E	1695
Uperin 2.8	Arg-C	**2.8F**	515[a,b]
		2.8G	1482

[a] CA MS/MS studies performed on this ion.

[b] Figure **4.4** is a CA MS/MS spectrum of this ion.

Table **4.5**: CA MS/MS Data for the Enzyme Digests from Uperin 2.8.

Peptide	*m/z*	Ion	Observed fragment ions [*m/z* (%)]	Peptide sequence
2.8A	517	'b'	215 (4); 314 (8); 371 (13); 499 (100)	Val Gly Lys
		'y'	416 (17); 303 (27); 204 (7)	Thr Leu Val
Sequence of **2.8A** by CA MS/MS data:				**Thr Leu Val Gly Lys (OH)**
2.8B	715	'b'	284 (8); 399 (17); 498 (29); 569 (36); 697 (100)	Val Ala Lys
		'y'	658 (38); 545 (28); 432 (16); 317 (10)	Gly Leu Leu Asp
Sequence of **2.8B** by CA MS/MS data:				**Gly Leu Leu Asp Val Ala Lys (OH)**
2.8C	784	'b'	384 (2); 483 (3); 596 (6); 653 (15); 766 (100)	Val Leu Gly Leu
		'y'	671 (55)	Leu
Sequence of **2.8C** by CA MS/MS data:				**Leu (Arg + Asn) Val Leu Gly Leu (OH)**
2.8D	302	'b'	171 (53); 284 (95)	Leu
		'y'	245 (58); 132 (3)	Gly Leu
Sequence of **2.8D** by CA MS/MS data:				**Gly Leu Leu (OH)**
2.8F	515	'b'	214 (9); 327 (96); 384 (46); 497 (100)	Leu Gly Leu
		'y'	401 (7); 302 (41); 189 (4)	Asn Val Leu
Sequence of **2.8F** by CA MS/MS data:				**Asn Val Leu Gly Leu (OH)**

From these experiments it can be deduced that **2.8D** contains the first three residues of uperin 2.8: it is a product of the Asp-N digest, and does not contain the Asp residue. Consequently, **2.8B** corresponds to the first seven residues of uperin 2.8. Similarly, **2.8F** must be the last five residues of uperin 2.8 since it is the product of the Arg-C digest. It follows that **2.8C** has the sequence Leu Arg Asn Val Leu Gly Leu (OH). Peptide **2.8A** contains residues 8 - 12 of uperin 2.8.

(f) The sequence from MS data gives Gly Leu Leu Asp Val Ala Lys Thr Leu Val Gly Lys Leu Arg Asn Val Leu Gly Leu (OH). Automated sequencing identifies Ile (2 and 19) and confirms the sequence.

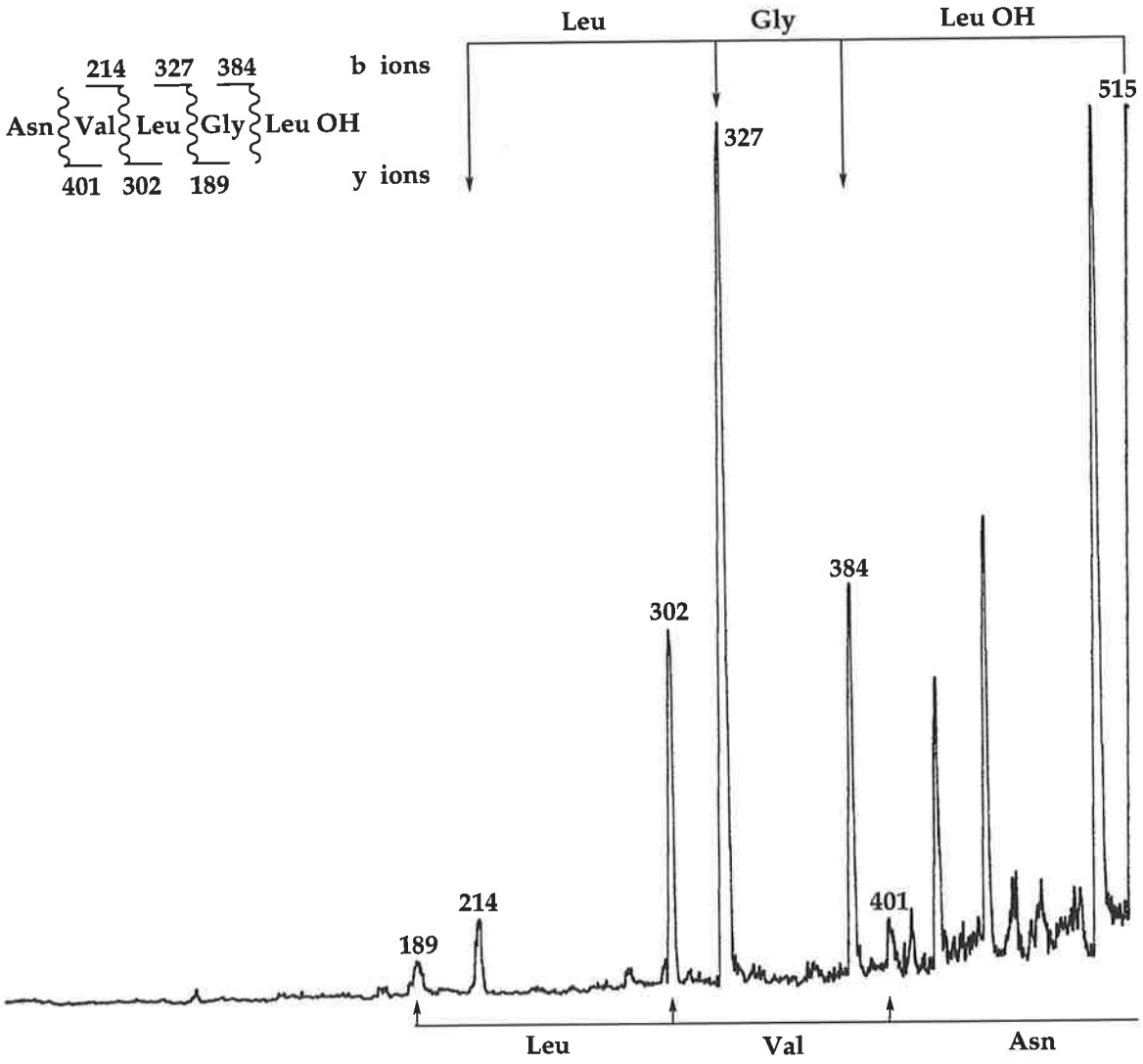


Figure **4.4**: CA FAB MS/MS spectrum of peptide **2.8F** ($[M + H]^+$ =515), derived from the Arg-C digest of uperin 2.8. Cleavages associated with 'b' ions are indicated at the top of the figure, 'y' ions at the bottom.

Uperin 3.4

(a) FAB MS: $[M + H]^+ = 1735$.

(b) HPLC retention time: 24.3 minutes (Figure **4.3**). Uperin 3.4 coelutes with uperin 2.7, which were separated by HPLC upon repurification. Uperin 3.4 is the major component in this fraction.

(c) Esterification of uperin 3.4 with methanol increased the mass by 44 Da, indicating the presence of one CO_2H group and two NH_2 groups.

(d) Tables **4.6** and **4.7** list the results of enzymic digests and CA MS/MS experiments, respectively.

Table **4.6**: Results of the Enzyme Digests of Uperin 3.4.

Peptide Digested	Enzyme Used	Resultant Peptide	$[M + H]^+$
Uperin 3.4	Lys-C	**3.4A**	344[a]
		3.4B	572[a]
		3.4C	857[b]
3.4C	Asp-N	**3.4D**	232[a]
		3.4E	644[a]

[a] CA MS/MS studies performed on this ion.

[b] CA MS/MS studies performed on this ion were inconclusive.

Table **4.7**: CA MS/MS Data or the Enzyme Products from Uperin 3.4.

Peptide	***m/z***	**Ion**	**Observed fragment ions [*m/z* (%)]**	**Peptide sequence**
3.4A	344	'a'	200 (3)	Val
		'b'	228 (75); 327 (100)	Val
		'y'	[none detected]	
Sequence of **3.4A** by CA MS/MS data:				**(Asn + Leu) Val (NH_2)**
3.4B	572	'b'	242 (11); 313 (42); 426 (56); 554 (100)	Ala Leu Lys
		'y'	501 (67); 402 (31); 331 (12); 260 (12)	Ala Val Ala Ala
Sequence of **3.4B** by CA MS/MS data:				**Ala Val Ala Ala Leu Lys (OH)**
3.4D	232	'b'	157 (100); 214 (50)	Gly
		'y'	175 (18)	Gly
Sequence of **3.4D** by CA MS/MS data:				**Gly (Val) Gly (OH)**
3.4E [a]	644	'b'	229 (31); 498 (63); 626 (100)	Leu Arg Lys
		'y'	529 (68); 303 (17)	Asp
		'z'	514 (59); 401 (26); 288 (21)	Asp Leu Leu
Sequence of **3.4E** by CA MS/MS data:				**Asp Leu Leu Arg Lys (OH)**

[a] CA MS/MS spectra of peptide **3.4E** also show peaks at *m/z* 201 (21%), 470 (54%) and 555 (77%) which are respectively associated with the cleavage ions 'a_2', 'a_4' and 'x_4'. Cleavage ions 'a', 'x' and 'z' are rarely observed in other spectra; in this particular case the culmination of all the data gives the sequence.

Since **3.4C** gave unsatisfactory spectra, it was digested with endoprotease Asp-N to give two smaller peptides. **3.4D** and **3.4E** are confirmed to be Gly Val Gly (OH) and Asp Leu Leu Arg Lys (OH) respectively, which are also residues 1 - 3 and 4 - 8 of uperin 3.4. **3.4A** must be the last three residues of uperin 3.4 since it is the product of the Lys-C digest and does not contain a C-terminal Lys residue. **3.4B** must be the central portion of the parent peptide.

(f) The sequence from MS data gives Gly Val Gly Asp Leu Leu Arg Lys Ala Val Ala Ala Leu Lys Asn Leu Val (NH_2). Automated sequencing identifies Asn (15) and Ile (6, 13 and 16) and confirms the sequence.

Uperin 3.5

(a) FAB MS: $[M + H]^+ = 1779$.

(b) HPLC retention time: 25.8 minutes (Figure **4.3**). Uperin 3.5 coelutes with uperin 2.8, which were separated by HPLC upon repurification. Uperin 3.5 is the major component in this fraction.

(c) Esterification of uperin 3.5 with methanol increased the mass by 44 Da, indicating the presence of one CO_2H group and two $CONH_2$ groups.

(d) The results of enzymic digests are listed in Table **4.8**, and the results of CA MS/MS experiments are recorded in Table **4.9**.

Table **4.8**: Results of the Enzyme Digests of Uperin 3.5.

Peptide Digested	Enzyme Used	Resultant Peptide	$[M + H]^+$
Uperin 3.5	Lys-C	**3.5A**	344[a]
		3.5B	616[a]
		3.5C	857[a]
3.5C	Asp-N	**3.5D**	232[a]
		3.5E	644

[a] CA MS/MS studies performed on this ion.

Table 4.9: CA MS/MS Data Observed for Peptides from Enzyme Digests.

Peptide	*m/z*	Ion	Observed fragment ions [*m/z* (%)]	Peptide sequence
3.5A	344	'a'	200 (8)	Val
		'b'	228 (100); 327 (93)	Val
		'y'	[none observed]	
Sequence of **3.5A** by CA MS/MS data:				**(Asn + Leu) Val (NH_2)**
3.5B	616	'b'	258 (3); 357 (7); 470 (22); 598 (100)	Val Leu Lys
		'y'	545 (50); 446 (14); 359 (15); 260 (6)	Ala Val Ser Val
Sequence of **3.5B** by CA MS/MS data:				**Ala Val Ser Val Leu Lys (OH)**
3.5C	857	'a'	527 (7); 683 (10)	Arg Lys
		'b'	711 (11); 839 (100)	Lys
		'y'	800 (32); 701 (44)	Gly Val
Sequence of **3.5C** by CA MS/MS data:				**Gly Val (398) Arg Lys (OH)**
3.5D	232	'b'	157 (100); 214 (61)	Gly
		'y'	175 (38)	Gly
Sequence of **3.5D** by CA MS/MS data:				**Gly (Val) Gly (OH)**

These experiments show that **3.5D** must be the first three residues of uperin 3.5, since it is a product of the Asp-N digest, and does not contain the Asp residue. Peptide **3.5C** must therefore have the partial sequence Gly Val Gly Asp (226) Arg Lys (OH). The unidentified portion could be either (Leu + Leu) or (Pro + Glu). It has been established from the methylation experiment that there is only one acidic residue in the sequence, thus eliminating the combination of (Pro + Glu). Peptide **3.5A** must be the last three residues of uperin 3.5 since it is the product of the Lys-C digest, and **3.5B** must be the central portion of the parent peptide comprising of residues 9 - 14.

(f) The sequence from MS data gives Gly Val Gly Asp Leu Leu Arg Lys Ala Val Ser Val Leu Lys Asn Leu Val (NH_2). Automated sequencing identifies Asn (15), Leu (5) and Ile (6, 13 and 16). The structure of uperin 3.5 is listed in Table **4.1.**

Uperin 3.6

(a) FAB MS: $[M + H]^+ = 1826$.

(b) HPLC retention time: 23.0 minutes (Figure **4.3**). Uperin 3.6 was the only identifiable compound in this fraction.

(c) Esterification of uperin 3.6 with methanol increased the mass by 44 Da, indicating the presence of one CO_2H group and two $CONH_2$ groups.

(d) The results of enzymic digests are listed in Table **4.10**, and the results of CA MS/MS experiments are recorded in Table **4.11**.

Table **4.10**: Results of the Enzyme Digests of Uperin 3.6.

Peptide Digested	Enzyme Used	Resultant Peptide	$[M + H]^+$
Uperin 3.6	Lys-C	**3.6A**	392[a]
		3.6B	671[a]
		3.6C	673[a]
		3.6D	799
		3.6E	801
Uperin 3.6	Asp-N	**3.6F**	288[a]
		3.6G	1557

[a] CA MS/MS studies performed on this ion.

Table **4.11**: CA MS/MS Data Observed for Peptides from Enzyme Digests.

Peptide	***m/z***	**Ion**	**Observed fragment ions [*m/z* (%)]**	**Peptide sequence**
3.6A	392	'b'	115 (3); 228 (29); 375 (100)	Leu Phe
		'y'	278 (6); 165 (11)	Asn Leu
Sequence of **3.6A** by CA MS/MS data:				**Asn Leu Phe (NH_2)**
3.6B	671	'b'	199 (20); 313 (46); 412 (80); 525 (84); 653 (100)	Asn Val Leu Lys
		'y'	572 (62); 473 (40); 359 (11); 260 (18); 147 (11)	Val Val Asn Val Leu
Sequence of **3.6B** by CA MS/MS data:				**Val Val Asn Val Leu Lys (OH)**
3.6C	673	'b'	385 (8); 456 (21); 527 (41); 655 (100)	Ala Ala Lys
		'y'	616 (77); 517 (28); 404 (12); 289 (7)	Gly Val Leu Asp
Sequence of **3.6C** by CA MS/MS data:				**Gly Val Leu Asp Ala Ala Lys (OH)**
3.6F	288	'b'	270 (100)	
		'y'	231 (23)	Gly
Sequence of **3.6F** by CA MS/MS data:				**Gly (Val Leu) (OH)**

Peptide **3.6A** is the last three residues of uperin 3.6, since it has a C-terminal Phe (NH_2). Peptide **3.6F** is the first three residues of uperin 3.6 as it does not contain the Asp residue, and thus peptide **3.6C** must be the first seven residues. Peptide **3.6B** must be the central portion of uperin 3.6. Two peptides of small abundance (relative to peptide **3.6B**) resulting from the incomplete Lys-C digest of uperin 3.6 were observed at *m/z* 799 (**3.6D**) and 801 (**3.6E**). Peptide **3.6D** corresponds to Lys + **3.6B**; peptide **3.6E** corresponds to **3.6C** + Lys.

(f) The sequence from MS data gives Gly Val Leu Asp Ala Ala Lys Lys Val Val Asn Val Leu Lys Asn Leu Phe (NH_2). Automated sequencing identifies Ile (3) and confirms the sequence shown in Table **4.1**.

Uperin 3.7

(a) FAB MS: $[M + H]^+ = 1844$.

(b) HPLC retention time: 23.5 minutes (Figure **4.3**). Uperin 3.7 coelutes with uperin 2.6 and two unidentified peptides. Repurification by HPLC separated uperins 2.6 and 3.7.

(c) Esterification of uperin 3.7 with methanol increased the mass by 44 Da, indicating the presence of one CO_2H group and two $CONH_2$ groups in the peptide.

(d) The results of enzymic digests are listed in Table **4.12**, and the results of CA MS/MS experiments are recorded in Table **4.13**.

Table **4.12**: Results of the Enzyme Digests of Uperin 3.7.

Peptide Digested	Enzyme Used	Resultant Peptide	$[M + H]^+$
Uperin 3.7	Lys-C	**3.7A**	330[a]
		3.7B	660[a]
		3.7C	891[b]
Uperin 3.7	α-Chymotrypsin	**3.7D**	607[a]
		3.7E	1256
Uperin 3.7	Asp-N	**3.7F**	232[a]
		3.7G	1631

[a] CA MS/MS studies performed on this ion.

[b] CA MS/MS studies performed on this ion were inconclusive.

Table **4.13**: CA MS/MS data observed for peptides from enzyme digests.

Peptide	***m/z***	**Ion**	**Observed fragment ions [*m/z* (%)]**	**Peptide sequence**
3.7A	330	'b'	214 (27); 313 (100)	Val
		'y'	[none observed]	
Sequence of **3.7A** by CA MS/MS data:				**(Asn + Val) Val (NH_2)**
3.7B	660	'b'	401 (12); 514 (17); 642 (100)	Leu Lys
		'y'	547 (21); 448 (16); 361 (6)	Leu Val Ser
Sequence of **3.7B** by CA MS/MS data:				**Leu Val Ser (Thr) Leu Lys (OH)**
3.7D	607	'b'	329 (18); 442 (98); 589 (100)	Leu Phe
		'y'	550 (72); 451 (37); 394 (12)	Gly Val Gly
Sequence of **3.7D** by CA MS/MS data:				**Gly Val Gly Asp Leu Phe (OH)**
3.7F	232	'a'	129 (10)	Gly
		'b'	157 (100); 214 (98)	Gly
		'y'	174 (38)	Gly
Sequence of **3.7F** by CA MS/MS data:				**Gly (Val) Gly (OH)**

Peptide **3.7A** has a C-terminal Val (NH_2) suggesting that this is the last residue of uperin 3.7, and peptide **3.7F** is the first three residues since it does not contain an Asp residue. Peptide **3.7D** begins with Gly Val Gly and does not end with Val (NH_2), thus **3.7D** must therefore be the first six residues of uperin 3.7. This also implies that **3.7C** must be **3.7D** + (156) + Lys. **3.7B** must be the central portion of the sequence spanning residues 9 - 14.

(f) The sequence from MS data gives Gly Val Gly Asp Leu Phe Arg Lys Leu Val Ser Thr Leu Lys Asn Val Val (NH_2). Automated sequencing identifies Asn (15), Val (16) and Ile (5, 9 and 13) and confirms the sequence which is shown in Table **4.1**.

4.3 Structural Features of the Uperin Peptides

The seven uperins isolated from *U. mjobergi* have similar structures to those isolated from the dermal secretions of *U. inundata*. Thus, these peptides have been classified in the same manner as the uperins reported in Chapter 3. The most surprising observation arising from this study is that the skin secretion of *U. mjobergi* does not contain a neuropeptide of the tachykinin family; neither uperolein (**15**), uperin 1.1 (**2**) nor any related peptide. This is most unusual since either uperolein or uperin 1.1 are major skin peptides of the three *Uperoleia* species so far studied.[275, 335] Equally surprising is the lack of peptides of the uperin 4-6 category. Fifteen of the twenty uperins isolated from the dermal glands of both frogs are uperin 2 type and 3 type peptides. This suggests that these particular peptides must play a significant role in the amphibians' integument.

There is a view that the structures of biologically active amphibian peptides should correlate with those of peptides from other organisms, in particular with mammalian peptides.[151, 174, 336] A GENINFO (R) BLAST Network Service (Cruncher) data bank search (Australian National Genomic Information Service)[332] has been carried out for the uperin 2 and 3 structures. This has shown no statistically significant correlation with any peptides or proteins so far reported from other organisms.

It was shown in Chapter 3 how uperin 2 and 3 type peptides could be fitted to well behaved α-helices using the Edmundson wheel and helical net diagrams: all of the uperins reported in this chapter, when projected in the same manner, also show well defined hydrophobic and hydrophilic zones. Such a property suggests the possibility of antimicrobial or antibacterial activity.

4.4 Biological Activity of the Uperin Peptides from *U. mjobergi*

Synthetic analogues* of uperins 2.8, 3.5 and 3.6 were submitted for antimicrobial testing. The results are listed in Table **4.14**, which also lists the activities of a select range of uperins from *Uperoleia inundata*, enabling a direct comparison of the activities between the two species. Table **4.14** clearly illustrates that uperins 3.5 and 3.6 (from *Uperoleia mjobergi*) are the most active antibiotic agents yet isolated from the *Uperoleia* genus. Further more, they show significant activity against a range of gram positive organisms. To test whether the uperins acted independently or as a 'cocktail' of peptides, uperins 3.5 and 3.6 were tested as a mixture. The results are also listed in Table **4.14**, which shows that the activity of the mixture neither greatly increases nor decreases in comparison to either of the authentic samples. Further experiments must be carried out with differing pairs and 'cocktails' of uperins in order to establish if there is a significant trend. If there is a general increase in the biological activity of a 'cocktail' of uperins, then it would most certainly be worth investigating 'cocktail' mixtures of peptides from other frog species.

* See Experimental (section **6.9**) for details of the synthetic procedures. Not all uperins could be synthesised due to financial constraints. Those which were selected were purified by HPLC and were also shown to have the same retention times, molecular weights and the same CA MS/MS spectra as the natural components.

Table **4.14**: Comparison of Uperin 2 and 3 Antibiotic Activities [MIC (μg/mL)].

Organism	**2.1**	**2.2**	**2.4**	**2.8**	**3.1**	**3.5**	**3.6**	**3.5 + 3.6**
Bacillus cereus	[a]			100		25	25	
Escherichia coli [c]								
Leuconostoc lactis	-[b]		6	6	25	3	3	6
Leuconostoc mesenteroides	6		-	-	-	-	-	-
Listeria innocua	1.5					50	25	50
Micrococcus luteus					100	25	12.5	25
Pasteurella haemolytica						-	-	-
Pasteurella multoci							100	50
Staphylococcus aureus						25	50	12.5
Streptococcus epidermis				100	100	12.5	12.5	25
Streptococcus uberis	50		25	50	25	12.5	12.5	50

[a] no figure means the MIC value is > 100 μg/mL.

[b] a dash (-) means not tested.

[c] this organism is gram negative, all others are gram positive.

4.5 Primary Sequence and Biological Activity Relationships

All the uperins 2 and 3 type peptides are listed in Table **4.15**. This table shows that there are distinct homologies between peptides of the same group. All of the amino acids common to the class of peptide are highlighted. The amino acids which differ between structures show a striking trend: hydrophobic residues replace hydrophobic residues at the same position and the same scenario occurs for hydrophilic residues. The exception being residue (12) of uperin 3.7, which has a threonine in place of a hydrophobic residue. Uperins 3.1 - 3.3 from *U. inundata* are very conservative, as the changes between the structures occur only at positions 7 and 16. These peptides show antimicrobial activity, albeit to a lesser degree than the uperin 3 peptides from *U. mjobergi*, which are less conserved. It is also interesting to note that the number of

Table 4.15: Amino Acid Sequence of Uperins 2.1 - 2.8 and 3.1 - 3.7.

Uperin	$[M + H]^{+}$ [a]	Amino Acid Sequence	Amphibian
2.1	1926	**Gly** Ile Val **Asp** Phe **Ala Lys** Lys Val Val **Gly** Gly Ile Arg **Asn** Ala **Leu Gly Ile (OH)**	*U. inundata*
2.2	1926	**Gly** Phe Val **Asp** Leu **Ala Lys** Lys Val Val **Gly** Gly Ile Arg **Asn** Ala **Leu Gly Ile (OH)**	*U. inundata*
2.3	1974	**Gly** Phe Phe **Asp** Leu **Ala Lys** Lys Val Val **Gly** Gly Ile Arg **Asn** Ala **Leu Gly Ile (OH)**	*U. inundata*
2.4	1913	**Gly** Ile Leu **Asp** Phe **Ala Lys** Thr Val Val **Gly** Gly Ile Arg **Asn** Ala **Leu Gly Ile (OH)**	*U. inundata*
2.5	1940	**Gly** Ile Val **Asp** Phe **Ala Lys** Gly Val Leu **Gly** Lys Ile Lys **Asn** Val **Leu Gly Ile (OH)**	*U. inundata*
2.6	1948	**Gly** Ile Leu **Asp** Ile **Ala Lys** Lys Leu Val **Gly** Gly Ile Arg **Asn** Val **Leu Gly Ile (OH)**	*U. mjobergi*
2.7	1948	**Gly** Ile Ile **Asp** Ile **Ala Lys** Lys Leu Val **Gly** Gly Ile Arg **Asn** Val **Leu Gly Ile (OH)**	*U. mjobergi*
2.8	1978	**Gly** Ile Leu **Asp** Val **Ala Lys** Thr Leu Val **Gly** Lys Leu Arg **Asn** Val **Leu Gly Ile (OH)**	*U. mjobergi*
3.1	1827	**Gly Val** Leu **Asp** Ala Phe Arg **Lys** Ile Ala Thr Val Val **Lys Asn** Val Val **(NH_2)**	*U. inundata*
3.2	1841	**Gly Val** Leu **Asp** Ala Phe Arg **Lys** Ile Ala Thr Val Val **Lys Asn** Leu Val **(NH_2)**	*U. inundata*
3.3	1813	**Gly Val** Leu **Asp** Ala Phe Lys **Lys** Ile Ala Thr Val Val **Lys Asn** Leu Val **(NH_2)**	*U. inundata*
3.4	1735	**Gly Val** Gly **Asp** Leu Ile Arg **Lys** Ala Val Ala Ala Ile **Lys Asn** Ile Val **(NH_2)**	*U. mjobergi*
3.5	1779	**Gly Val** Gly **Asp** Leu Ile Arg **Lys** Ala Val Ser Val Ile **Lys Asn** Ile Val **(NH_2)**	*U. mjobergi*
3.6	1826	**Gly Val** Ile **Asp** Ala Ala Lys **Lys** Val Val Asn Val Leu **Lys Asn** Leu Phe **(NH_2)**	*U. mjobergi*
3.7	1843	**Gly Val** Gly **Asp** Ile Phe Arg **Lys** Ile Val Ser Thr Ile **Lys Asn** Val Val **(NH_2)**	*U. mjobergi*

[a] Nominal mass

acidic and basic residues (*i.e.* Arg, Lys and Asp) remains constant and that the nett charge* is also the same. The nett charge of uperins 2.1 - 2.8 is +1, and for uperins 3.1 - 3.7 it is +2.† The charge of a peptide plays a key role in the electrostatic attraction of the peptide for phospholipids. This property may help explain why position 4 in these uperins is occupied by aspartic acid; similarly for lysine and arginine which occupy positions 7 and 14: conclusive studies using modified peptide analogues will clarify this matter.

4.6 Summary of the Research in Chapters 3 and 4

The characterisation of all of the amphibian peptides were facilitated by the use of positive ion fast atom bombardment mass spectrometry. Collisional activation in conjunction with tandem mass spectrometry techniques and enzymic digests provided the sequences of the peptides, which were confirmed by automated sequencing.

The investigation in to the isolation and characterisation of the peptides from *Uperoleia inundata* and *U. mjobergi* has shown that there are a number of peptides in the glandular secretions. Some of which are structurally related, although no peptide is common to both frogs. The results of the biological activities show that the uperin 2 and 3 type peptides are host defence peptides: in particular the uperin 3 type peptides exhibit significant activity against gram positive organisms. The role of the uperins which showed little activity in the testing program, is as yet unknown.

---

* The nett charge is the summation of the individual charges of all the acidic and basic residues in the sequence if they were deprotonated and protonated respectively.

† All uperins have a positive nett charge, except for uperin 5.1 which has a nett charge of -2, and uperin 1.1, which is -1.

The secretion from *Uperoleia mjobergi* contains no neuropeptide of the tachykinin class of peptides, which should be compared with the high abundance of such compounds isolated from other *Uperoleia* species:[275, 335] *U. inundata* has a neuropeptide, named uperin 1.1, which is one of the most active amphibian tachykinins which have so far been isolated.

Chapter 5 GAS-PHASE FRAGMENTATIONS OF DEPROTONATED TETRAPEPTIDES

"I believe there is no branch of science where promise of great discoveries is more hopeful than those which will result by researches which involve the application of physical measurements to chemical phenomena."

Sir J. J. Thomson[337]

5.1 Introduction to Gas-Phase Negative Ion Fragmentations

The negative ion mass spectra reported in this thesis were generated by CA MIKES of even electron organic anions, produced by FAB in the ion source of the VG ZAB 2 HF mass spectrometer operated in the negative ion mode. This instrumentation has been discussed in Chapter 1. Fragmentation processes of even electron organic anions have been classified into five main fragmentation categories:[338]

(i) loss of a radical to form a stabilised radical anion;

(ii) direct fragmentation through an intermediate ion-complex with subsequent elimination of a neutral molecule from the ion-complex;

(iii) fragmentation preceded by proton transfer to the initial site of deprotonation followed by elimination of a neutral molecule;

(iv) fragmentations which are preceded by skeletal rearrangement;

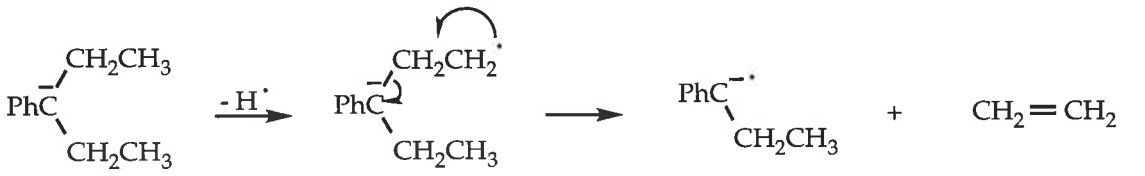
(v) charge remote fragmentations.

A brief summary of these processes are outlined on the following pages.

5.2 Fragmentation *via* loss of a Radical

Collision-induced mass spectra of most deprotonated organic anions involve the loss of a radical, typically a hydrogen radical, *via* a simple homolytic cleavage. The fragmentations are pronounced in the spectra if the radical forms a stabilised radical anion [scheme (**5.1**)].[339] In some systems, the elimination of an alkyl chain has been shown to occur *via* a two step process [scheme (**5.2**)].[340, 341]

(**5.1**)



(**5.2**)

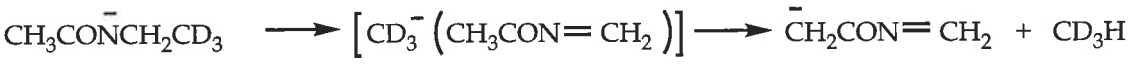
5.3 Fragmentation *via* an Ion-molecule Complex

An ion-molecule complex may be formed directly from the deprotonated ion. These complexes fragment *via* three possible pathways:

(i) direct displacement of the 'bound' anion from the ion-complex, *e.g.* scheme (**5.3**).[339]

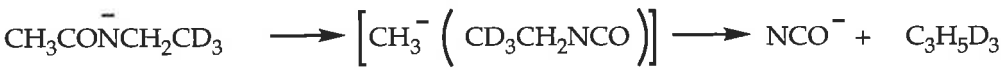
(**5.3**)

(ii) deprotonation of the 'neutral' molecule by the 'bound' anion in the ion-complex, illustrated by scheme (**5.4**).[339]



(**5.4**)

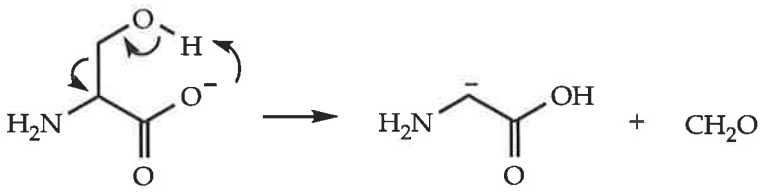
(iii) an elimination reaction or S_N2 displacement of the 'neutral' molecule effected by the 'bound' anion in the ion-complex, *e.g.* scheme (**5.5**).[339]



(**5.5**)

5.4 Proton Transfer Preceding Fragmentation

Direct elimination of a neutral species can occur following proton transfer to the initial anionic centre [scheme (**5.6**)].[342]



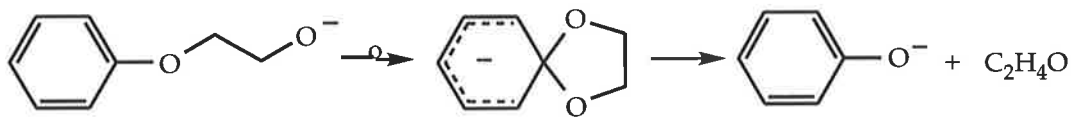
(**5.6**)

Also, proton transfer to the initial anionic centre can precede fragmentation of the ion-complex, as discussed in section **5.3** above.[338]

5.5 Fragmentations Following Rearrangement

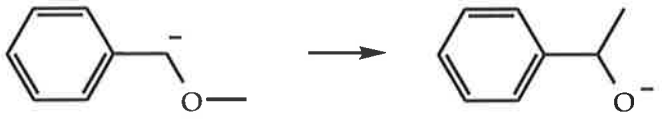
Processes involving a skeletal rearrangement form interesting and important pathways for the fragmentation of collisionally activated organic anions. Such processes have been grouped as:

(i) Cyclization processes, where there is a nucleophilic attack by the initially formed anion, as in scheme (**5.7**).[343]



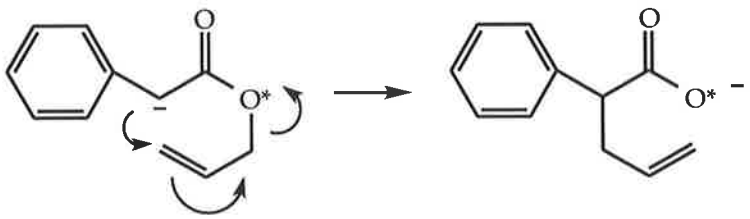
(**5.7**)

(ii) 1, 2- anionic rearrangements, *e.g.*, Wittig [scheme (**5.8**)].[344]



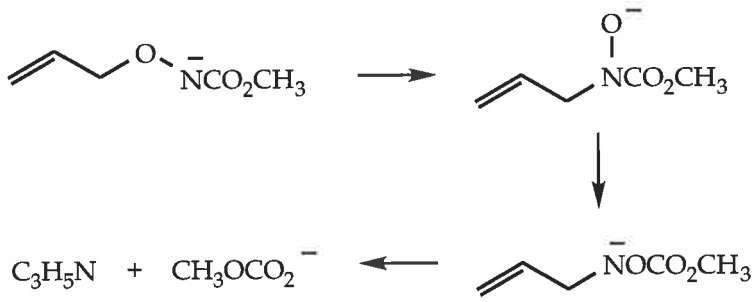
(**5.8**)

(iii) Six-centred sigmatropic rearrangements such as the anion Claisen ester type rearrangement [scheme (**5.9**)].[345]



(**5.9**)

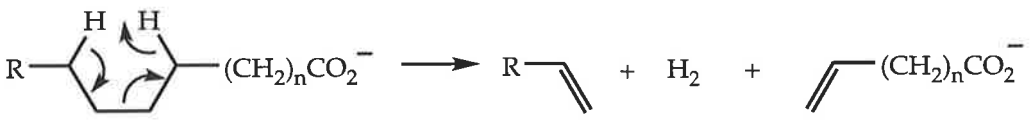
(iv) Novel gas-phase rearrangements can occur as detailed in scheme (**5.10**).[346]



(**5.10**)

5.6 Charge Remote Fragmentations[347]

Charge remote fragmentations are defined[348] as gas-phase ion decompositions that occur at sites in the gas-phase molecule that are physically removed from the charged centre. Gross has reported a classical example of a charge-remote reaction using the anion of stearic acid [$CH_3(CH_2)_{16}CO_2^-$], which is illustrated in scheme (**5.11**), where R is the remainder of the organic chain.[349] Elimination of an alkene in these reactions has been confirmed by ionisation of the eliminated neutral alkene.[350]

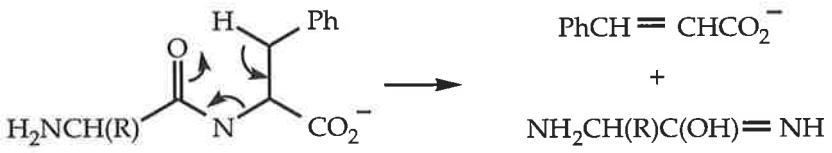


(**5.11**)

The mechanism involves a 1, 4- loss of H_2 to give a terminally unsaturated fatty acid carboxylate and a 1- alkene. Charge remote fragmentations have also

been claimed to occur for prostaglandins, phospholipids, peptides, steroids, glycosphingolipids, fatty acids, and carbohydrates [for a review see ref (348)].

It should be noted that some anionic cleavages discussed in this thesis could possibly occur by charge remote mechanisms. This is illustrated in scheme (**5.12**) for the particular example of cinnamate anion formation from phenylalanine occupying the C-terminal position of a dipeptide.



(**5.12**)

5.7 Introduction to Peptide Negative Ions

FAB MS has facilitated the structural analysis of peptides.[329] Current research into peptide sequencing by mass spectrometry has primarily been dominated by positive ion analysis. The application of negative ion mass spectrometry to peptide analysis has principally been used to detect deprotonated molecular ions $[M - H]^-$;[351, 352] there are a few exceptions which involve fragmentation studies.[353-356]

The collision induced fragmentation behaviour of deprotonated peptide ions has, until recently, been little understood and there is rapid acceptance that such data can provide complementary information to the conventional positive ion spectra.[351-362] The Bowie research group has been investigating this area and to date has proposed mechanisms for the collision induced dissociations of underivatised deprotonated and dedeuterated amino acids,[342] dipeptides,[363-370] and tripeptides.[363, 369, 370]

5.8 Deprotonated Amino Acids[342]

A previous study[342] reported that deprotonated amino acids which contain a specific functional group fragment quite differently from simple amino acids that have alkyl side chains. The spectra of these deprotonated amino acids show pronounced fragmentations through their α-side chains;[342] the losses and formations associated with such fragmentations are listed in Table **5.1**.

5.9 Dipeptides and Tripeptides[363-370]

There are two very important general types of collision-induced fragmentations observed for deprotonated ions from dipeptides and tripeptides which can be used analytically. These are: (i) backbone cleavages which provide the primary sequencing information; and (ii) fragmentations characteristic of particular amino acids which are either: (a) in specific positions within the peptide, *i.e.* at the N-terminal or C-terminal position only; or (b) non-specific, *i.e.* these fragmentations occur irrespective of the position of the amino acid within the peptide. Table **5.1** also lists side chain fragmentations observed for amino acid residues at either position in a dipeptide.

(i) Backbone Cleavages

The backbone cleavages of deprotonated dipeptide ions are considered to proceed through the intermediate enolate ion formed by proton transfer to the

Table **5.1**: Losses (or Formations) Observed for Amino Acids and Dipeptides.

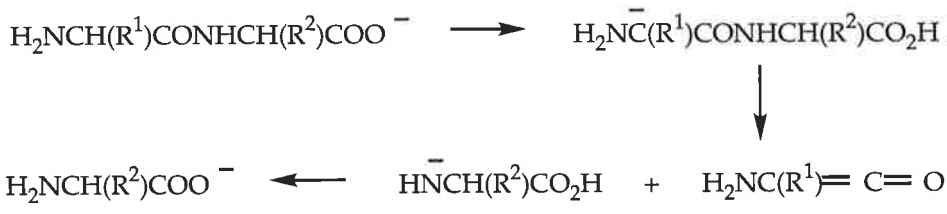
Amino acid [a]	Loss (or formation)	Amino acid*	Dipeptide	
			N-terminal	C-terminal
Ala	Me·	Y [b]	Y	Y
Val	*i*Pr·	Y	Y	Y
Leu	*t*Bu·	Y	Y	Y
Ile	*s*Bu·	Y	Y	Y
Ser	CH_2O	Y	Y	Y
Thr	CH_3CHO	Y	Y	Y
Cys	H_2S	Y	Y	Y (+ CO_2)
Gln / Asn	NH_3	Y	Y (+ CO_2)	N
Asn	H_2O	Y	N	Y
Asp / Glu	H_2O	Y	Y	Y
Glu	pyroglutamate anion	N	Y	Y
Met	MeSH	Y	Y	Y
	MeSMe	Y	Y	Y
	$\cdot CH_2CH_2SMe$	Y	Y	Y
His	H_2O	Y	Y (+CO_2)	Y (+CO_2)
Phe	$PhCH_3$	m	Y	N
	($PhCH_2^-$)	m	Y	N
Phe / Tyr	NH_3	Y	Y (+ CO_2)	N
	($ArCH{=}CHCO_2^-$)	N	N	Y
Tyr	$O{=}C_6H_4{=}CH_2$	N	Y	Y
	(p-$HOC_6H_4CH_2^-$)	N	Y	Y
Arg	NH_3	Y	Y	m
	$HN{=}C{=}NH$	Y	m	Y (+ CO_2)
Trp	C_9H_7N	N	Y	Y

[a] Amino acids are represented by the three letter code.

[b] The letters correspond to: Y = observed in the spectra; N = not observed in the spectra; m = observed, but is a minor fragmentation pathway.

* Column 3 summarises whether (or not) the particular side chain fragmentation listed in Column 2, occurs from the (M-H)$^+$ ion of the amino acid designated in Column 1.

carboxylate centre,* but direct proton abstraction to yield the enolate anion is also possible. The spectra are generally simple and are of analytical applicability and importance. As an example, the characteristic backbone cleavage of dipeptides that identifies the C-terminal amino acid is shown in scheme (**5.13**),[363] where R^1 and R^2 = H or alkyl.



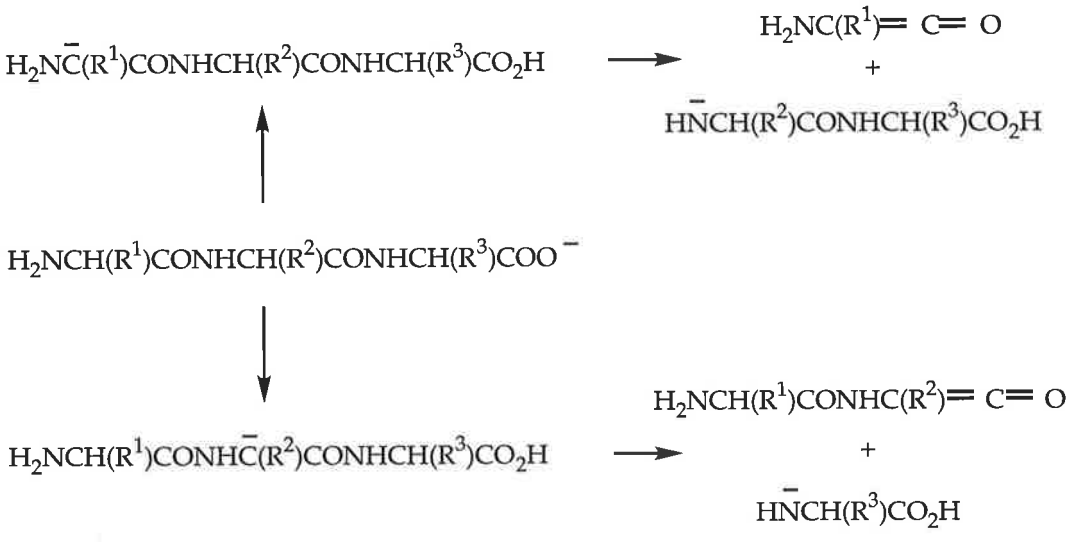
(**5.13**)

There are a number of acidic sites in peptides, *e.g.* R-CO_2**H**, R-N**H**CO-R and R-C**H**CONH-R. Therefore, the position(s) of deprotonation in a peptide ionised by FAB cannot be determined qualitatively, even if deuterium labelling is used. An added complication is the presence of both 'neutral' and 'zwitterion' forms of the peptide in the FAB matrix. When an unlabelled peptide is analysed by FAB, minor amounts of enolate ions (together with R-CON$^-$-R species) will be formed together with the more abundant carboxylate species. However, in a typical labelling experiment, exchange of the acidic hydrogens [*e.g.* in the dipeptide Ala-Gly (OH)] with D_2O gives $ND_2CH(Me)CONDCH_2CO_2D$ and FAB ionisation yields mainly an $[M - D]^-$ species, which can have no direct enolate ion contributor. The characteristic fragmentation of the $[M - D]^-$ ion must occur following proton transfer.

---

* This is a general feature of the fragmentation behaviour of simple alkyl carboxylic acids.[371] For example, in the prototypical case of acetic acid, deprotonation forms mainly $MeCO_2^-$, but $(CH_2CO_2H)^-$ is also identified.[372] ΔG°_{acid} ($C\mathbf{H}_3CO_2H$) = 363 kcal mol^{-1} [372, 373] and ΔG°_{acid} ($MeCO_2\mathbf{H}$) = 341.5 kcal mol^{-1} [374]; and the two ions are interconvertible upon collisional activation.

In the case of tripeptides, an additional and characteristic cleavage proceeds through the enolate ion formed at the N-terminal end of the peptide.(363) This results in the observation of two backbone cleavages as illustrated in scheme (**5.14**), and the central residue is easily determined by deductive intuition.

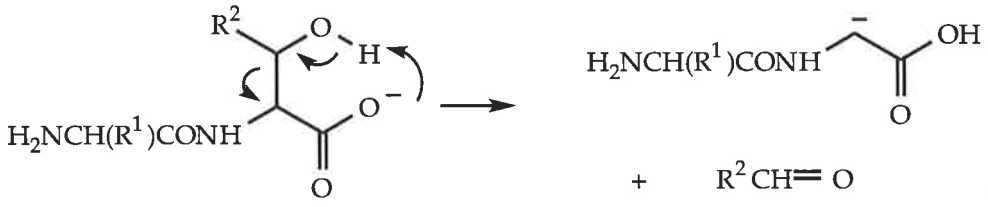


(**5.14**)

(ii) Fragmentations Characteristic of Amino Acids in Di- and Tripeptides

As with backbone cleavages, fragmentations characteristic of particular amino acids are also observed in deprotonated spectra of dipeptides and tripeptides, and are either independent of location in the peptide, or occur exclusively at a particular position in the peptide. As a general rule, fragmentations that are independent of location (*i.e.* non-specific), occur *via* processes where initial fragmentation occurs in the side chain after either: nucleophilic attack on the side chain by the carboxylate anion; or from proton transfer to the carboxylate centre from the α-side chain only. There are many examples of such processes; these include the loss of formaldehyde and acetaldehyde, which occur for serine and threonine respectively. The proposed mechanism is outlined in

scheme (**5.15**)[364] using a dipeptide, but readily occurs for tripeptides regardless of the residue's position. This mechanism is similar to scheme (**5.6**).



(**5.15**)

where R^2 = -H (*i.e.* amino acid is serine)

R^2 = $-CH_3$ (threonine)

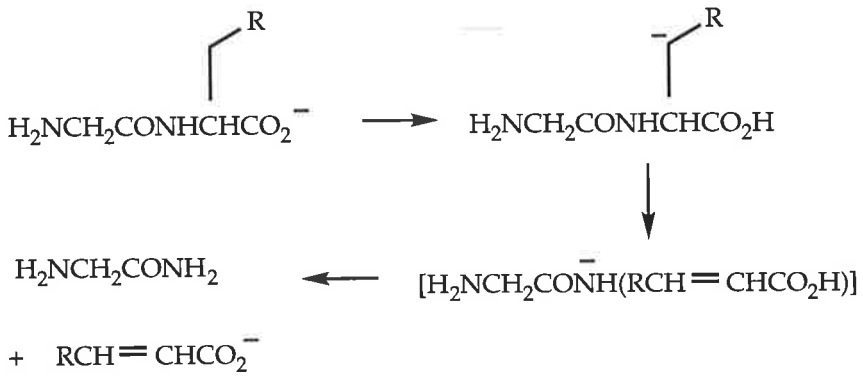
As a general rule, fragmentations that are characteristic of amino acids in specific location occur by processes which fragment the peptide backbone. In such cases, proton transfer to the carboxylate centre from either the peptide skeleton or β-carbon will result in different fragmentations, depending on the location of the specified amino acid in the peptide. For example, the isomeric dipeptides glycyl-phenylalanine and phenylalanyl-glycine can both be deprotonated at the β-carbon, with fragmentation generating the cinnamate anion and loss of NH_3 respectively [schemes (**5.16**) and (**5.17**)].[365]* The proposed mechanisms can also be applied to residues such as tyrosine† and asparagine.[367]‡ Occasionally the loss of NH_3 is accompanied by loss of CO_2.

---

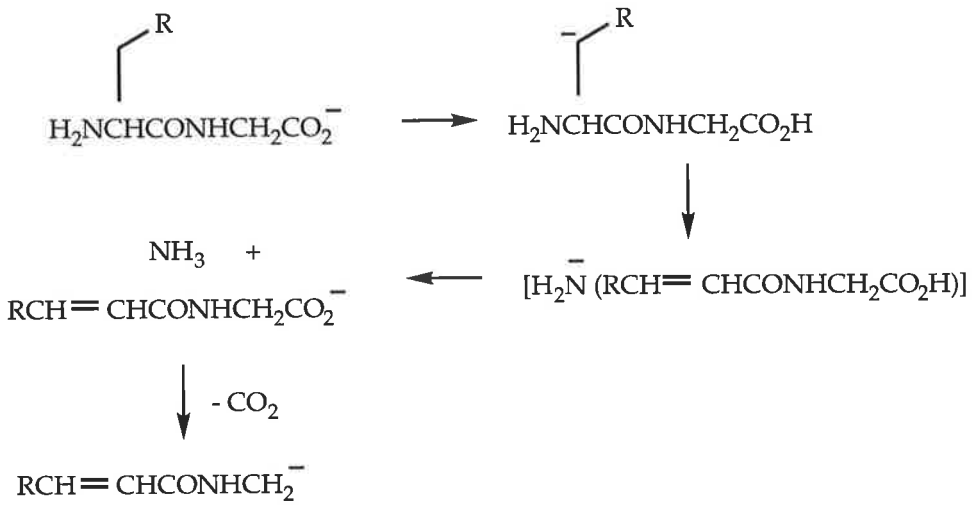
* This mechanism may be charge remote for Gly Phe (see section 5.6). However, Phe Gly cannot eliminate NH_3 *via* a charge remote process.

† Tryptophan and histidine also have β-carbons bearing acidic hydrogens, however deprotonation at these sites yields trace amounts of the expected products.[365]

‡ The loss of NH_3 from asparagine can result from the loss of the side chain amide, or the N-terminal amine.[367]



(**5.16**)



(**5.17**)

where R = $-C_6H_5$ (phenylalanine)
$-C_6H_4OH$ (tyrosine)
$-CH_2CONH_2$ (asparagine)

Since dipeptides and tripeptides have been extensively studied and an understanding in their fragmentation behaviour is now known, the next step of the investigation into deprotonated peptide analysis is to analyse the fragmentations of tetrapeptides. Kulik and Heerma,[355] and Tsunematsu *et al*,[356] have commenced investigations into this area; the latter studied tetrapeptides derivatised at both termini; the former investigated

underivatised tetrapeptides, but mechanisms justifying the fragmentation behaviour were not reported. Marzluff *et al* have investigated the low energy dissociations of dipeptides, tripeptides and a few tetrapeptides.[375]

The objectives of this research are to:

(i) classify the negative ion fragmentation behaviour of deprotonated underivatised tetrapeptides;

(ii) observe if backbone cleavages provide full sequencing information for the tetrapeptides;

(iii) determine if the fragmentations of the amino acid α-side chains are still observed and if so, what is the relative yield of these ions in comparison to those formed by backbone cleavages?;

(iv) determine whether negative-ion mass spectrometry can be used to determine the structures of complex peptides.

The fragmentation mechanisms presented in this chapter have been rationalised by 'anionic' processes based on the fragmentation of even electron anions of simple organic compounds.[338] In certain cases, other reaction types may occur, such as remote charge processes.

5.10 Fragmentations of Tetrapeptides

The negative-ion CA MS/MS fragmentations of 24 tetrapeptides which contain the majority of common amino acids have been investigated and are reported in this chapter. To clarify the proposed mechanisms of fragmentations, the acidic hydrogens in all tetrapeptides have been exchanged with deuterium and the CA MS/MS spectra of the $[M - D]^-$ ion recorded. The dedeuterated spectra of these tetrapeptides give fragmentations which:

(i) are in accord with the proposed mechanisms discussed in this chapter for the respective unlabelled tetrapeptide;
(ii) indicate that predominant dedeuteration occurs at the carboxyl centre;
(iii) confirm that hydrogen and deuterium transfer either precedes or accompanies cleavage. This is a feature observed in the spectra of the corresponding labelled ions derived from di- and tripeptides.[370]

The extent of hydrogen scrambling observed complicates the spectra of the dedeuterated tetrapeptides, and for brevity, they have been omitted from the Tables. Four examples of dedeuterated spectra have been included to illustrate this point. It must be emphasised that, in essence, the fragmentations observed for the deprotonated tetrapeptides are in accord with the fragmentations of the dedeuterated analogues.

The 24 tetrapeptides studied are grouped according to their side chain functionalities, *viz*:

(i) tetrapeptides containing non-polar or no side chains;
(ii) tetrapeptides containing amino functionality;
(iii) tetrapeptides with aromatic or heterocyclic functionality;
(iv) tetrapeptides containing acid or amide side chains;
(v) tetrapeptides with side chains containing oxygen or sulphur;

5.11 Tetrapeptides Containing Non-Polar or no Side Chains

Four of the tetrapeptides studied are classified in this group, which includes the simplest tetrapeptide, tetra-glycine. Any significant cleavages observed for tetra-glycine can only involve fragmentation of the tetrapeptide backbone. The negative-ion CA MS/MS data are listed in Table **5.2** and an example spectrum of tetra-glycine is shown as Figure **5.1**.

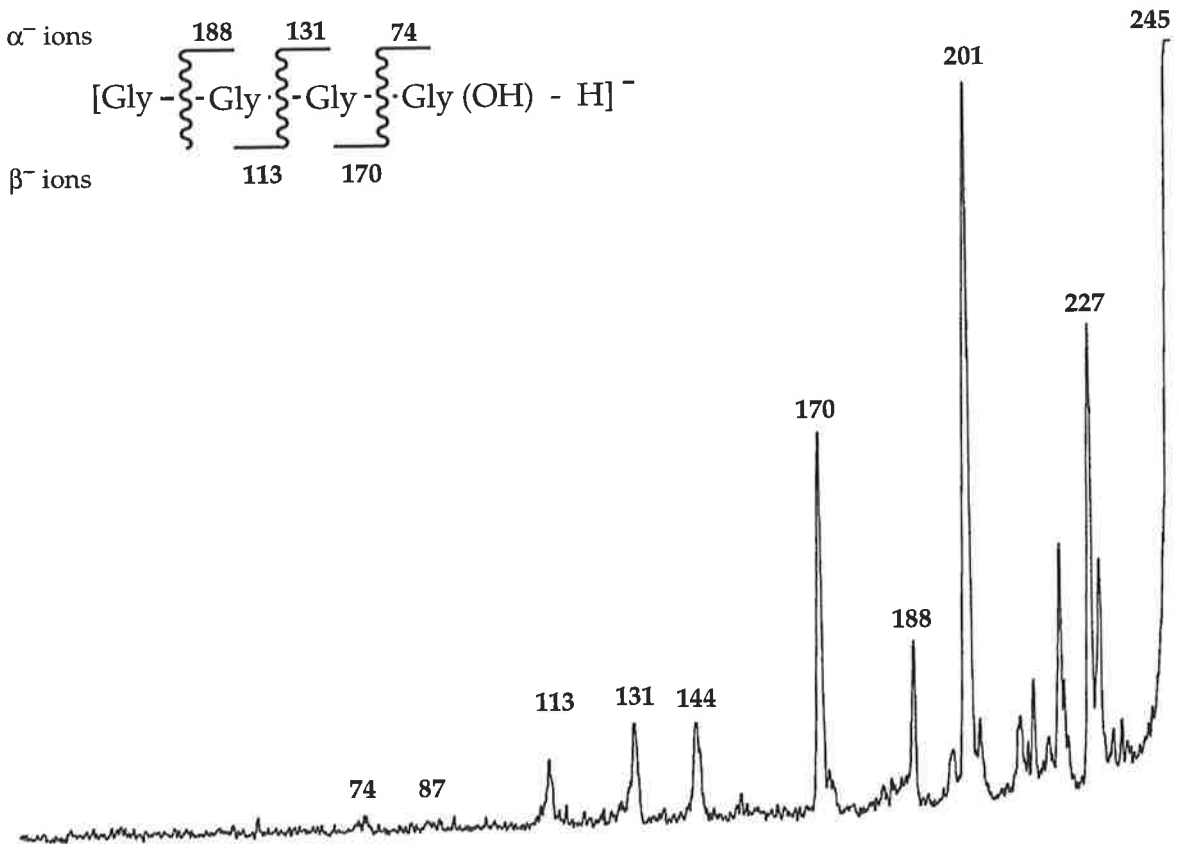


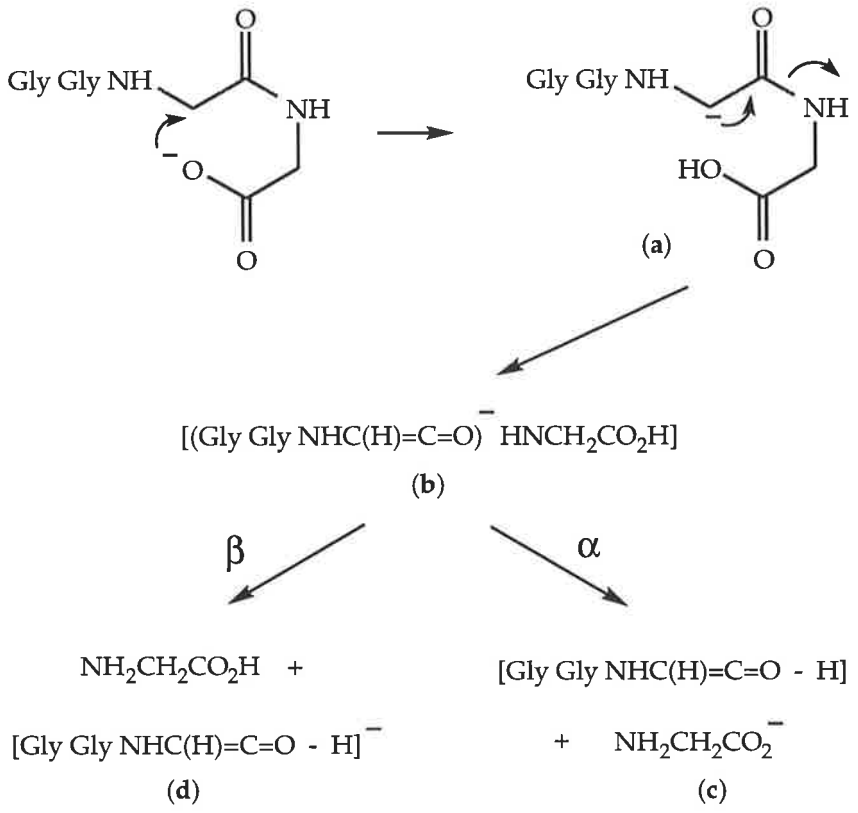
Figure **5.1**: CA MS/MS spectrum of [Gly Gly Gly Gly(OH) – H]⁻, with the α and β ions displayed. Peaks either side of *m/z* 227 are considered to be artefacts (1st field free region decompositions)

Table 5.2: CA MS/MS Data for $[M - H]^-$ Ions of Tetrapeptides Containing Gly (G), Ala (A), Leu (L), Pro (P), Arg (R) and Lys (K). [a]

Precursor Ion	Loss				Formation						
(*m/z*)	H_2O	CO_2	Me	Bu	α_1^-	$(\alpha_1^- - CO_2)$	α_2^-	$(\alpha_2^- - CO_2)$	α_3^-	β_2^-	β_3^-
$[GGGG(OH) - H]^-$	70	100			21	13	15	4	4	9	48
(245)	(227)	(201)			(188)	(144)	(131)	(87)	(74)	(113)	(170)
$[AAAA(OH) - H]^-$	22	100	58		43	58	63	26	59	26	38
(301) [b]	(283)	(257)	(286)		(230)	(186)	(159)	(115)	(88)	(141)	(212)
$[ALAL(OH) - H]^-$	5	100		6	5	21	39	11	50 [d]	10	24
(385) [c]	(367)	(341)		(328)	(314)	(270)	(201)	(157)	(130)	(183)	(254)
$[GPGG(OH) - H]^-$	13	100			36	9	8		5		68
(285) [e]	(267)	(241)			(228)	(184)	(131)		(74)		(210)
$[GPRP(OH) - H]^-$	31	51			5	10	11		4	3	100
(424) [f]	(406)	(380)			(367)	(323)	(270)		(114)	(153)	(309)
$[RPKP(OH) - H]^-$	22	38							5	9	100
(495)	(477)	(451)							(114)	(252)	(380)
$[PFGK(OH) - H]^-$	63	79			4		24		16	12	19
(446) [g]	(428)	(402)			(349)		(202)		(145)	(244)	(300)

(a) Relative abundance (base peak = 100%). Losses of $H^\cdot$ and H_2 do occur, but since they are irrelevent to sequence information they have been omitted.
(b) The genesis of a peak at *m/z* 241 (73%) is unknown.
(c) There are peaks at *m/z* 297 (7%) and 166 (14%) which are not identified.
(d) The α_3^- ion loses CO_2 to form a peak at *m/z* 86 (4%).
(e) There is a peak at *m/z* 171 (7%) whose genesis is unknown.
(f) The genesis of (238, 11%) and (281, 17%) are not known.
(g) This spectrum also shows peaks at: *m/z* 257, $(\alpha_1^- - PhMe)$, 11%; *m/z* 113, $(Pro\ NH^-)$, 4%; *m/z* 332, [loss of $Pro(NH_2)$], 13%; *m/z* 186, {$PhCH{=}CHCONHC{\equiv}CO^-$ [*i.e.* *m/z* 332 - Lys(OH)]}, 26%; *m/z* 288, {*m/z* 332 - CO_2 [*i.e.* loss of $Pro(NH_2) + CO_2$]}, 100%.

The backbone fragmentations observed for tetra-glycine are typical of the fragmentations observed for tetrapeptides. The spectrum shows the expected losses of H_2O and CO_2, (which forms the base peak in the majority of cases), and include some typical features as well as cleavage ions which give sequence information. The sequence information is given by backbone cleavages which are very dramatic: the normal cleavages of dipeptides and tripeptides for the type shown in schemes (**5.13**) and (**5.14**) are very pronounced, with a second and related cleavage operating which is not observed in the spectra of dipeptides and only occasionally in tripeptides. The related cleavage ion is formed as proposed in scheme (**5.18**) Proton transfer to the carboxylate centre forms enolate anion **a** which may cleave to form the ion complex **b**. The complex may (i) cleave to give deprotonated glycine **c**, or (ii) neutral glycine **d** may be eliminated. The observation of these product ions arising from a competitive proton transfer reaction implies the possibility of formation of a long-lived dissociation complex.



(**5.18**)

Since the two processes forming the anions **c** and **d** are widely observed in the spectra of tetrapeptides, they are named α and β backbone cleavages, so as to avoid confusion with any of the documented cleavages observed in the positive ion mode.[329] In principle, three enolate anions may undergo α and β cleavage in the case of tetrapeptides. Since each pair arises from the same ion complex, the α and β cleavages are numbered from the N-terminal end of the peptide. The backbone cleavages observed in Figure **5.1** are summarised in Figure **5.2**.

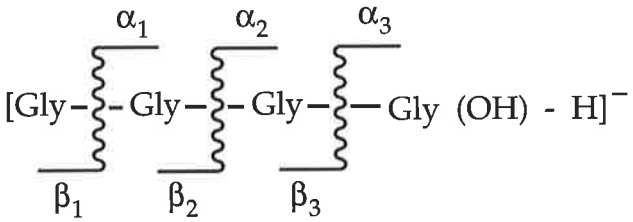


Figure **5.2**: Schematic representation of the backbone cleavages of tetrapeptides.*

It must be emphasised that this pictorial representation is not intended to depict the actual mechanisms of the reactions, since the α ions almost certainly undergo proton transfer to form the product anions, while the β ions are formed following deprotonation in the anion complex [see scheme (**5.18**)].

The three other tetrapeptides studied which only have polar side chains also behave similarly, with significant fragmentations generating sequence information. The data from these spectra are also listed in Table **5.2**, and examples of these spectra are shown in Figures **5.3**, **5.4** and **5.5**.

* The β_1 anion has been included here to complete the set of α and β cleavages. However, it must be noted that tetra-glycine does not form a β_1 ion.

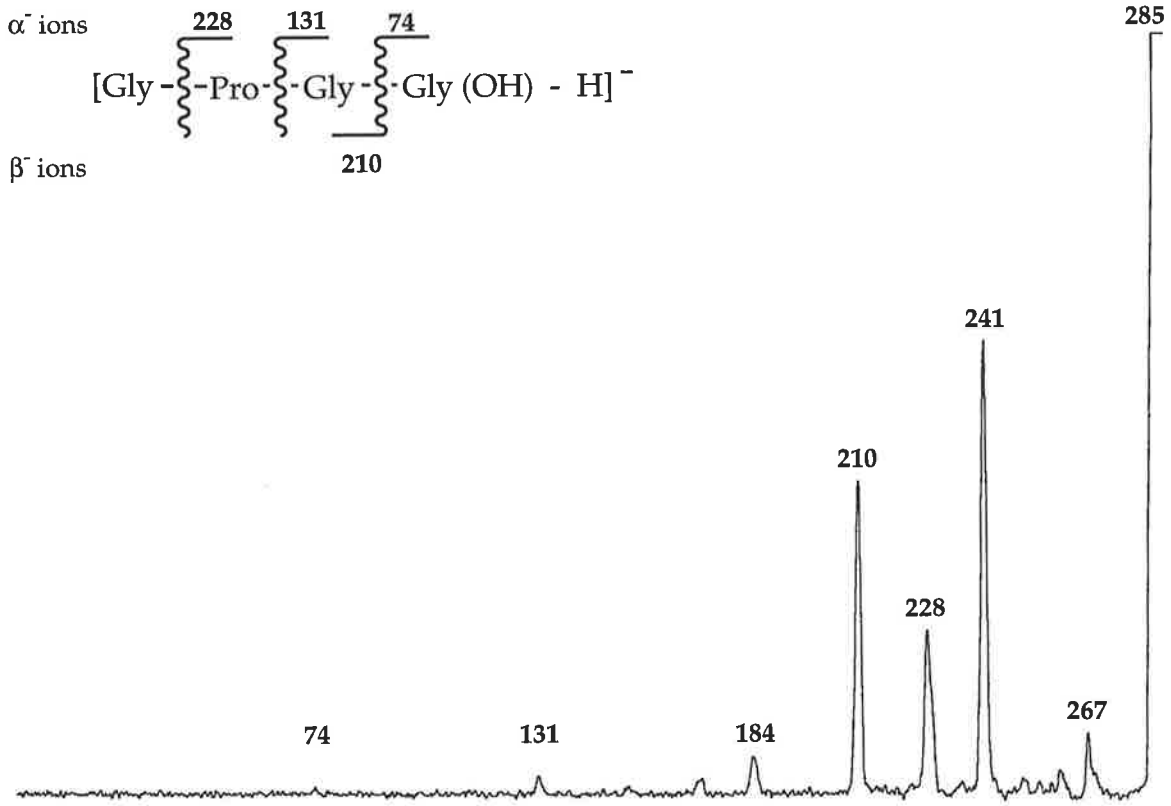


Figure **5.3**: CA MS/MS spectrum of [Gly Pro Gly Gly(OH) – H]⁻.

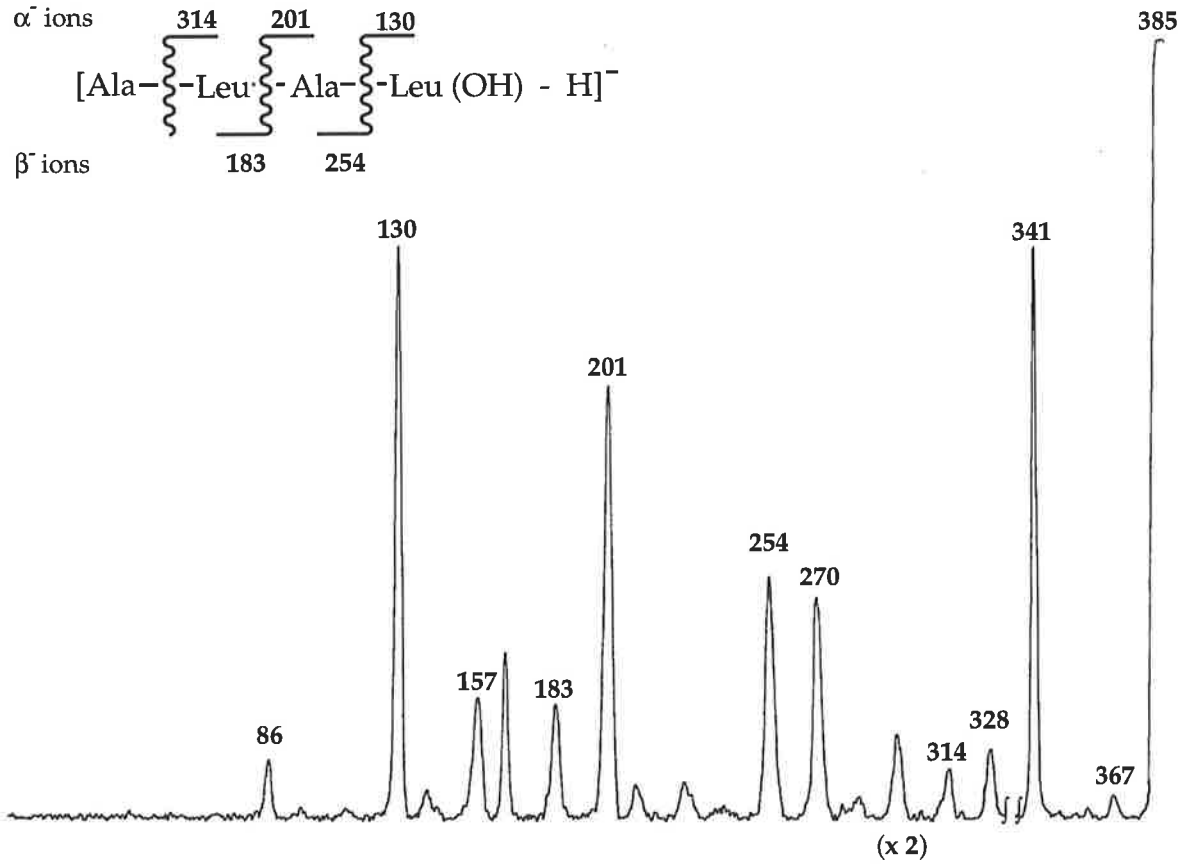


Figure **5.4**: CA MS/MS spectrum of [Ala Leu Ala Leu(OH) – H]⁻.

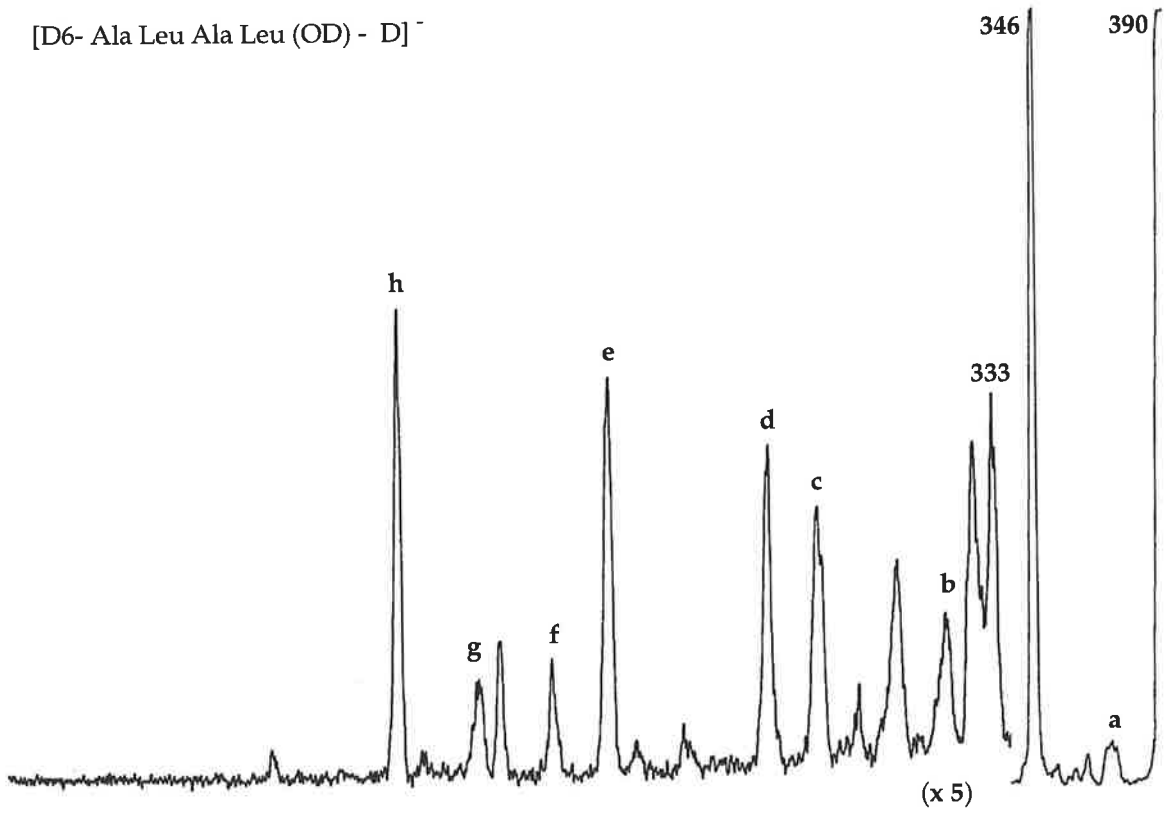


Figure **5.5**: CA MS/MS spectrum of $[D_6$- Ala Leu Ala Leu(OD) – D$]^-$. The spectrum is characterised as follows (the majority of peaks are not entirely resolved) [*m/z* (loss or formation) relative abundance]: [**a**] 372/371 (H_2O/HOD) 7; 346 (CO_2) 100; 333 ($Bu^{\cdot}$) 9; [**b**] 317/318 (α_1^-, D_3, D_4) 4; [**c**] 273/274 [$(\alpha_1 - CO_2)^-$, $D_3>D_4$] 7; [**d**] 257/258 (β_3^-, $D_3<D_4$) 8; [**e**] 203/204 (α_2^-, $D_2<D_3$) 10; [**f**] 185/186 (β_2^-, D_2, D_3) 4; [**g**] 159/160 [$(\alpha_2 - CO_2)^-$, D_2, D_3] 3; and [**h**] 131/132 (α_3^-, D_1, D_2) 12.

5.12 Tetrapeptides Containing Amino Functionality

Five of the tetrapeptides investigated contain amino functional groups. The CA MS/MS spectral data are recorded in Table **5.2**. Representative spectra are shown as Figures **5.6** and **5.7**.

It has been established that dipeptides containing lysine do not have fragmentations associated with the side-chain,[367] and arginine dipeptides show characteristic loss of HN=C=NH in the spectra (see Table **5.1**).[367]

Surprisingly, no such fragmentation corresponding to this characteristic loss is observed for the tetrapeptide examples. The tetrapeptide spectra show that, in the negative-ion mode, Arg and Lys behave in a similar manner to the amino acids Ala, Val and Leu in undergoing α and β cleavage to such an extent that backbone cleavage occurs to the exclusion of α-side chain fragmentation.

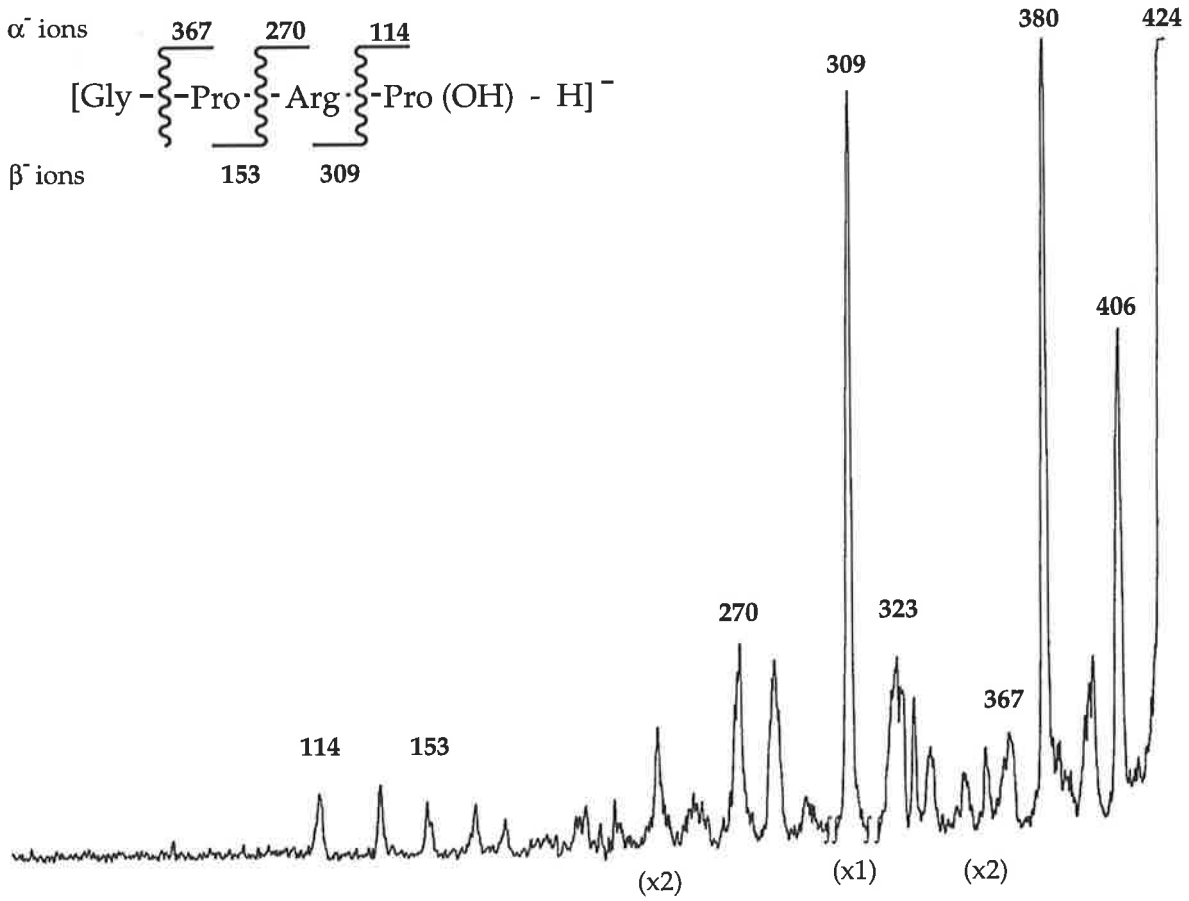


Figure **5.6**: CA MS/MS spectrum of [Gly Pro Arg Pro(OH) - H]⁻.

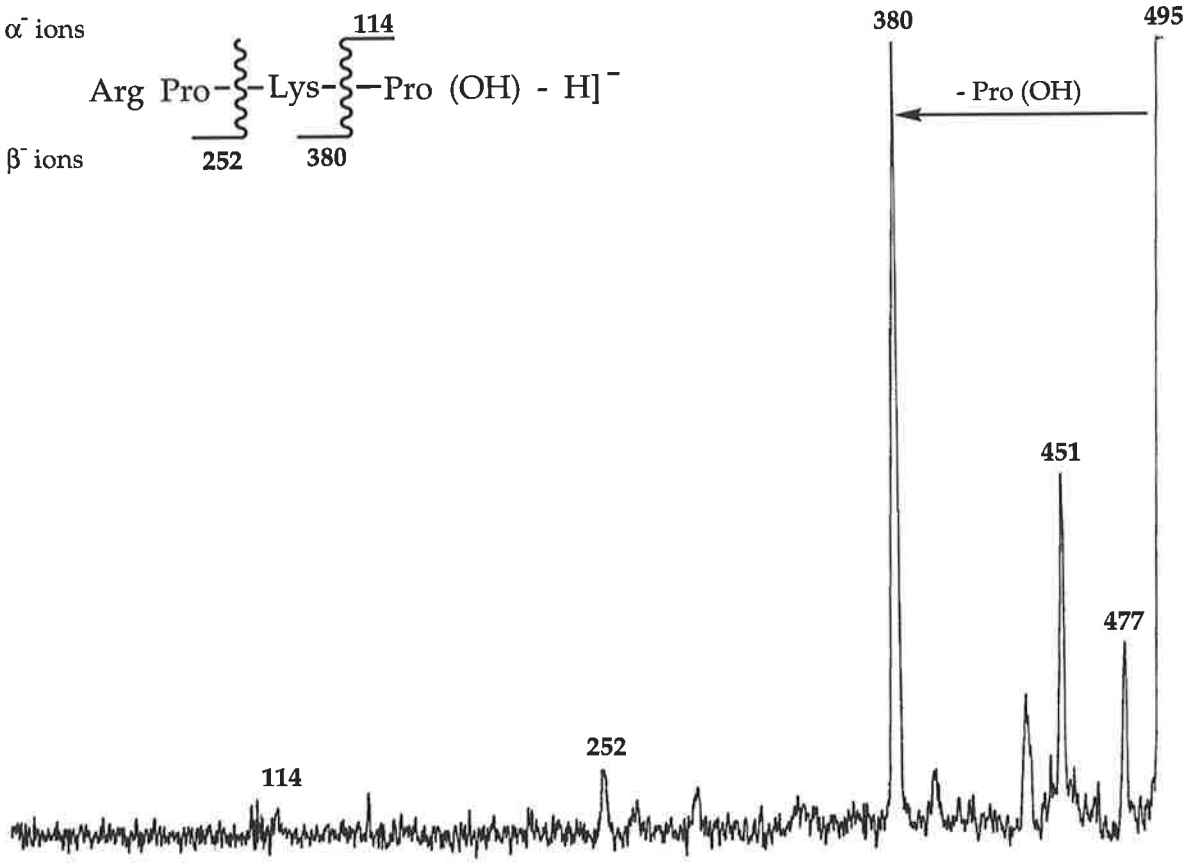


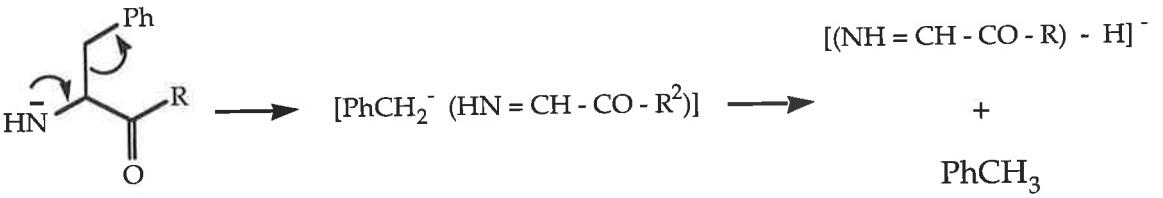
Figure 5.7: CA MS/MS spectrum of [Arg Pro Lys Pro(OH) - H]$^-$. Loss of C-terminal proline dominates spectra. Arginine and lysine show no side-chain fragmentations.

5.13 Tetrapeptides with Aromatic or Heterocyclic Functionality

Eleven of the tetrapeptides studied are classified into this group which include the amino acids Phe, Tyr, Trp and His. The negative ion spectra are recorded in Tables **5.3** and **5.4**, and representative spectra are depicted in Figures **5.8** - **5.15**. The α and β cleavages are prominent fragmentations, which allows for the complete sequence of each peptide to be determined. Also there are a number of fragmentations characteristic of aromatic and heterocyclic amino acid residues in tetrapeptides. The extent of these side chain fragmentations diminish considerably upon lengthening of the peptide, albeit they are still noticeable. These side chain fragmentations are discussed in the following pages for each of the individual amino acids.

(i) Phenylalanine

Dipeptides containing phenylalanine show a side chain cleavage which forms $PhCH_2^-$.[365] This fragmentation pathway still occurs in the spectra of tetrapeptides, however the $PhCH_2^-$ ion formed in the ion complex deprotonates the neutral, and consequently the loss of toluene is detected. This process is characteristic of tetrapeptides containing N-terminal phenylalanine, and is illustrated in scheme (**5.19**) and in Figures **5.8** and **5.9**. A spectrum of a deuterated tetrapeptide is shown in Figure **5.10**.



(**5.19**)

Table 5.3: CA MS/MS Data for $[M - H]^-$ Ions of Tetrapeptides Containing Phe (F). [a]

Precursor Ion	Loss						Formation						
(*m/z*)	NH_3	H_2O	CO_2	PhMe	$PhCH{=}CHCO_2H$	$PhCH{=}CHCO_2^-$	α_1^-	$(\alpha_1^- - CO_2)$	α_2^-	$(\alpha_2^- - CO_2)$	α_3^-	β_2^-	β_3^-
$[GGFL(OH) - H]^-$		26	100				23	3	34 [b]	4	24 [c]	5	56 [c]
(391)		(373)	(347)				(334)	(290)	(277)	(233)	(130)	(113)	(260)
$[VAAF(OH) - H]^-$		9	100		14	2	3 [e]	5	12	3	17	5	9
(405) [d]		(387)	(361)		(257)	(147)	(306)	(262)	(235)	(191)	(164)	(169)	(240)
$[FGGF(OH) - H]^-$	7 [g]		100 [h]	13	[i]	4	36	8	21	3	29	14	21
(425) [f]	(408)		(381)	(333)		(147)	(278)	(234)	(221)	(177)	(164)	(203)	(260)
$[FGFG(OH) - H]^-$	15 [g]		100 [k]	21			6	3	36 [bl]	4	4	8 [l]	45
(425) [j]	(408)		(381)	(333)			(278)	(234)	(221)	(177)	(74)	(203)	(350)

(a) Relative abundance (base peak = 100%). Losses of $H^{\cdot}$ and H_2 do occur, but since they are irrelevent to sequence information they have been omitted.

(b) The α_2^- ion loses PhMe to form *m/z* 185 (25%) for GGFL (OH) and *m/z* 129 (3%) for FGFG (OH).

(c) The α_3^- and β_3^- ions have the same integral mass as the ions formed by deprotonation at the β-carbon.

(d) This spectrum shows peaks at *m/z* 214 (3%), 186 (2%) and 115 (2%) for the respective formations of {$[H_2NCH(iPr)CONHCH(Me)CONHCH_2CH_3 - H]^-$, *i.e.* *m/z* 257 - HNCO}, {$[H_2NCH(iPr)CONHCH(Me)CONH]^-$, *i.e.* *m/z* 257 - Ala} and {$[H_2NCH(iPr)CONH]^-$, *i.e.* *m/z* 257 - (Ala Ala)}. The formation of these ions have no analogy in this or previous work. The mechanism of formation is not understood.

(e) The α_1^- ion loses H_2O to form *m/z* 288 (4%).

(f) There is a peak at *m/z* 322 (13%) which is unidentified.

(g) These spectra contain unresolved peaks corresponding to the losses of H_2O and NH_3.

(h) There is a peak at *m/z* 289, (7%) corresponding to the loss of (PhMe + CO_2); also a peak at *m/z* 364, (NH_3 + CO_2), 17%.

(i) Loss of $PhCH{=}CHCO_2H$ (*m/z* 277), unresolved from the α_1^- ion.

(j) Unidentified peaks include: (322, 12%), (306, 12%), (185, 21%) and (163, 17%).

(k) There are peaks corresponding to: *m/z* 289, (PhMe + CO_2), 5%; *m/z* 364, (NH_3 + CO_2), 17%.

(l) The anions formed from the β-carbon (*m/z* = 220 and 204) are unresolved from the α_2^- and β_2^- ions.

Table 5.4: CA MS/MS Data for $[M - H]^-$ Ions of Tetrapeptides Containing Phe (F), Tyr (Y), His (H) and Trp (W). [a]

Precursor Ion	Loss					Formation						
(*m/z*)	H_2O	CO_2	$CH_2{=}C_6H_4{=}O$	$(CO_2 + CH_2{=}C_6H_4{=}O)$	$RCH{=}CHCO_2H$	α_1^-	$(\alpha_1^- - CO_2)$	α_2^-	$(\alpha_2^- - CO_2)$	α_3^-	β_2^-	β_3^-
$[YGGF(OH) - H]^-$	22[c]	100	33	19	12	4[d]	7	4[e]	4	8	4	18
(441) [b]	(423)	(397)	(335)	(291)	(294)	(278)	(234)	(221)	(177)	(164)	(219)	(276)
$[GGYR(OH) - H]^-$	24 [c]	100	12	10		23	4	10		17	2	13
(450) [f]	(432)	(406)	(344)	(300)		(393)	(349)	(336)		(173)	(113)	(276)
$[GAVH(OH) - H]^-$	32	100			16	6	12	15	13	53		
(381) [g]	(363)	(337)			(243)	(324)	(280)	(253)	(209)	(154)		
$[FPWL(OH) - H]^-$	100	6				25		19				
(560) [h]	(542)	(516)				(413)		(316)				
$[FPWP(OH) - H]^-$	32	15				19		92				100
(544) [i]	(526)	(500)				(397)		(300)				(429)

(a) Relative abundance (base peak = 100%). Losses of $H^\cdot$ and H_2 do occur, but since they are irrelevent to sequence information they have been omitted.

(b) This tetrapeptides also loses $(H_2O + CO_2)$, (379, 6%).

(c) The loss of H_2O is unresolved due to the loss of NH_3 from this tetrapeptide.

(d) Unresolved, shoulder on *m/z* 276. (e) Unresolved.

(f) This spectrum also shows *m/z* 130, Gly Gly NH^-, 4% and 319, $\{HOC_6H_4CH{=}CHCONHCH[(CH_2)_3NHC(NH_2){=}NH]CO_2H - H\}^-$, 3%; There is a peak at *m/z* 201 (7%) which is unidentified.

(g) This spectrum shows two peaks at *m/z* 217 (34%) and 234 (22%) which are unidentified.

(f) The characteristic loss of C_9H_7N results in the peak at *m/z* 431, which is followed by the β_3^- cleavage, forming a peak at *m/z* 300.

(g) This scpectrum also shows peaks at: *m/z* 415, (loss of C_9H_7N), 84%; *m/z* 371, (loss of $C_9H_7N + CO_2$), 57; *m/z* 300, (β_3^- ion from *m/z* 431, *i.e.* formation of $[(Phe\ Pro\ NHCH{=}C{=}O) - H]^-$.

When phenylalanine is an internal residue of a tetrapeptide, loss of toluene is observed, not from the molecular anion, but from the α-cleavage which generates N-terminal phenylalanine as the fragment anion. For example, (i) loss of toluene is observed from the α_1 ion of deprotonated Pro Phe Gly Lys(OH) (Table **5.2**), and (ii) the α_2 cleavage of [Gly Gly Phe Leu(OH) - H]$^-$ is also followed by loss of toluene (Table **5.3**).

When phenylalanine is located at the C-terminal position, loss of toluene is never observed. In such cases a different side chain fragmentation may occur as detailed in the following section.

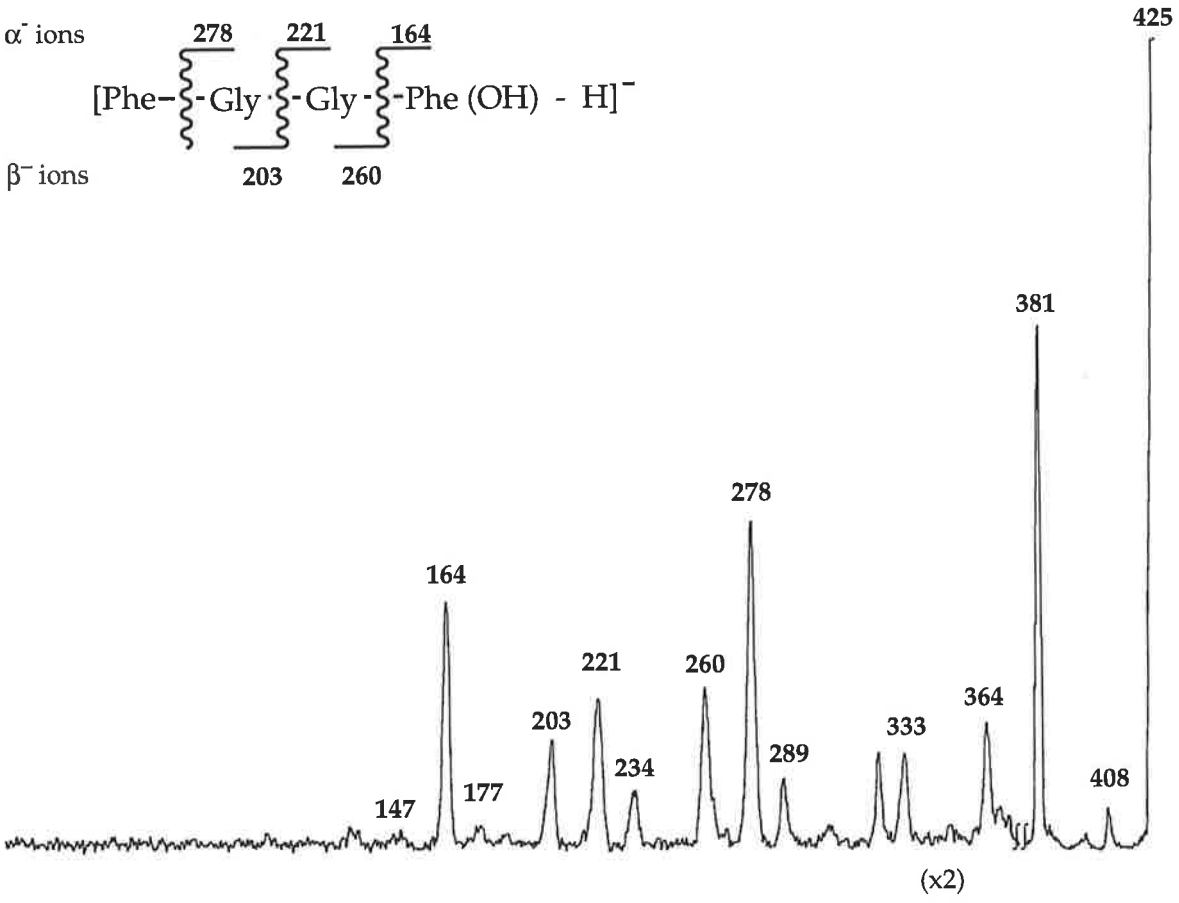


Figure **5.8**: CA MS/MS spectrum of [Phe Gly Gly Phe(OH) - H]$^-$.

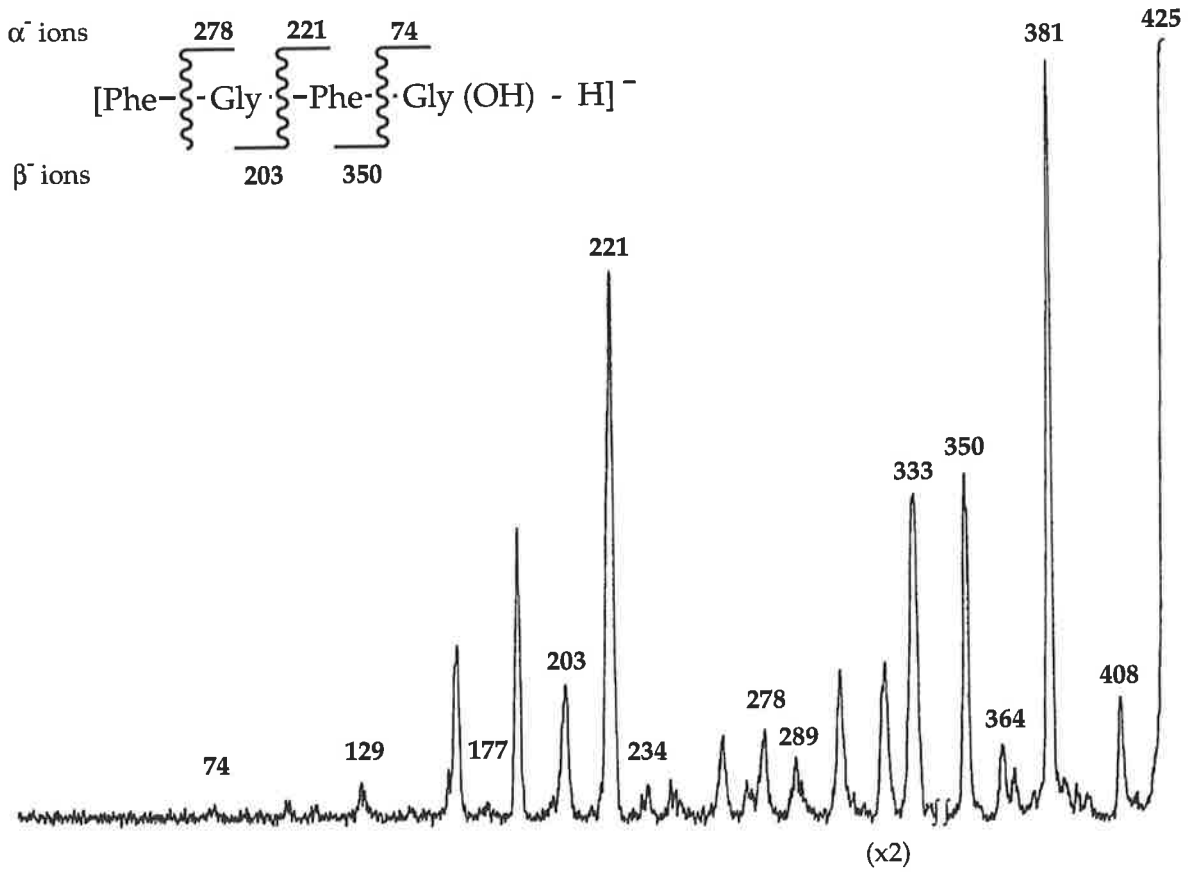


Figure **5.9**: CA MS/MS spectrum of [Phe Gly Phe Gly(OH) - H]⁻. (Table **5.3**)

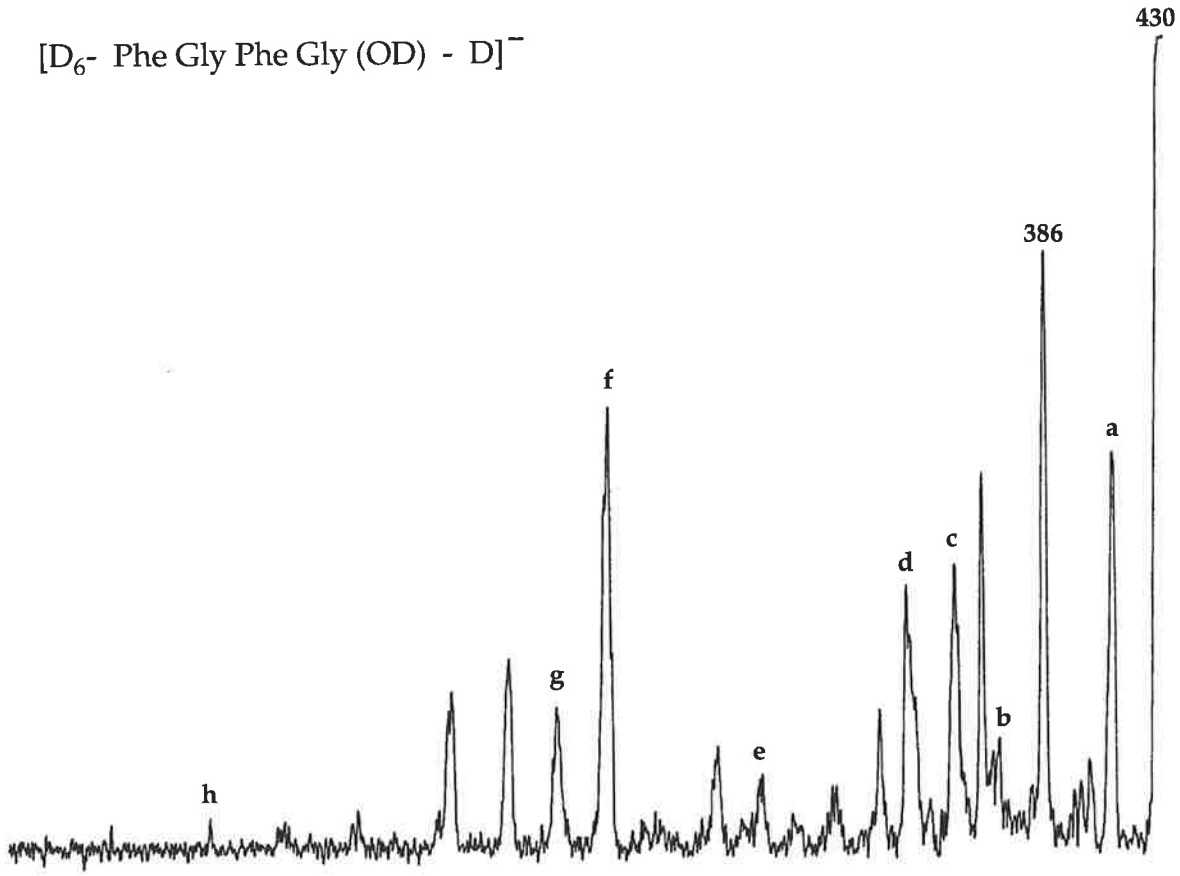
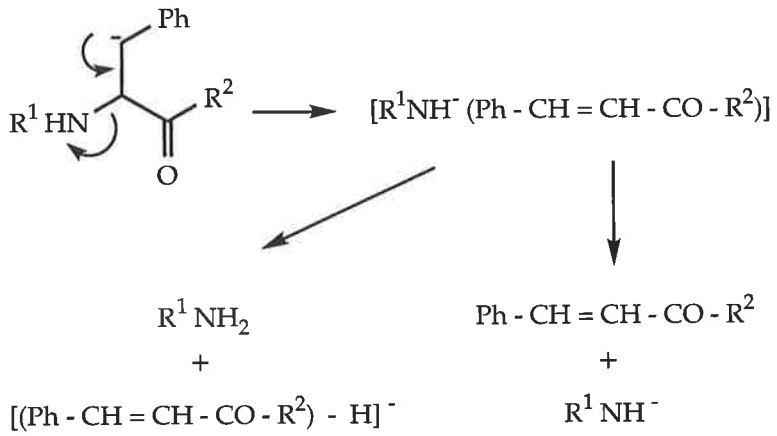


Figure **5.10**: CA MS/MS spectrum of [D_6- Phe Gly Phe Gly(OD) - D]$^-$. The spectrum is characterised as follows (the majority of peaks are not fully resolved) [*m/z* (loss or formation) relative abundance]: **[a]** 410/411 (ND_3, HOD, D_2O) 66; 386 (CO_2) 100; **[b]** 366/367 (410/411 - CO_2) 18; **[c]** 353/354 (β_3^-, D_3>D_4) 47; **[d]** 337/338 ($PhCH_2D$>$PhCH_3$) 44; **[e]** 280/281 (α_1^-, D_3>D_2) 12; **[f]** 222/223 (α_2^-, D_2>D_1) 74; **[g]** 205/206 (β_2^-, D_3>D_2) 24; **[h]** 75 (α_3^-) 4. This spectrum also shows unidentified peaks at: *m/z* 362 (63%); 325 (23); 185/186 (32); 163/164 (27). Peaks associated with fragmentation from the β-carbon are unresolved.

(ii) Fragmentations Resulting from the Formation of a Benzyl Anion in Tetrapeptides Containing Phenylalanine

Tetrapeptides containing phenylalanine residues have fragmentations resulting from deprotonation at the β-carbon, resulting in the formation of an ion-complex as described by scheme (**5.20**).



where the precursor ion is a tetrapeptide, such that:

R^1 = N-terminal residues, or H

R^2 = C-terminal residues, or OH (**5.20**)

The ion-complex fragments *via* two ways, either (i) dissociation, which generates the loss of the neutral cinnamyl derivatives, or (ii) fragmentation after proton transfer within the ion-complex forming cinnamate anion derivatives. Such a process may provide sequencing information complementary to the α and β cleavage ions as it identifies the position at which the process occurs. As will be seen shortly, it doesn't necessarily indicate

which particular residue is lost, since other amino acids which are capable of being deprotonated at the β-carbon also fragment *via* this pathway. Fragmentations derived from deprotonation at the β-carbon are observed in the spectra of dipeptides and tripeptides, however they become less pronounced as the length of the peptide increases. It has been suggested that this process may proceed *via* a charge remote mechanism [see scheme (**5.12**), page 128], but in the case of N-terminal residues this is not possible.

The following observations are noted for phenylalanine occupying each of the four positions in a tetrapeptide. In principle other aromatic or heterocyclic amino acid residues may also behave similarly to phenylalanine in these positions. These residues are discussed separately.

(a) When phenylalanine occupies the N-terminal position, fragmentations associated with side chain cleavages occur for deprotonation at the amino functionality, or the benzyl position. The former results in the loss of $PhCH_3$, and occasionally the loss of CO_2 accompanies such a process (*e.g.* see Figure **5.9**). The latter results in the loss of NH_3 and is frequently accompanied by the loss of CO_2. For example, initial deprotonation at the β-carbon of the N-terminal residue of the tetrapeptide Phe Gly Phe Gly(OH), (Figure **5.9**) results in the losses of NH_3 (m/z = 408) and $NH_3 + CO_2$ (m/z = 364).*

(b) Pro Phe Gly Lys(OH) (Figure **5.11**, Table **5.2**) is the only tetrapeptide studied which can be deprotonated at the β-carbon of residue two. This tetrapeptide shows remarkable fragmentations associated with deprotonation at this site. Formation of Pro NH^- (m/z = 113) and loss of Pro (NH_2) (m/z = 332) are rationalised by scheme (**5.21**). The latter anion (deprotonated

* Deprotonation at the β-carbon of the internal phenylalanine residue results in the loss of cinnamylglycine (m/z = 220) and the formation of deprotonated cinnamyl glycine (m/z = 204), and both are unresolved from the α_2^- and β_2^- ions.

cinnamyl-glycyl-lysine) can fragment further either by decarboxylation, forming the base peak ($m/z = 288$), or can lose lysine *via* a β-cleavage ($m/z = 186$). The loss of lysine from this ion was an unexpected observation, as there are no similar fragmentations occurring when the N-terminal residue is aromatic. It is not understood why the loss of CO_2 from this ion should form the base peak. Also, this spectrum shows that the α_1^- ion then loses PhMe since the tripeptide Phe Gly Lys(OH) which is formed, is deprotonated at the N-terminus.

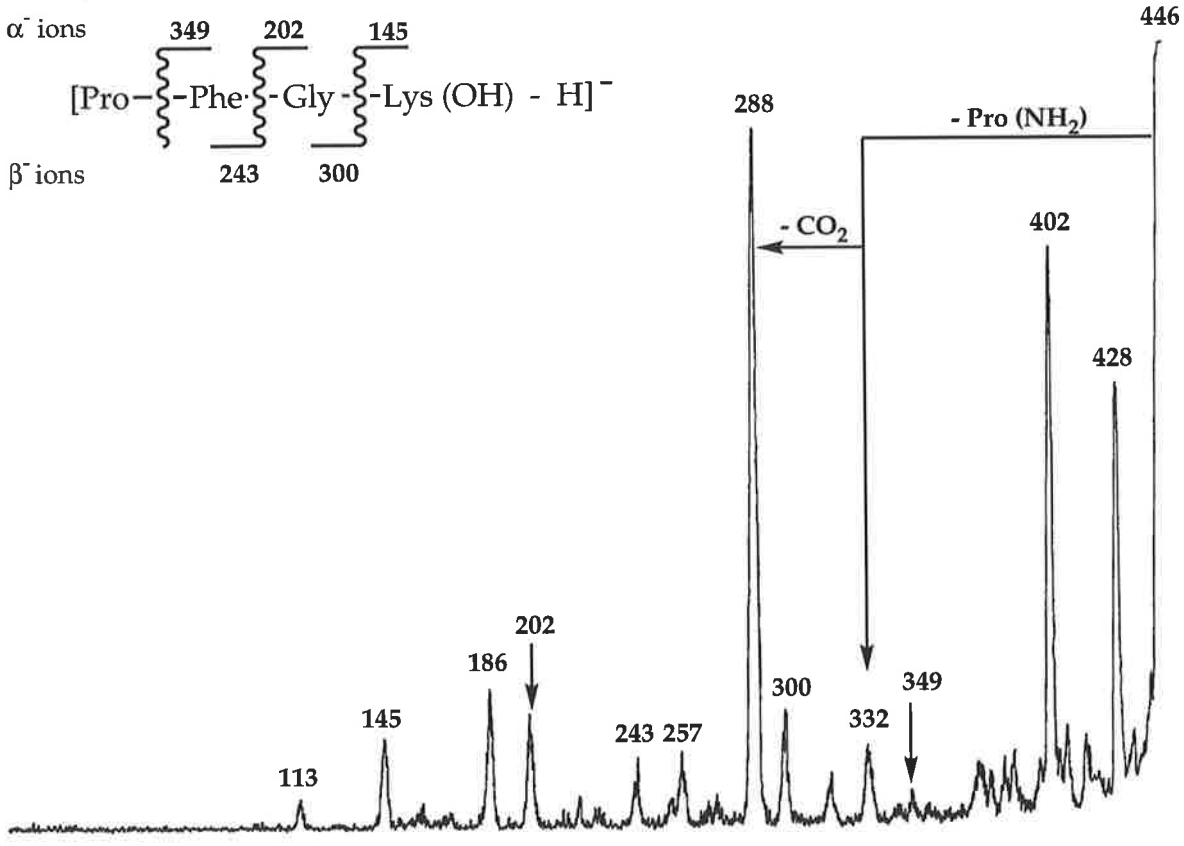
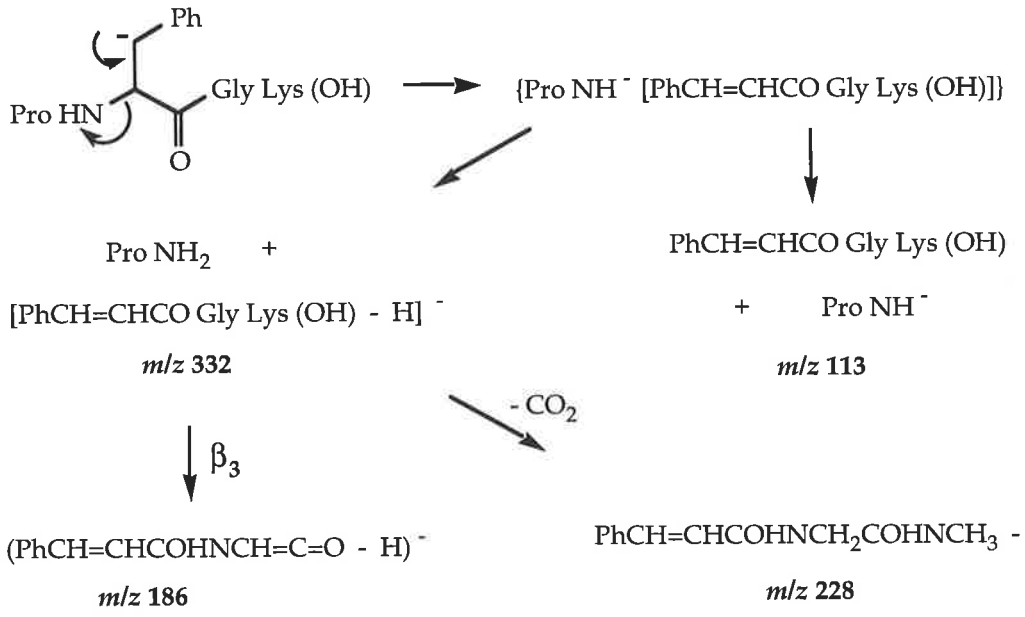
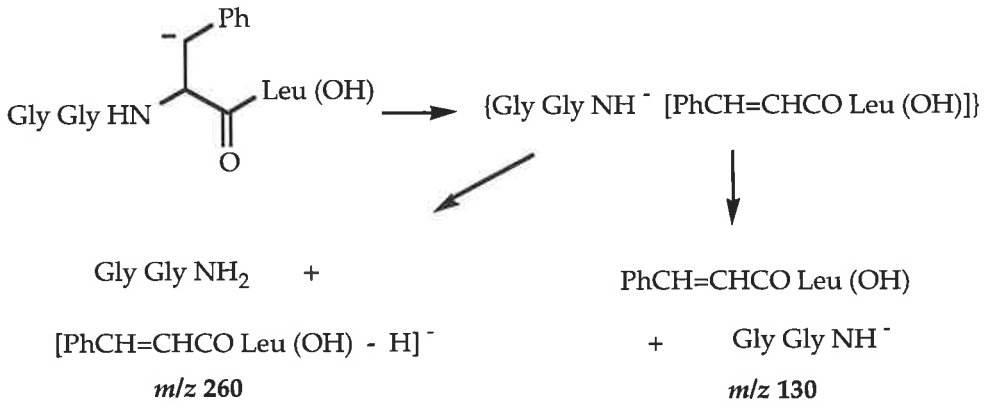


Figure **5.11**: CA MS/MS spectrum of [Pro Phe Gly Lys(OH) - H]⁻.



(5.21)

(c) When position three is occupied by phenylalanine, proton transfer from the β-carbon to the carboxylate centre results in an ion-complex [(scheme (5.22)] which may fragment by two competitive pathways: (i) dissociation yields a deprotonated dipeptide amide (loss of a neutral cinnamyl amino acid); or (ii) proton transfer preceding fragmentation results in the formation of deprotonated cinnamyl amino acid (loss of neutral dipeptide amide). These processes account for the peaks at *m/z* 130 and 260 in the spectra of deprotonated Gly Gly Phe Leu(OH) (Figure **5.12**), and **g** and **d** in the spectra of dedeuterated d6- Gly Gly Phe Leu(OD) (Figure **5.13**).



(5.22)

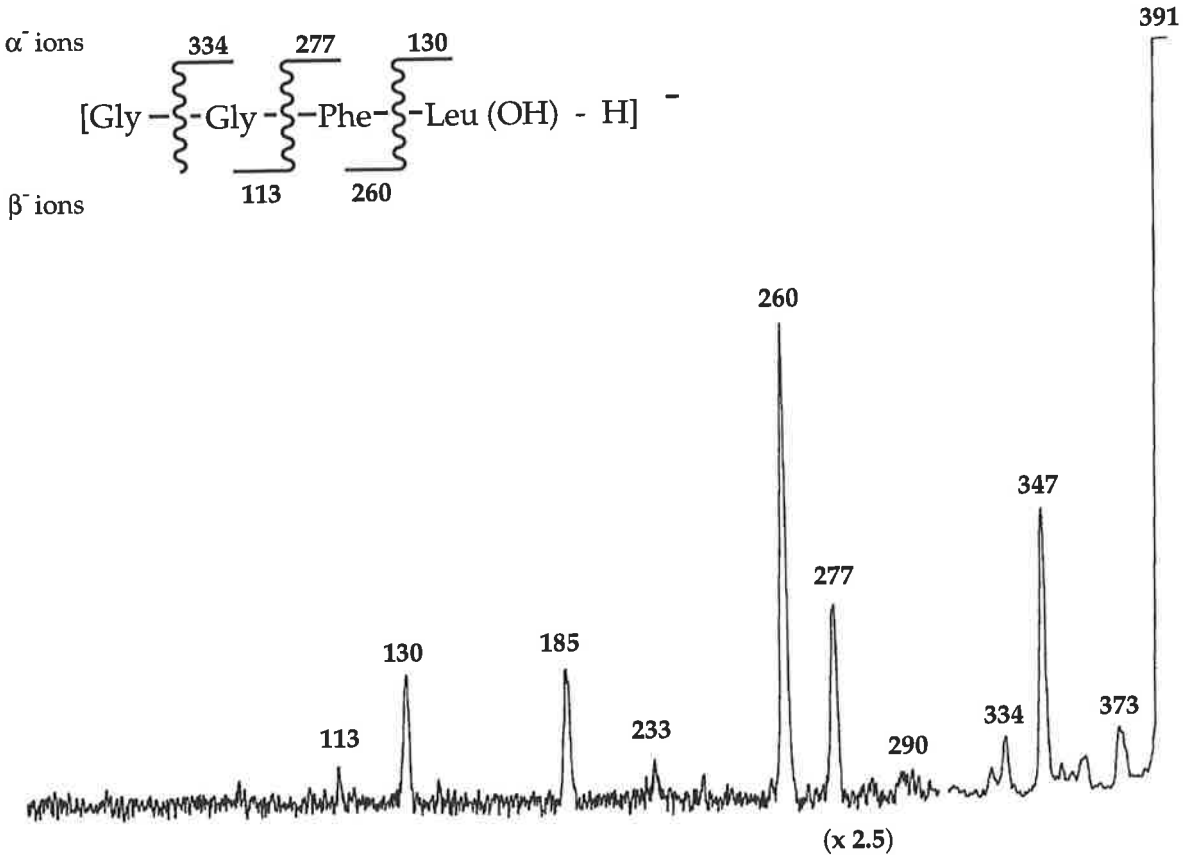


Figure **5.12**: CA MS/MS spectrum of [Gly Gly Phe Leu(OH) - H]⁻.

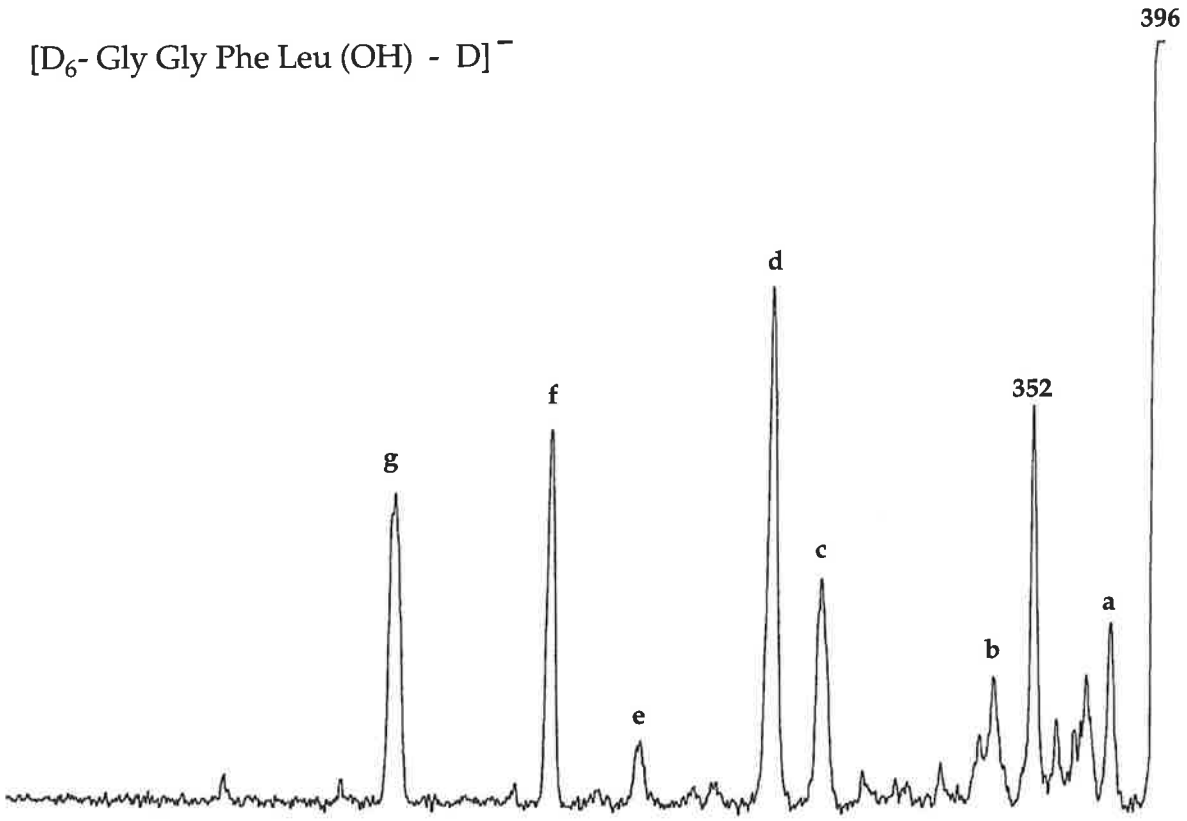


Figure **5.13**: CA MS/MS spectrum of [d_6- Gly Gly Phe Leu(OD) - D]$^-$. The spectrum is characterised as follows (the majority of peaks are not fully resolved) [*m*/*z* (loss or formation) relative abundance]: [**a**] 376/377 (ND_3, HOD, D_2O) 35; 352 (CO_2) 78; [**b**] 337/338 (α_1^-, D_3>D_4) 25; [**c**] 279/280 (α_2^-, D_2>D_3) 47; [**d**] 263/264 (β_3^-, D_3>D_4) 100; [**e**] 218/219 (362/363 - CO_2) 12; [**f**] 187 [(α_2 - PhMe)$^-$] 72; [**g**] 131/132 (α_3^-, D_1, D_2) 60. Cleavages associated with fragmentation from the β-carbon are unresolved from the α_3^- and β_3^- ions, as noticed by thickness of these peaks.

(d) Of the tetrapeptides investigated, Phe and His are the only residues with aromatic or heterocyclic functionality at the C-terminus. These tetrapeptides fragment *via* the β-carbon and show: competitive cinnamate anion formation and loss of cinnamic acid for C-terminal phenylalanine, *e.g.* Figure **5.14** [Val Ala Ala Phe(OH)]; and only the loss of uracanic acid for C-terminal histidine (see Table **5.4**).

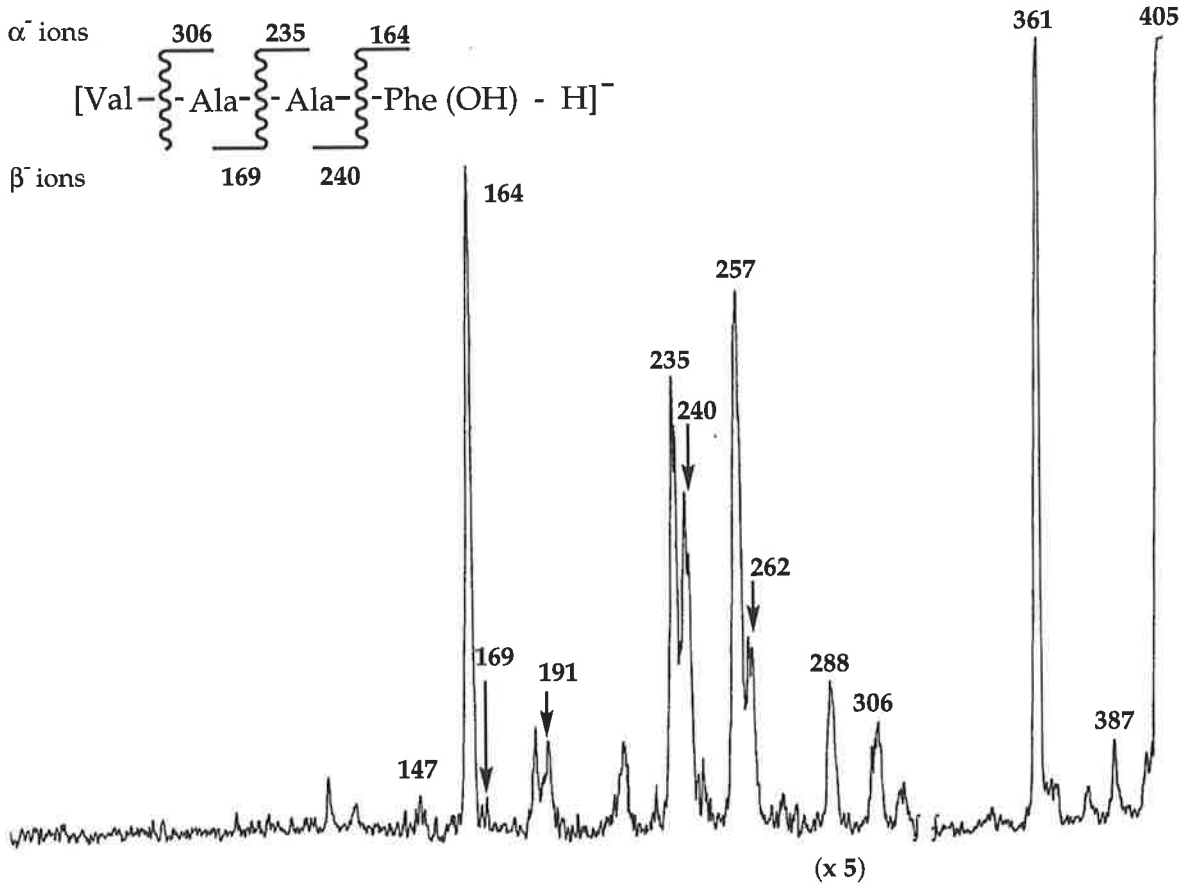
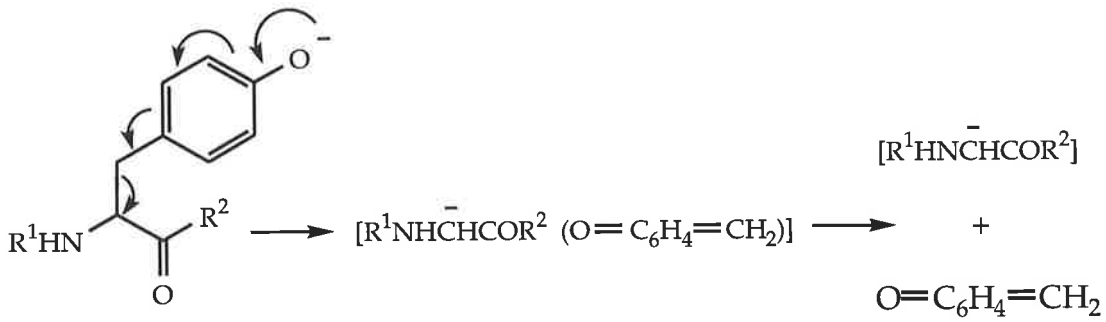


Figure **5.14**: CA MS/MS spectrum of [Val Ala Ala Phe (OH) - H]⁻.

(iii) Tyrosine

Spectra of tetrapeptides containing tyrosine can be readily identified by the loss of *p*-benzoquinone methide [scheme (**5.23**)], which is also a loss observed in the spectra of the amino acid and tyrosine dipeptides;[342] however, the formation of deprotonated *p*-benzoquinone is not detected in any of the tyrosine containing tetrapeptide spectra.

From a sequencing point of view, the loss of the *p*-benzophenone methide indicates the presence of tyrosine, but since the loss is independent of position in the tetrapeptide, it does not disclose any detail regarding its specific position. Such information may be gained from fragmentations observed after deprotonation at the β-carbon. Such a process is analogous to the mechanisms observed for phenylalanine being deprotonated at this position. Such a mechanism can confirm the position of tyrosine (or phenylalanine) in a peptide.



where R^1 = N-terminal residues, or H.

R^2 = C-terminal residues, or OH (**5.23**)

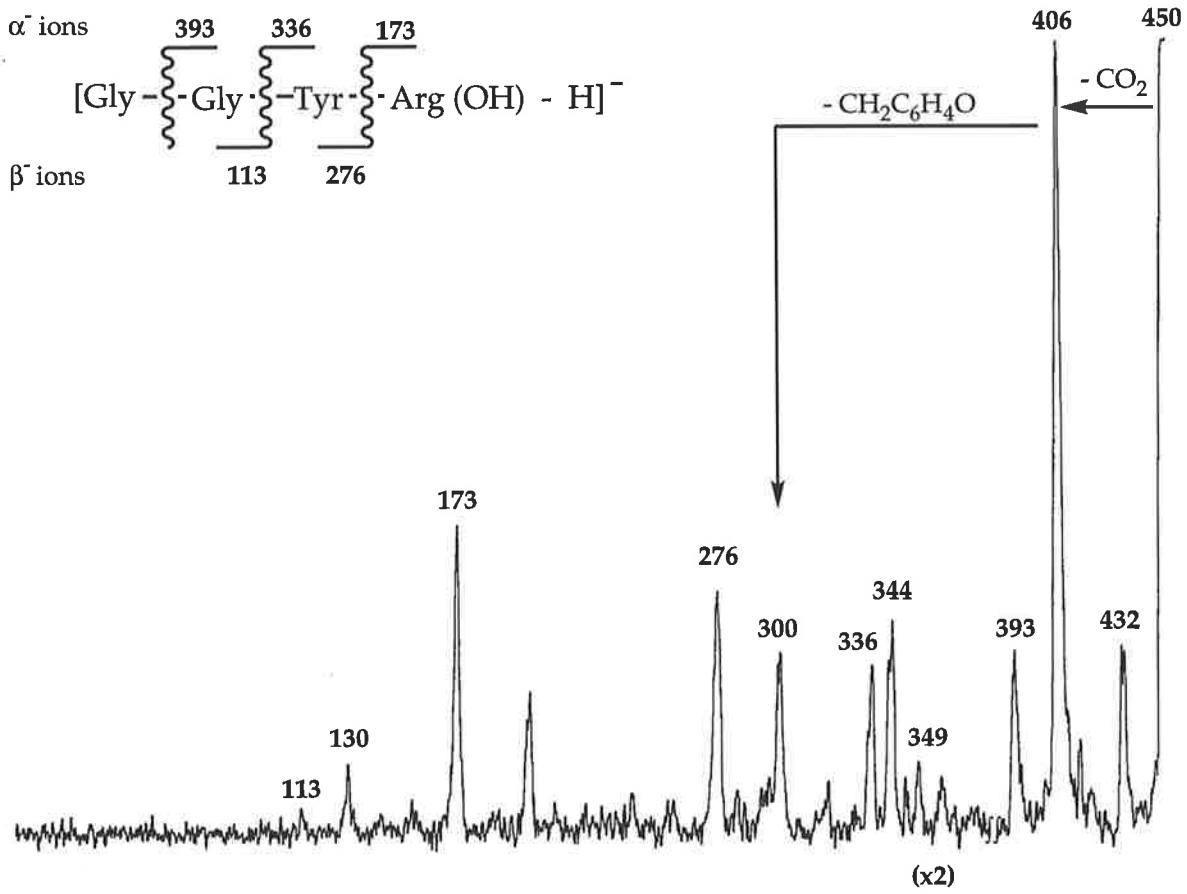


Figure **5.15**: CA MS/MS spectrum of [Gly Gly Tyr Arg (OH) - H]⁻, with the loss of *p*-benzoquinone methide illustrated.

(iv) Tryptophan

Tetrapeptides containing tryptophan show α and β cleavages as well as the characteristic side-chain loss of C_9H_7N, which dominates the spectra.[342] The tetrapeptides studied have the tryptophan residue at position three and show loss of the α-side chain together with a β_3 cleavage* (see Table **5.4** and Figure **5.16**).

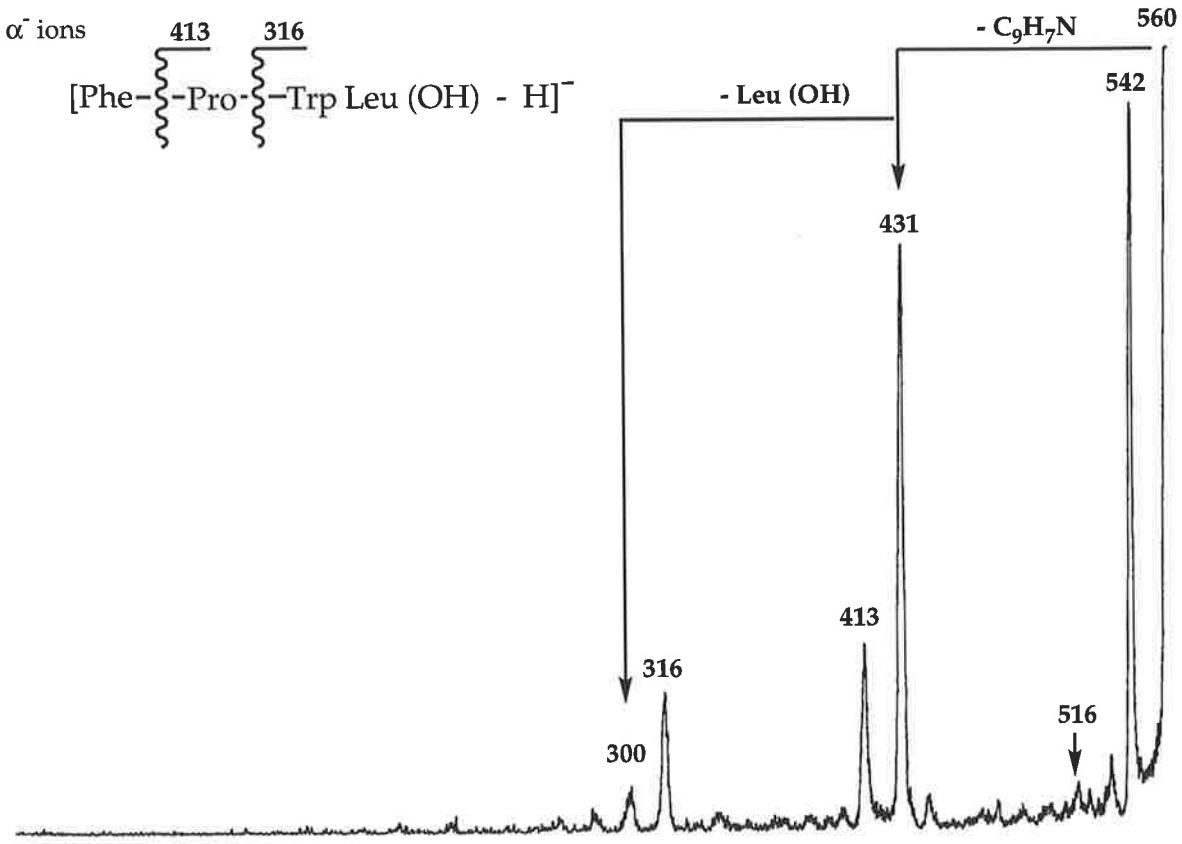


Figure **5.16**: CA MS/MS spectrum of [Phe Pro Trp Leu(OH) - H]⁻ showing that the side chain fragmentation dominates this spectrum.

* It is unclear whether the peptide first loses the side-chain forming [Phe Pro Gly X (OH) - H]⁻, (where X = Pro or Leu) which then forms the β_3^- ion, or *vice versa*. Either way, the formation of [Phe Pro NHCH=C=O - H]⁻ (m/z = 300) is detected.

(v) Histidine

Figure **5.17** is a CA MS/MS spectrum of [Gly Ala Val His (OH) - H]$^-$ (see also Table **5.4**), and shows no cleavages associated with loss of the histidine α-side chain, which is consistent with that reported for dipeptides.(365) The spectrum shows the formation of the three α cleavage ions, but lacks peaks corresponding to β ions. Deprotonation at the β-carbon results in the loss of uracanic acid.

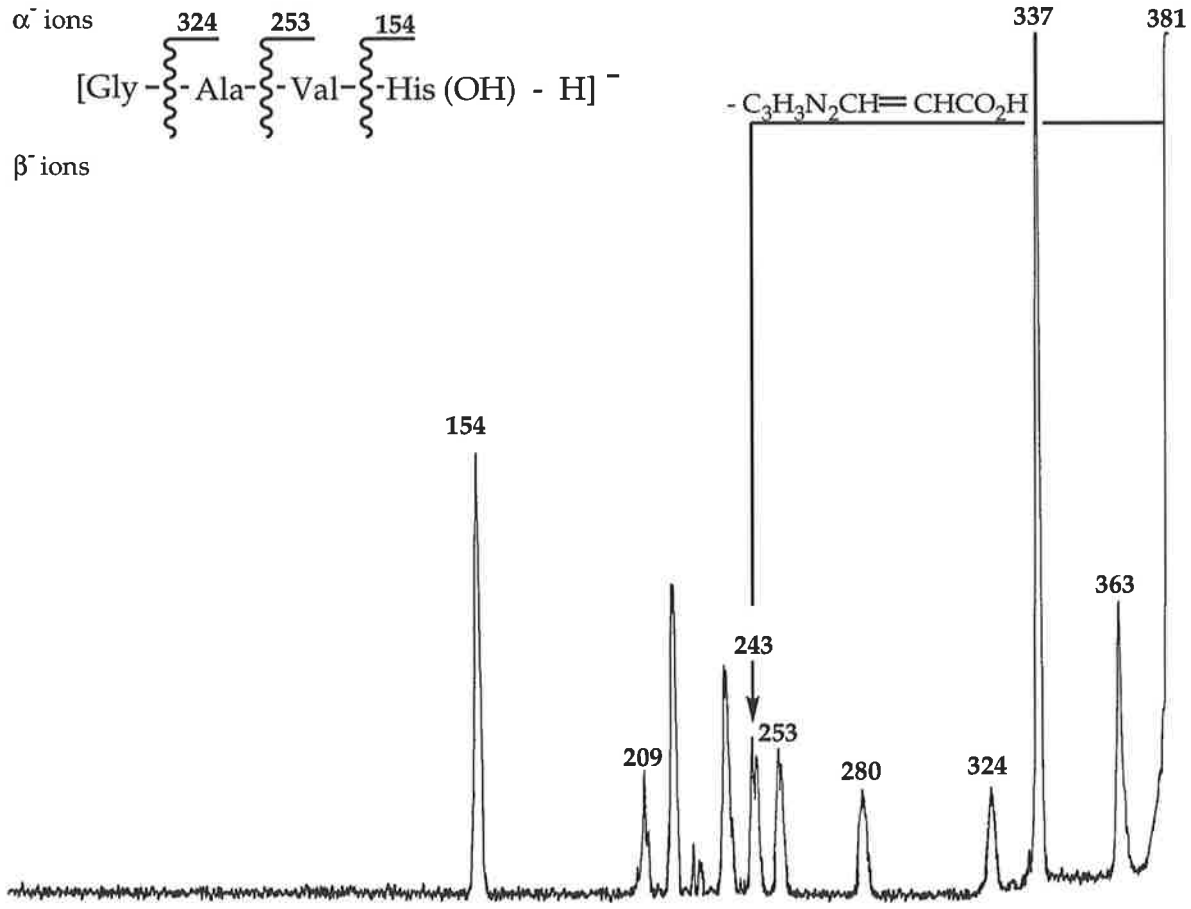
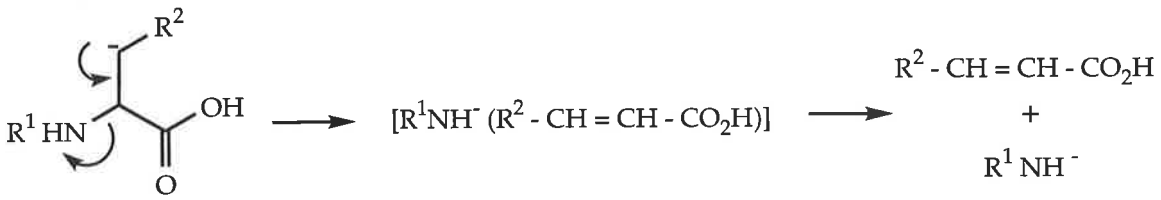


Figure **5.17**: CA MS/MS spectrum of [Gly Ala Val His(OH) - H]$^-$.

5.14 Tetrapeptides Containing Acid or Amide Side Chains

Four of the tetrapeptides studied have acid or amide side chains. The spectra of these tetrapeptides are recorded in Tables **5.5** and **5.6** (page 172) and Figures **5.18 - 5.20** and show quite remarkable differences. When Asp and Asn are C-terminal, no α and β cleavages are observed at all, and the fragmentations observed in the spectra are associated with the α-side chain. Negative ion spectra of dipeptides containing C-terminal Asp and Asn show loss of H_2O[366] (see Table **5.1**), whereas the spectra of tetrapeptides are dominated by the respective losses of $HO_2CCH{=}CHCO_2H$ and $HO_2CCH{=}CHCONH_2$ (see Figures **5.18** and **5.19**). These fragmentations are not observed in all of the $[M - H]^-$ dipeptide spectra.[366] Both aspartic acid and asparagine have acidic hydrogens on the β-carbons, thus these fragmentations are rationalised by scheme (**5.24**).



where the precursor ion is a tetrapeptide, such that:

R^1 = N-terminal residues

R^2 = -COOH, or $-CONH_2$ (**5.24**)

Table 5.5: CA MS/MS Data for $[M - H]^-$ Ions of Tetrapeptides Containing Glu (E), Asp (D), Asn (N). [a]

Precursor Ion (m/z)	Loss				Formation				
	NH_3	H_2O	CO_2	$RCOCH{=}CHCO_2H$ (R=OH or NH_2)	α_1^-	α_2^-	α_3^-	β_2^-	β_3^-
$[GRGD(OH) - H]^-$		34	38	100					
(402) b		(384)	(358)	(286)					
$[EAEN(OH) - H]^-$	100	d	10 [e]	65	5	5		11	26
(460) [c]	(443)		(416)	(345)	(331)	(260)		(199)	(328)
$[NAGA(OH) - H]^-$	100	d	5		8				5
(330) [f]	(313)		(286)		(216)				(241)

(a) Relative abundance (base peak = 100%). Losses of $H^{\cdot}$ and H_2 do occur, but since they are irrelevent to sequence information they have been omitted.

(b) There is a peak at m/z 336 (32%) whose genesis is not known.

(c) This spectrum also shows: m/z 217, [loss of pGlu Asn(OH)], 11%; m/z 242, {loss of Glu Ala(OH), *i.e.* formation of [pGlu Asn (OH) - $H]^-$}, 46%; m/z 128, $[pGlu(OH) - H]^-$, 8%; m/z 313, (α_1 - H_2O), 7%; m/z 398/399, (CO_2 + H_2O/NH_3), unresolved (5%). There are also several peaks whose genesis is unknown: (223, 5%), (298, 8%).

(d) The loss of H_2O and NH_3 are unresolved.

(e) A further loss of CO_2 is observed at m/z 372 (18%), loss of tris-CO_2 has the same integral mass as the β_3^- ion.

(f) This spectrum shows peaks at: m/z 295, (H_2O + NH_3), 23%; m/z 268, (CO_2 + NH_3), 43%.

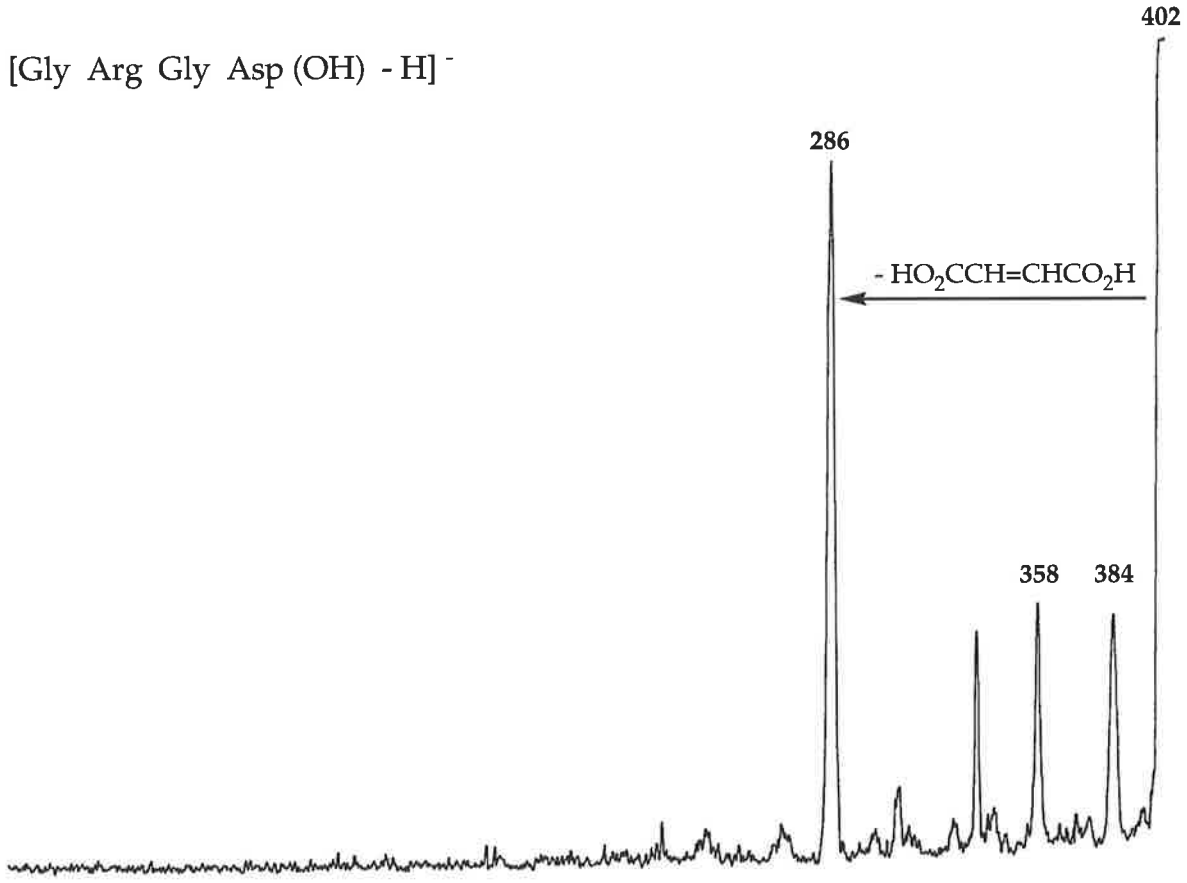
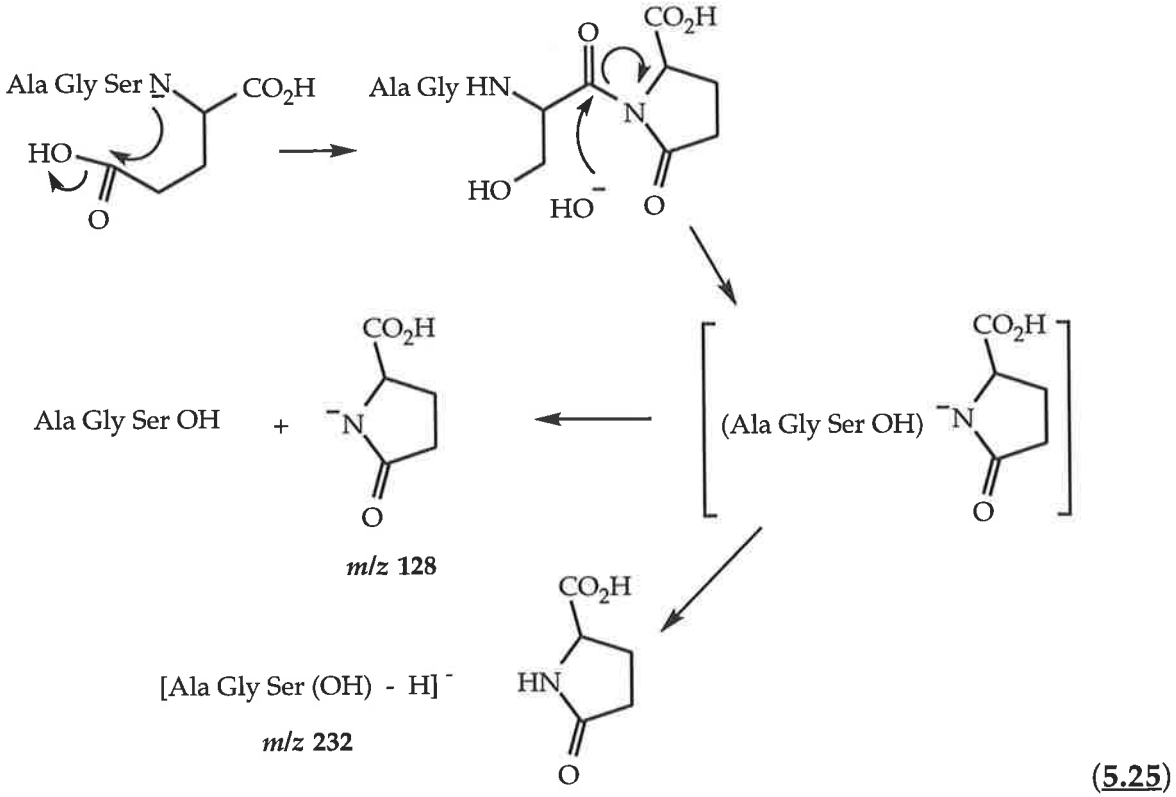


Figure **5.18**: CA MS/MS spectrum of [Gly Arg Gly Asp(OH) - H]$^-$.

In marked contrast, the deprotonated ion from Asn Ala Gly Ala(OH) shows several α and β backbone cleavages (Table **5.5**). This tetrapeptide, which cannot lose $HO_2CCH=CHCONH_2$, shows the characteristic loss of NH_3 associated with asparagine residues. Dipeptides with N-terminal asparagine show significant loss of NH_3 (cf. Table **5.1**). The loss of NH_3 may result from a number of possible pathways. Two of the more plausible include the process described in scheme (**2.24**) for the formation of the maleamide derivative. Alternatively, the formed enolate anion in scheme (**2.24**) may eliminate NH_3 from the amide side chain.

The behaviour of glutamic acid residues are intriguing. Previous studies show that glutamic acid forms the pyroglutamate anion irrespective of its position in the peptide sequence, and does so at high and low energy dissociation pathways.[366, 375] Glu Ala Glu Asn(OH) and Ala Gly Ser Glu(OH) are the two tetrapeptides investigated which have glutamic acid residues. The deprotonated spectra are recorded in Tables **5.5** and **5.6** and are shown in Figures **5.19** and **5.20**. Both show α and β cleavages, but are of small abundance in comparison to the side chain fragmentations. The spectra of Ala Gly Ser Glu (OH) has an intense peak (m/z = 232) which is accounted for by the loss of pyroglutamic acid. The proposed mechanism of formation is shown in scheme (**5.25**) and involves the formation of an ion-complex generated by nucleophilic attack by the amide. The hydroxide ion then displaces the pyroglutamate ion which is observed at m/z 128. (The peaks at m/z 128 and 232 are unresolved from the α_2^- and β_2^- cleavage ions at m/z 127 and 233). Alternatively, the pyroglutamate may be formed by nucleophilic attack on the amide carbonyl by the side chain carboxylate anion, with the elimination of pyroglutamic acid ensuing.[375] Attack by the C-terminal carboxylate is unlikely since this process has not been observed in any spectra of other tetrapeptides. Loss of pyroglutamic acid from the C-terminal residue generates a deprotonated tripeptide.



(5.25)

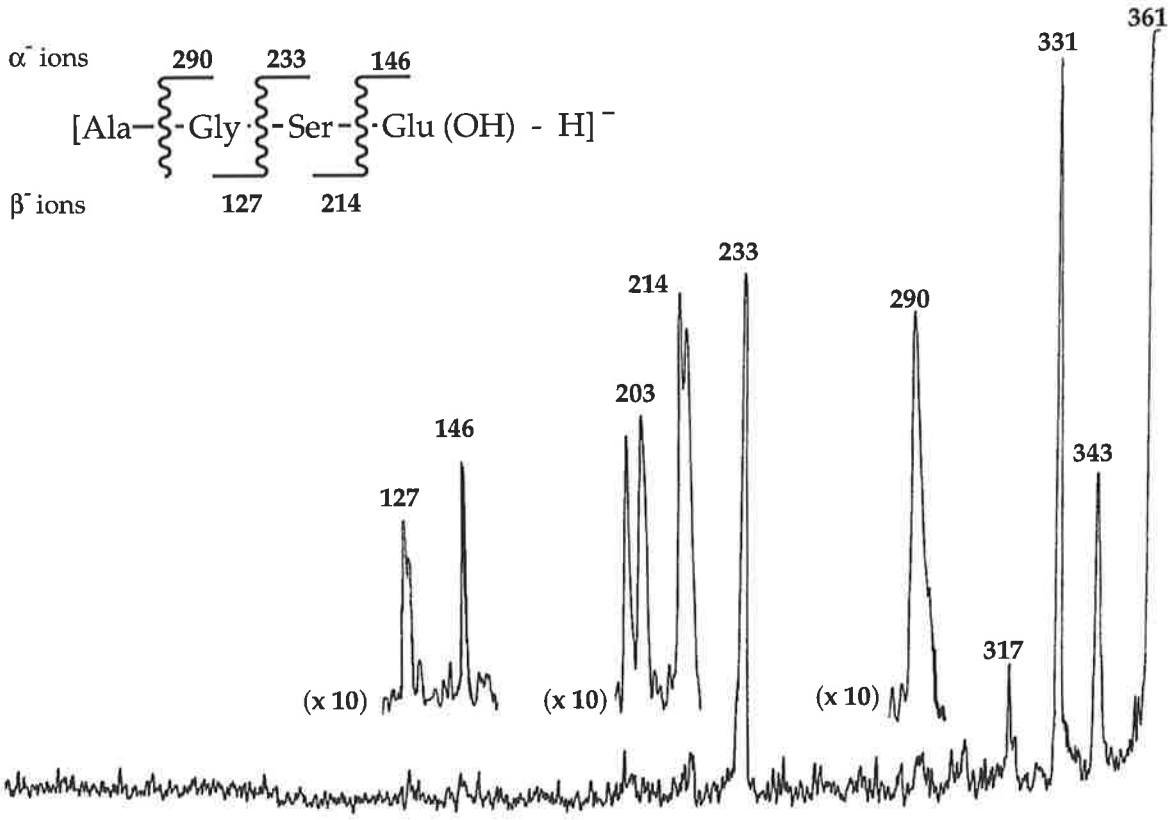
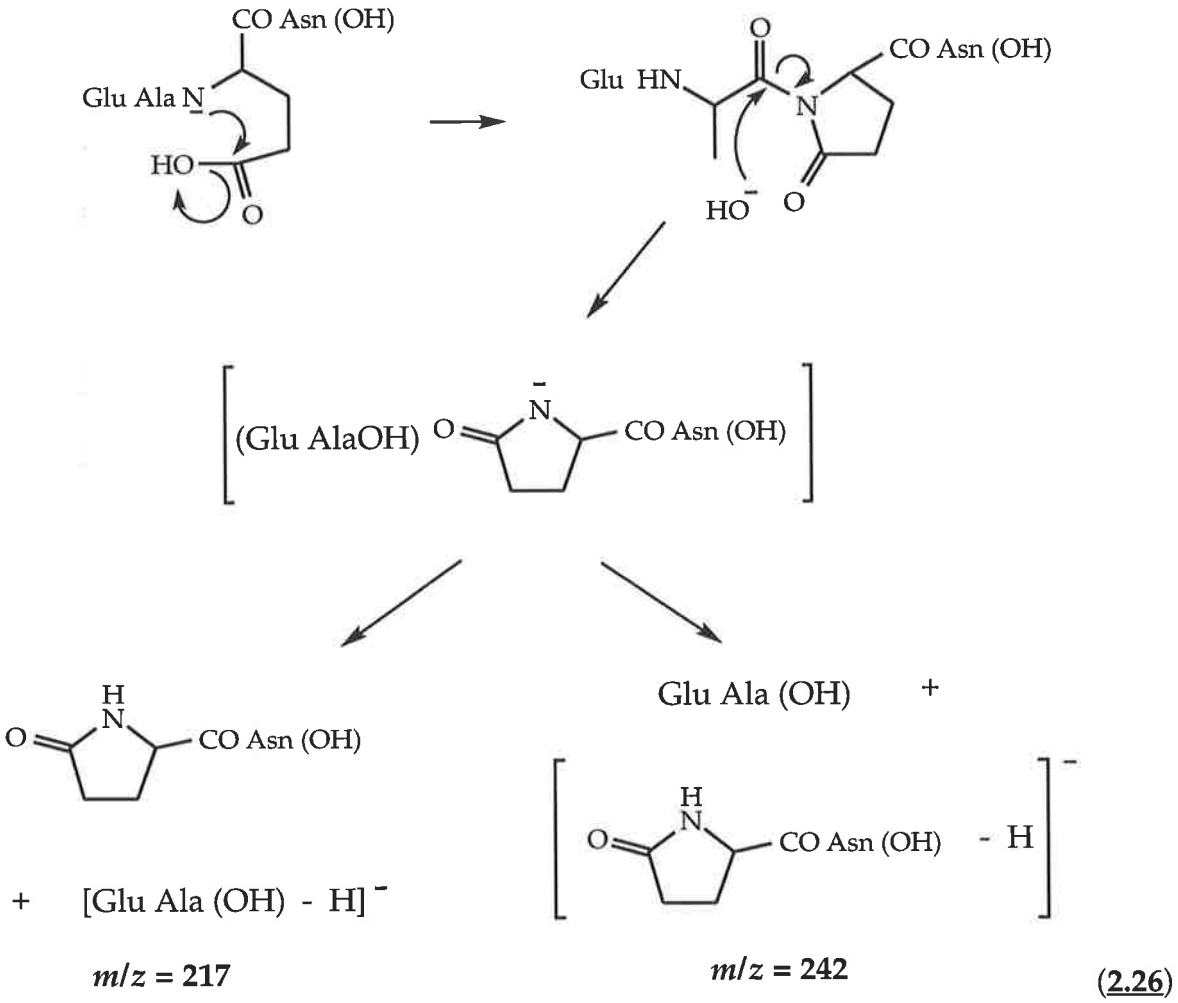


Figure **5.19**: CA MS/MS spectrum of [Ala Gly Ser Glu (OH) - H]$^-$.

When glutamic acid is an internal residue in a peptide, the formation of pyroglutamate still occurs. This is evident in the spectra of [Glu Ala Glu Asn (OH) - H]$^-$ (Table **5.5** and Figure **5.20**), where m/z = 242 results from the loss of neutral Glu Ala (OH), and m/z = 217 results from the elimination of the deprotonated dipeptide as shown in scheme (**5.26**).



(**2.26**)

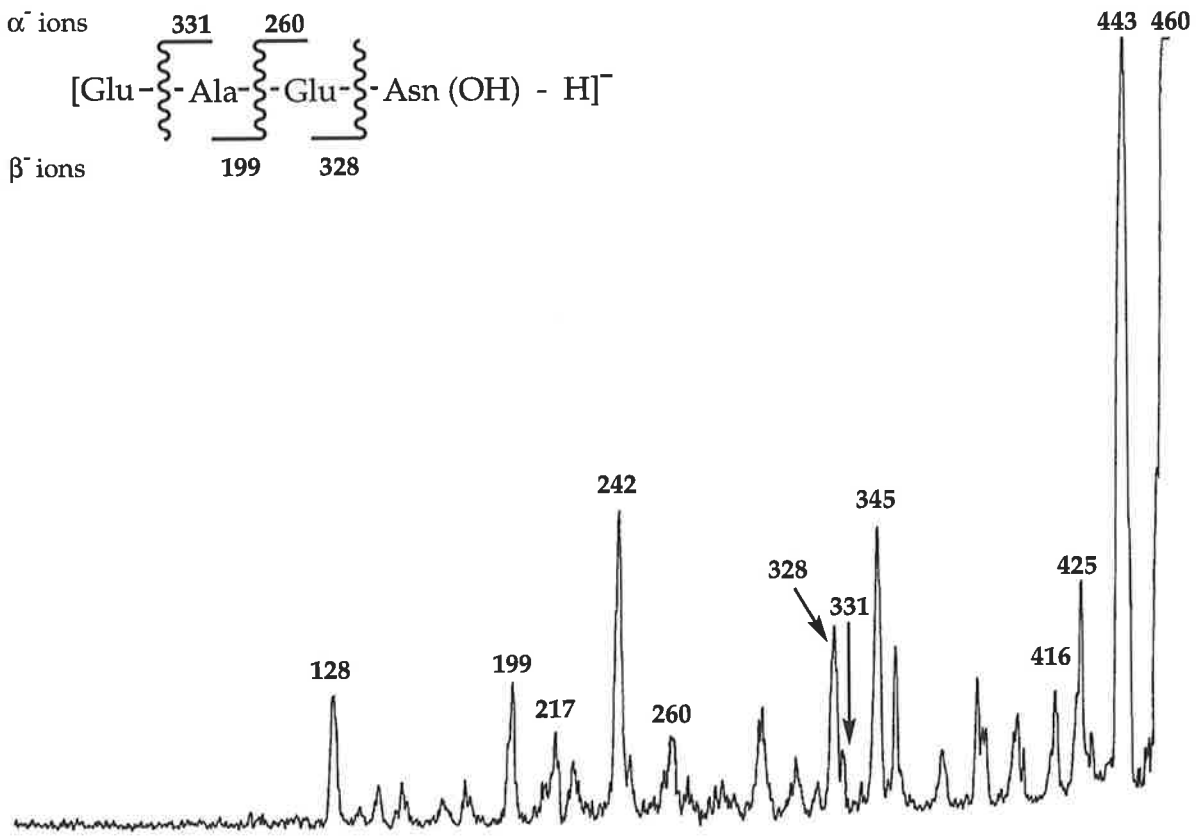
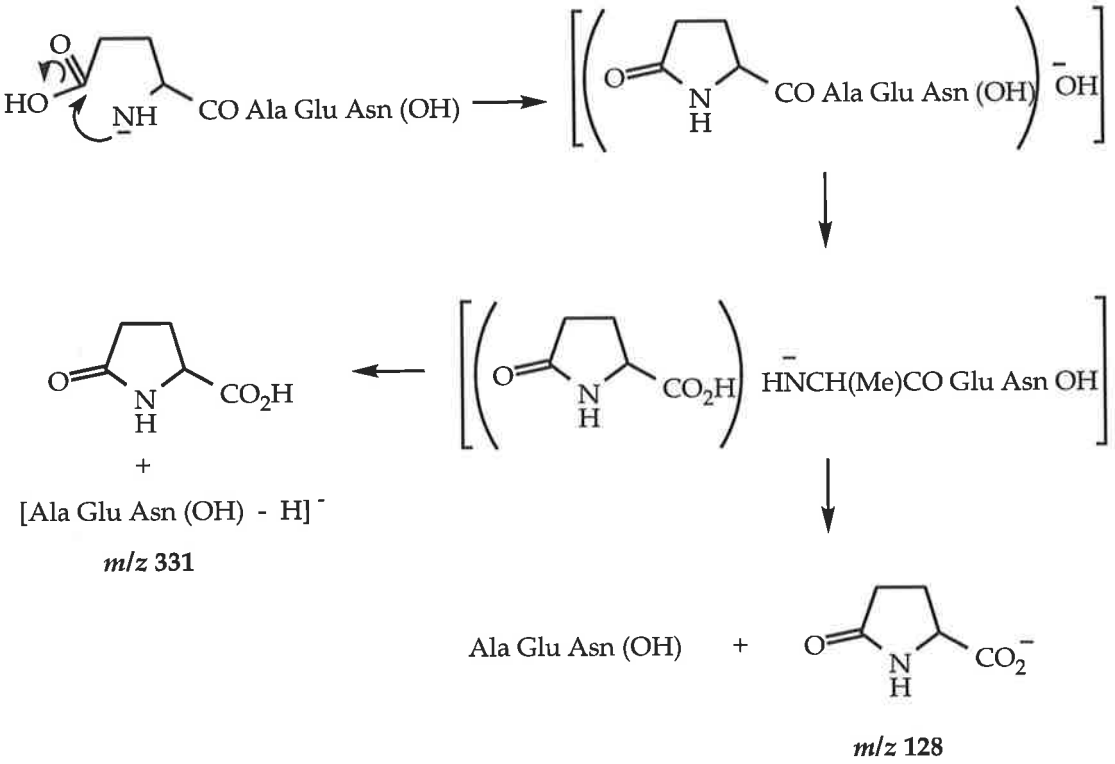


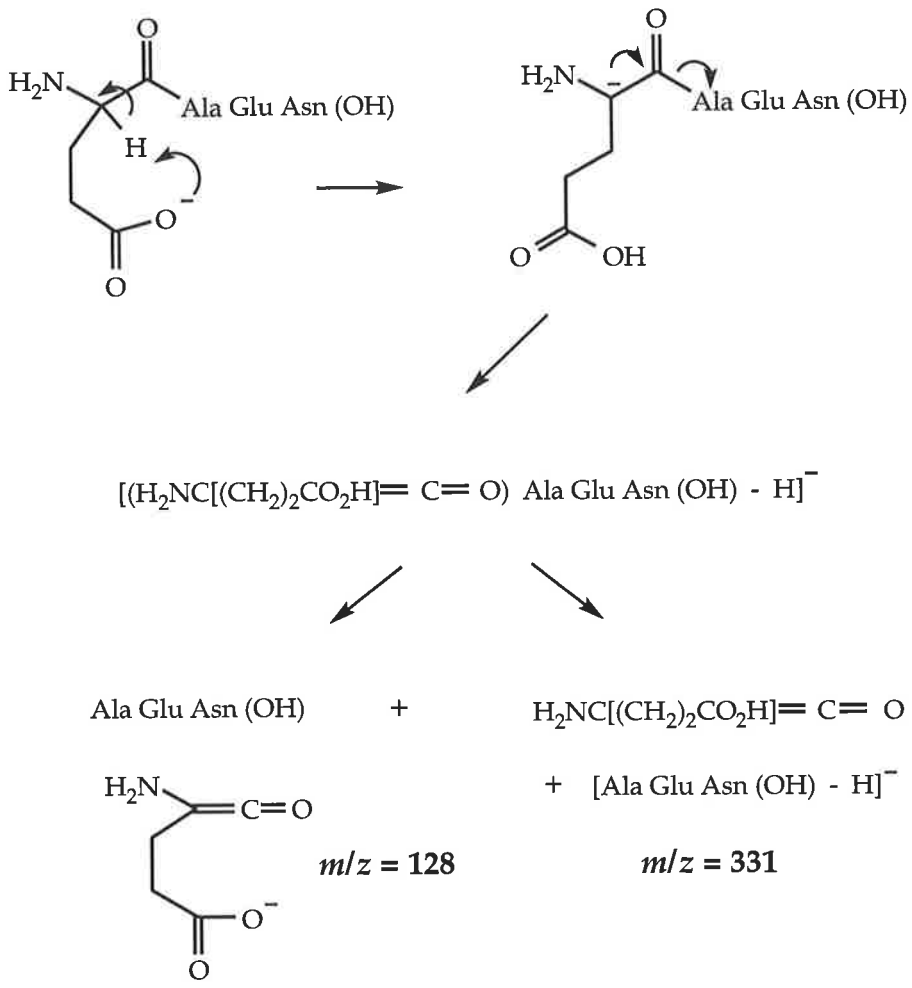
Figure **5.20**: CA MS/MS spectrum of [Glu Ala Glu Asn(OH) - H]⁻.

When glutamic acid is the N-terminal residue of a tetrapeptide, the formation of the pyroglutamate anion (m/z = 128) as well as the loss of pyroglutamic acid are observed [scheme (**5.27**)]. This is rationalised for the peptide Glu Ala Glu Asn (OH) shown in Table **5.5** and Figure **5.20**.



(**2.27**)

The peak at m/z 128 could also arise from a process equivalent to the production of a β_1^- ion. However, β_1^- ions are typically not observed in the spectra of tetrapeptides. In this case the α_1^- ion generated in the ion-complex may deprotonate the carboxyl group of the N-terminal glutamic acid residue [$\Delta H°_{acid\ (side\ chain)}$ = 344 kcal mol^{-1},(376) scheme (**5.28**) adapted from ref (375)]. Overall, the detection of a peak at m/z 128 indicates that a glutamic acid residue occupies a terminal position in a peptide.



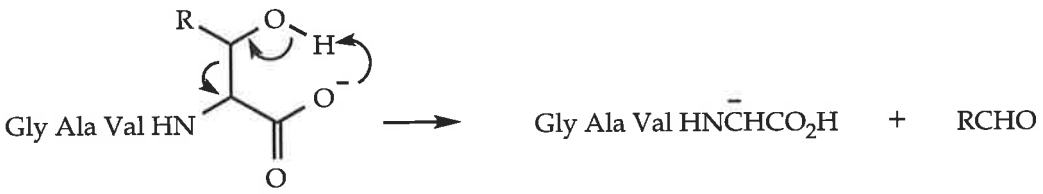
(5.28)

The spectra of deprotonated Glu Ala Glu Asn (OH) illustrates the complexity of analysing negative ion spectra of tetrapeptides which have a number of α-side chains bearing functional groups, all of which may participate in the formation of product ions. A lot of the unassigned peaks are due to a combination of losses of CO_2, H_2O and NH_3 from the α and β cleavage ions. These ions are in small abundance (<5%) compared to the ions which generate sequence information.

5.15 Tetrapeptides with Side Chains Containing Oxygen or Sulphur

Five of the tetrapeptides studied contain Ser, Thr, Cys or Met amino acids. The spectra for these peptides are recorded in Table **5.6** and illustrated in Figures **5.21 - 5.24**.

The characteristic fragmentations of Ser and Thr, involve the losses of CH_2O and MeCHO respectively, and are independent of their positions in a peptide. These fragmentations produce the base peak in the spectra of di-, tri- and tetrapeptides. When Ser and Thr occupy the C-terminal position in a tetrapeptide, these fragmentations occur to the complete exclusion of α and β backbone cleavages [see Table **5.6** and scheme (**5.29**)]. The loss of formaldehyde still forms the base peak of the spectrum when Ser is not C-terminal, but in such a case (*e.g.* Figure **5.21**), some α and β cleavage ions are observed.



Where R = H [*i.e.* Gly Ala Val Ser (OH)]
Me [Gly Ala Val Thr (OH)] (**5.29**)

Table 5.6: CA MS/MS Data for $[M - H]^-$ Ions of Tetrapeptides Containing Glu (E), Ser (S), Thr (T), Cys (C), Met (M). [a]

Precursor Ion	Loss										Formation				
(m/z)	H_2O	CO_2	CH_2O	MeCHO	H_2S	$(H_2S + CO_2)$	MeSH	$(MeSH + CO_2)$	MeSMe	$\cdot(CH_2)_2SMe$	α_1^-	α_2^-	α_3^-	β_2^-	β_3^-
$[AGSE(OH) - H]^-$	43	17	100								6	72[cd]	5	3[d]	7
(361) [b]	(343)	(317)	(331)								(290)	(233)	(146)	(127)	(214)
$[GAVS(OH) - H]^-$	10	4	100												
(331) [e]	(313)	(287)	(301)												
$[GAVT(OH) - H]^-$	10	g		100											11
(345) [f]	(327)			(301)											(226)
$[GAVC(OH) - H]^-$	18	9			100	25									
(347)	(329)	(303)			(313)	(269)									
$[GGFM(OH) - H]^-$	15	100					47	32	26	46	19	23	7		31
(409) [h]	(391)	(365)					(361)	(317)	(347)	(334)	(352)	(295)	(148)		(260)

(a) Relative abundance (base peak = 100%). Losses of $H^{\cdot}$ and H_2 do occur, but since they are irrelevent to sequence information they have been omitted.

(b) Peaks at m/z 216 (7%) and 199 (5%) are unidentified.

(c) This spectrum shows a peak at m/z 203 (5%) corresponding to $(\alpha_2^- - CH_2O)$.

(d) unresolved

(e) This spectrum also shows m/z 269, $(CO_2 + H_2O)$, 9%; m/z 283, $(CH_2O + H_2O)$, 3%.

(f) This spectrum also shows the following peaks: m/z 283, $[H_2O + (CO_2$ or $CH_3CHO)]$, 19%; m/z 244, (α_1^- cleavage from m/z 301), 7%, m/z 200, {m/z 244 - CO_2, *i.e.* formation of $[H_2NCH(Me)CONHCH(iPr)CONHCH_3 - H]^-$}, 6%.

(g) The integral masses of CO_2 and CH_3CHO are both 44.

(h) This spectrum also shows peaks at m/z 130, $\{NH_2CH_2CONHCH_2CONH^-$ [*i.e.* loss of $PhCH{=}CHCONHCH(CH_2CH_2SMe)CO_2H$]}, 11% and m/z 278, $[PhCH{=}CHCONHCH(CH_2CH_2SMe)CO_2^-$ (*i.e.* loss of $NH_2CH_2CONHCH_2CONH_2$)], 13%.

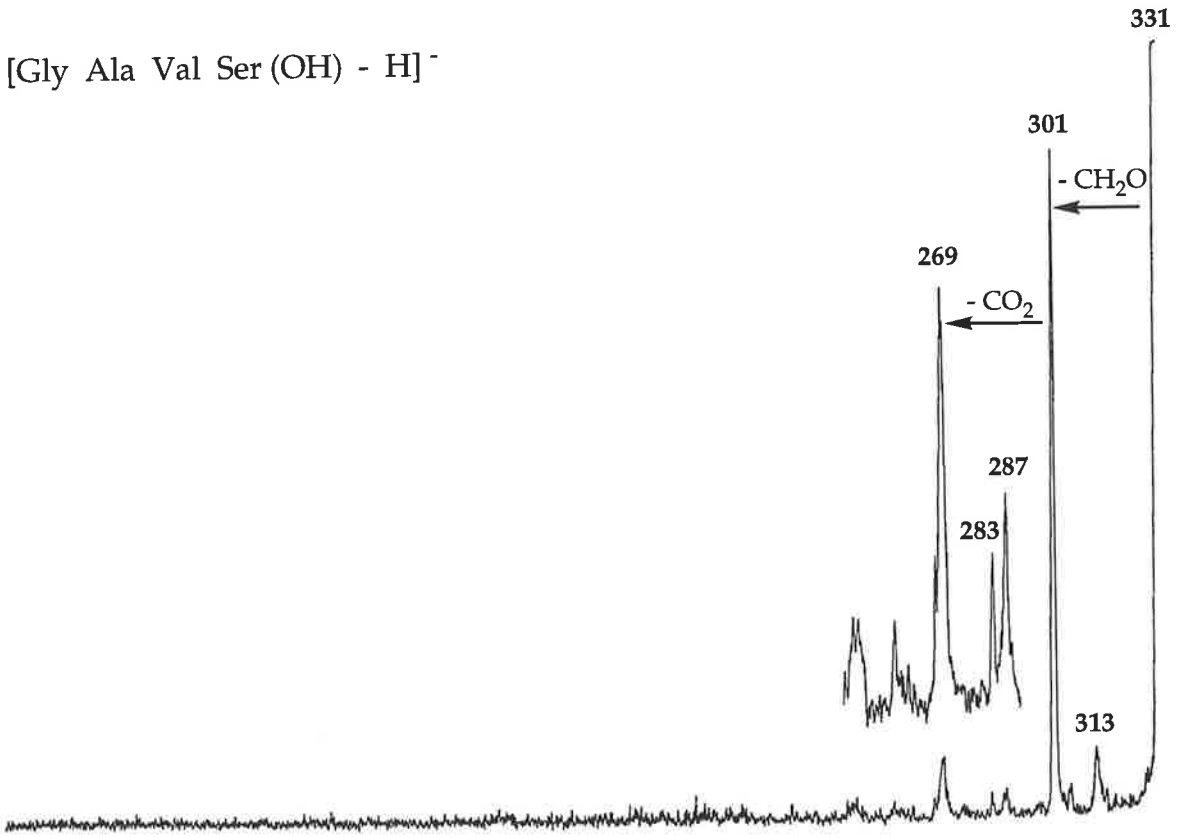


Figure **5.21**: CA MS/MS spectrum of [Gly Ala Val Ser(OH) - H]-.

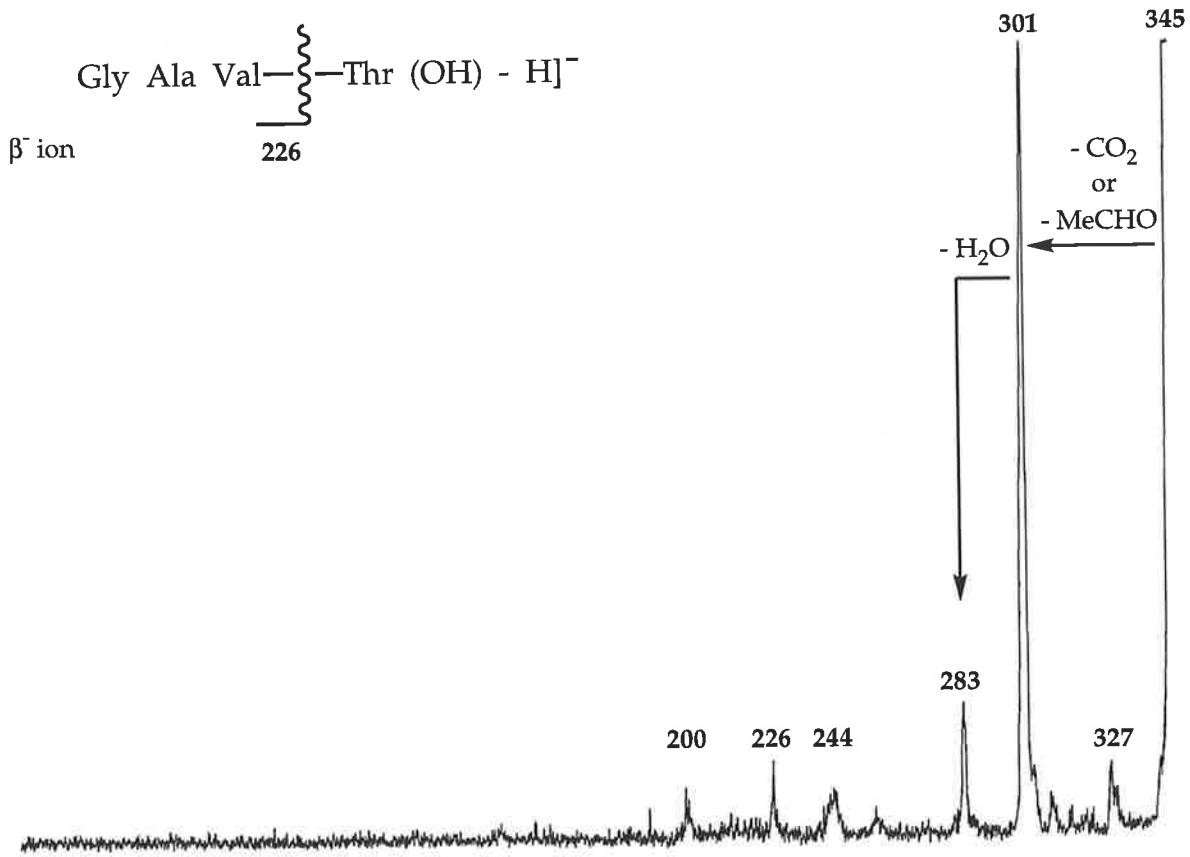


Figure **5.22**: CA MS/MS spectrum of [Gly Ala Val Thr(OH) - H]-.

The presence of a C-terminal Cys residue in a tetrapeptide causes the spectrum of its deprotonated ion to be totally dominated by peaks produced by losses of H_2S and (H_2S + CO_2), [see Table **5.6**, Figure **5.23** and scheme (**5.30**)]. No α and β cleavage ions are formed in this tetrapeptide system. This is a feature that C-terminal Cys has in common with C-terminal Asp, Asn, Ser and Thr.

(**5.30**)

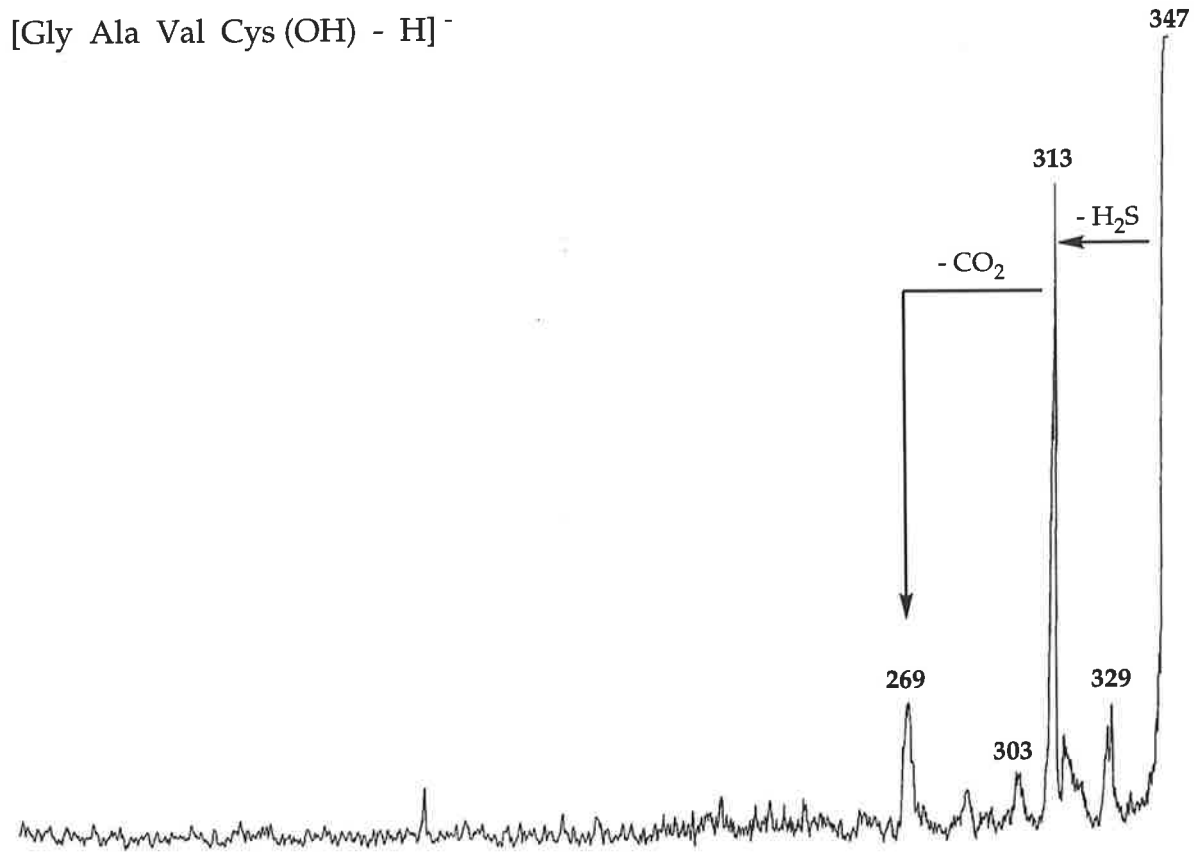


Figure **5.23**: CA MS/MS spectrum of [Gly Ala Val Cys(OH) - H]$^-$.

Finally, methionine is readily identified by negative ion mass spectrometry.[368, 370] Deprotonated ions from di-, tri- and tetrapeptides containing Met all show characteristic losses of MeSH, MeSMe and $MeSCH_2CH_2^{\cdot}$. In the particular case when Met is C-terminal in a tetrapeptide, these fragmentations are observed together with α and β cleavages (see Figure **5.24**). The loss of MeSH is unusual, and has been discussed in previous studies.[368] Such a fragmentation illustrates the difficulty of determining the mechanism of an unusual gas phase reaction in a molecule of this complexity.

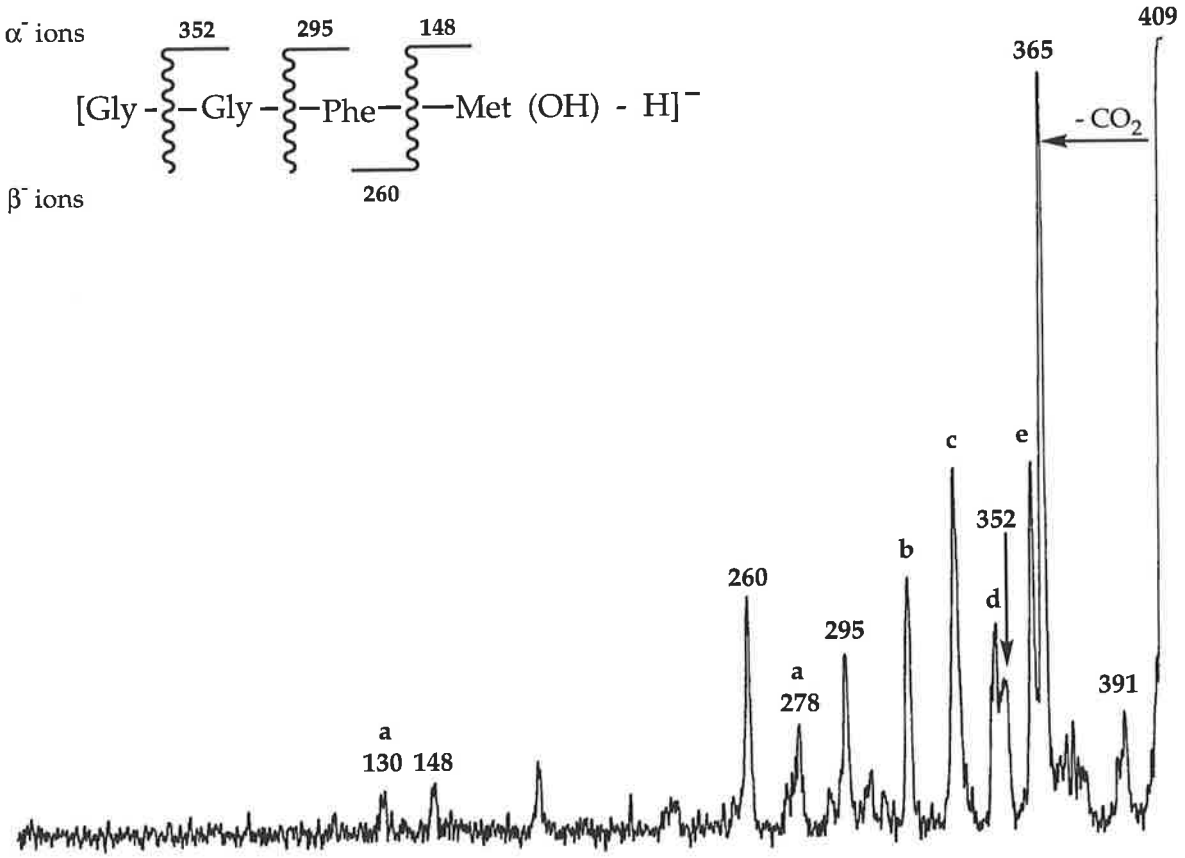


Figure **5.24**: CA MS/MS spectrum of [Gly Gly Phe Met(OH) - H]⁻. Further characterisation is given by: [**a**] peaks at *m/z* 130 and 278 which correspond respectively to $NH_2CH_2CONHCH_2CONH^-$ [loss of PhCH=CHCONHCH-$(CH_2CH_2SMe)CO_2H$] and $PhCH{=}CHCONHCH(CH_2CH_2SMe)CO_2^-$ [loss of $NH_2CH_2CONHCH_2CONH_2$; [**b**] *m/z* 317 [– (MeSH + CO_2)]; [**c**] *m/z* 334 [- ($^{\cdot}CH_2CH_2SMe$)]; [**d**] *m/z* 347 (– MeSMe); and [**e**] *m/z* 361 (– MeSH).

The spectrum of the D_6-derivative of Gly Gly Phe Met(OD) (Figure **5.25**) loses MeSD and MeSH from its $[M-D]^-$ ion in the ratio 3:1, both of these product anions then lose CO_2. The process either has to be an elimination/ deprotonation reaction and/or a charge-remote reaction: both are high energy processes. Using the labelled compound as an illustration, there are a number of possible mechanisms for the losses of MeSD and MeSH within the two scenarios outlined above. Two of the more plausible are:

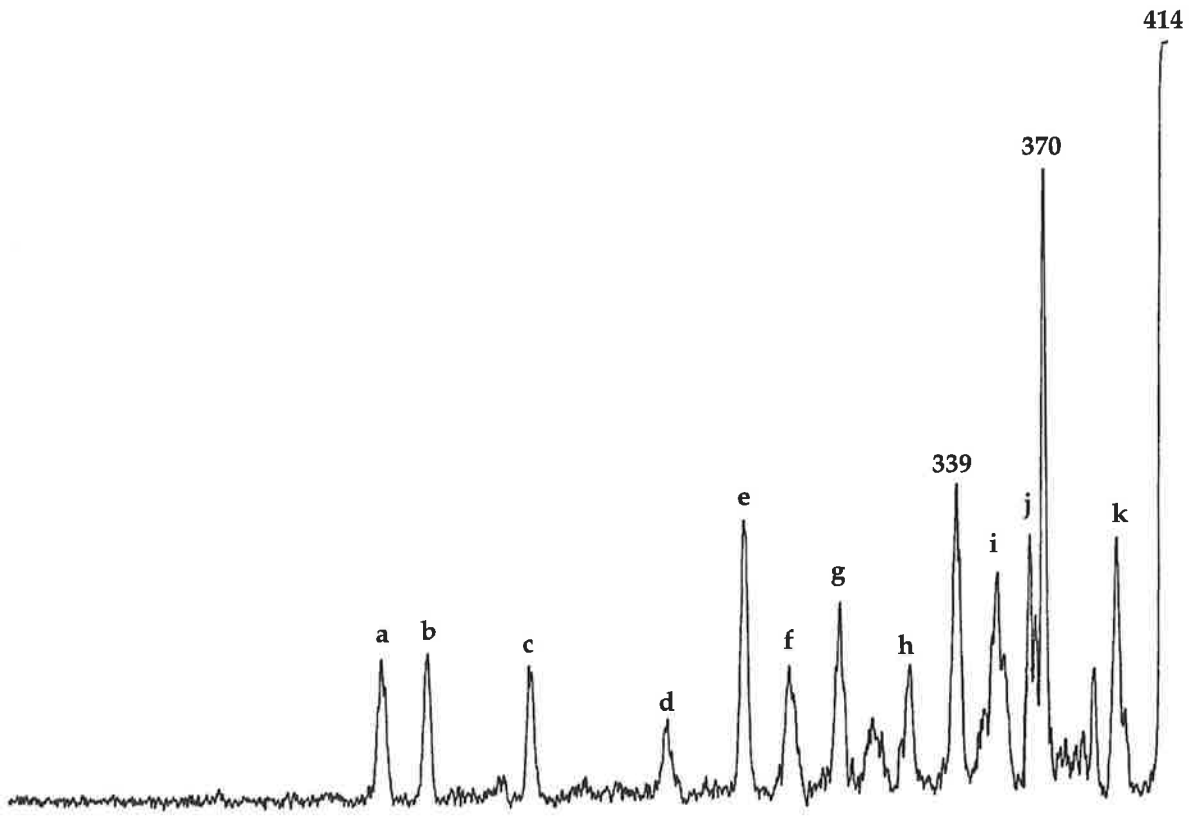
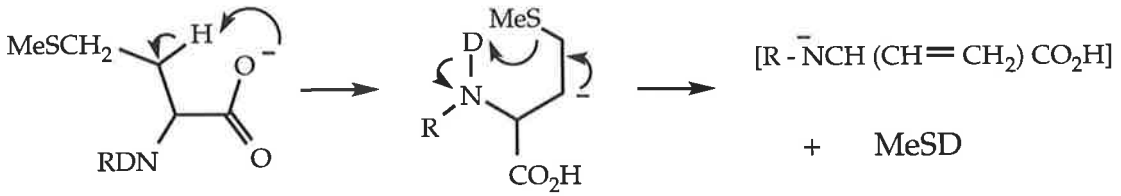


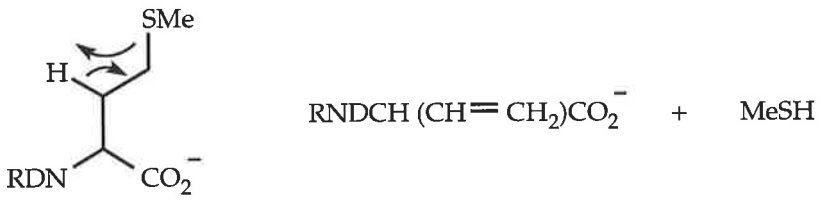
Figure **5.25**: CA MS/MS spectrum of the [D6- Gly Gly Phe Met (OD) – D]$^-$. The spectrum is characterised as follows (the majority of the peaks are not fully resolved) [*m*/*z* (loss or formation) relative abundance(%)]: [**a**] 133/134 [$(ND_2CH_2CONDCH_2COND)^-$, D_3, D_4] 23; [**b**] 149/150 (α_3^-, D_1, D_2) 24; [**c**] 187/188 (genesis unknown) 21; [**d**] 235/236 (**f** - CO_2) 13; [**e**] 263/264 (β_3^-, D_3, D_4) 46; [**f**] 279/280 [$PhCH{=}CHCONDCH(CH_2CH_2SMe)CO_2^-$] 21; [**g**] 297/298 ($\alpha_2^-$, D_2, D_3) 31; [**h**] 321/322 (**1** - CO_2) 21; 339 ($^{\cdot}CH_2CH_2SMe$) 50; [**i**] 355/356 (α_1^-, $D_3<D_4$) 36; [**j**] 365/366 (MeSD/MeSH) 40, 13; 370 (CO_2) 100; 395/396 (HOD/H_2O) 41.

(i) elimination/deprotonation requiring proton transfer to form E_{1cb} intermediate **e** followed by elimination through a six-centre state to form an anion [Scheme (**5.31**)] which may then proton transfer to form the carboxylate anion;



(**5.31**)

(ii) the concerted four-centre mechanism shown in Scheme (**5.32**).



(**5.32**)

5.16 Conclusions

This study of tetrapeptides has shown conclusively that backbone fragmentations do occur and generate a series of ions, which unequivocally provide sequencing information: these cleavages have been named α and β cleavage ions. This study has also shown that the process which forms these two species of ions are generated by a related mechanism by way of an ion-complex. Dissociation of this ion-complex yields α^- ions; whereas proton transfer in the ion-complex generates β^- ions. There are some exceptions to the formation of α and β cleavage ions particularly when the C-terminal residue of the peptide is either Cys, Ser, Thr, Asp or Asn. In these cases the α

and β ions are completely excluded from the spectrum, as fragmentation occurs preferentially through the C-terminal residue side chain. Such fragmentations dominate the spectrum and are characteristic for particular residues; hence they provide a useful insight into the amino acid constituents of the peptide.

Some side chain fragmentations observed in the spectra of dipeptides and tripeptides are still observed for particular residues in tetrapeptide spectra, although the intensity of these fragmentations are diminished. There are specific side chain fragmentations which allow the identification of a residue (*e.g.* Ser, Thr, Cys, Met, Tyr and Trp) irrespective of its position in the peptide. Side chain fragmentations for residues (*e.g.* Phe, His, Asp, Asn and Tyr) in particular positions generally result from proton transfer from the β-carbon, followed by elimination of the N-terminal residues. Also there are particular fragmentations which appear to be due to a juxtaposition of several residues, *e.g.* Pro Phe (Table **5.2** and **5.3**). Additional losses of CO_2, H_2O or NH_3 from any of these ions can also occur, but are usually in small abundance in comparison to ions which generate sequence information [except in the case of Pro Phe Gly Lys(OH), see Table **5.2**].

In many cases, fragmentations in the negative-ion CA MS/MS spectra of tetrapeptides produced by FAB MS are as informative as those from the cognate positive-ion spectra. At the very least, the negative and positive methods provide complementary information. As part of the continuation and development of this exciting field, Steinborner and Bowie have compared the positive and negative ion mass spectra of amphibian peptides.[377] It now remains for researchers to utilise this knowledge of the behaviour and chemistry of peptides in the gas phase to characterise unknown sequences of peptides in both the positive and negative ion modes of operation.

Chapter 6 EXPERIMENTAL

6.1 Mass Spectrometry Analysis

All mass spectra presented in this thesis were acquired using a Vacuum Generators ZAB 2HF reverse sector geometry mass spectrometer[115] (Department of Chemistry, University of Adelaide) equipped with an Ion Tech FAB gun operating with argon gas at an accelerating potential of 7 kV, with a current of 1 mA.[378] It has been discussed in section **1.13** and a schematic diagram is on page 27.

Fast Atom Bombardment Mass Spectrometry (FAB MS)

Routine positive ion FAB mass spectra (required for the detection of protonated molecular ions of peptides, including molecular weight information after Edman degradations and enzyme digests) were obtained by scanning the magnet over a mass range from m/z 3300 to 300 Daltons using a mass resolution of approximately 1000 - 1500, and either: recorded on UV sensitive paper using a Thorn EMI 6150 Mk II UV Oscillograph; or recorded using a PC data station with Maspec Data System for Windows® software (Mass Spectrometry Services software Ltd., Manchester, England), using a scan rate of approximately 70 sec/decade.

Samples for FAB MS analysis were dissolved in water (typically 50 µL) and 1 µL lyophilised onto the FAB probe tip in the pre-vacuum system of the mass spectrometer. The matrix (1 µL), usually either glycerol or magic bullet, was added to the probe tip and mixed thoroughly with the sample before insertion into the ion source.

Collisionally Activated Mass Analysed Ion Kinetic Energy Spectroscopy

Positive ion collisionally activated mass analysed ion kinetic energy spectra (CA MIKES) were obtained by setting the magnet to transmit only the protonated ion under investigation into the second field free region, where the collision cell was filled with argon to a pressure of 2×10^{-6} Torr and the electric sector scanned. The pressure of argon was measured by an ion gauge situated in the second field free region between the collision cell and the electric sector (see Figure **1.3**). This produced a reduction in the main ion beam of 10%, which corresponds essentially to single collision conditions.(110) All slits were fully open to maximise sensitivity and to reduce energy resolution effects.(379) Samples were dissolved in water (typically 50 µL) and 3 x 2 µL lyophilised onto the FAB probe tip in the pre-vacuum system of the mass spectrometer. The matrix (2 µL), usually glycerol or magic bullet,** was added to the probe tip and mixed thoroughly with the sample before insertion into the FAB ion source. Occasionally a trace amount of a matrix additive (typically trifluoro-, trichloro-, or acetic acid) would be mixed thoroughly with the sample prior to analysis, so as to enhance the formation of protonated ions in the mass spectrometer. The spectra were recorded on a paper chart recorder. The ions usually lasted long enough for the collection of at least two quality complete spectra. The masses of the resultant fragment ions were determined manually using a scan range of about 7020 V (*i.e.* E_p^{*} = ~~0~~ maxm V) and a scan rate of 2×10^2 sec/decade. Because of the kinetic shift, the peaks in the ion kinetic energy spectra may be up to one Dalton lower than the expected theoretical mass. However, it is the difference between the peaks that provide the sequencing information and these differences may be reliably determined to one Dalton.

---

* E_p is the electric sector voltage required to transmit the parent ion beam.

** dithiothreitol : dithioerythritol (5:1)

Negative ion CA MIKES spectra were obtained with similar facility, except that the polarity of the magnet and electric sectors were set to transmit negative ions formed from the deprotonated molecular ion $[M - H]^-$ generated in the FAB ion source. Samples were dissolved in water (typically 100 μL) and 2 x 3 μL lyophilised onto the FAB probe tip in the pre-vacuum system of the mass spectrometer. Magic bullet, thioglycerol,* glycerol or d_3-glycerol (for deuterated peptides) were used as matrices (2 μL) and occasionally a trace amount of tetramethylammonium hydroxide in methanol would be added to enhance the formation of deprotonated ions. The spectra were recorded on a paper chart recorder and the ions usually lasted long enough for the collection of at least two quality complete spectra.

6.2 Surface Electrical Stimulation and Preparation of Frog Skin Secretions

The granular secretions from the frogs were obtained through the courtesy of Associate Professor M. J. Tyler, Department of Zoology, University of Adelaide. Surface electrical stimulation[148] of the frog skin glands was achieved by using a C. F. Palmer 'Student Model' electrical stimulator, with a bipolar electrode of 21 G platinum. The frogs were held by the back legs, the skin moistened with deionised water and surface electrical stimulation was performed by gently rubbing the electrode in a circular motion over the entire skin. Operating conditions usually consisted of: a pulse duration of 2.5 msec; a pulse repetition of 50 Hz; and a stimulus strength of 2 volts. The resulting secretion was collected by washing the frog with a stream of deionised water (approximately 20 mL). The volume was doubled with methanol and stored at

* Thioglycerol was used for the tetrapeptide Glu Ala Glu Asn(OH) ($[M-H]^-$ = 460), since oligomers of glycerol and magic bullet appear at m/z 459 and 461 respectively.

-10° C for 24 hours (to precipitate and deactivate any proteins, proteases and enzymes). The secretion was centrifuged at 3000 rpm for 5 mins (Clements GS 100 centrifuge) and the supernatant liquid collected. The crude precipitate was washed with fresh methanol / acetic acid 1:1, centrifuged and the supernatants combined. The volume of supernatant was carefully reduced to about 5 mL *in vacuo* (using a rotary evaporator, < 40° C, 14 mm Hg). The residue was passed through a 0.45 µm Millex filter unit (Millipore, Bedford, MA., USA), and lyophilised. This extract can be stored for up to 12 months at -10° C, without any detectable changes in the composition of the peptide mixture.

Seven specimens of *Uperoleia inundata* afforded on average 3 mg of crude peptide mixture per frog; whereas 9 specimens of *Uperoleia mjobergi* afforded on average 7 mg per frog.

6.3 Analytical and Preparative HPLC

Analytical HPLC separations were achieved using a VYDAC 218TP54 protein and peptide c18 HPLC column (5 µ, 300 Å, 4.6 mm id x 250 mm, Separations Group, Hesperia, CA., USA), fitted with a VYDAC 218TP 300 Å guard column, equilibrated with acetonitrile / water [1:9] contaminated with 0.1% trifluoroacetic acid as an ion pairing agent. Analytical traces were primarily used for (i) purifying fractions; and (ii) as a "guide map" for preparative work. The elution profiles were generated using a linear gradient produced by an ICI DP 800 Data Station controlling two ICI LC 1110 HPLC pumps, increasing typically from 10 - 75% acetonitrile over 30 mins at a flow rate of 1 mL/min. The eluent was monitored by ultraviolet absorbance at 214 nm light using an ICI LC 1200 variable wavelength detector. Samples were injected into a Rheodyne injector fitted with a 200 µL injection loop using the appropriate sized Hamilton syringe equipped with the appropriate sized Rheodyne needle.

Preparative HPLC separations were achieved using a VYDAC 218TP510 protein and peptide c18 HPLC column (5 μ, 300 Å, 10 mm id x 250 mm),* fitted with a VYDAC 218TP 300 Å guard column, equilibrated with acetonitrile / water [1:9], contaminated with 0.1% trifluoroacetic acid as an ion pairing agent. The elution profiles were generated using a linear gradient produced by a Waters Automated Gradient Controller controlling two Waters HPLC pumps (pump A, model 501, aqueous; pump B, model 510, organic) increasing typically from 10 - 75% acetonitrile over 30 mins at a flow rate of 4.5 mL/min. The eluent was monitored by ultraviolet absorbance at 214 nm light using a Waters Lambda-Max model 481 LC variable wavelength detector. Samples were injected into a Waters UK6 injector, fitted with a 2 mL injection loop using an appropriate sized Hamilton syringe equipped with the appropriate sized Waters needle. Data were monitored by an ICI DP 800 Data Interface and an ICI DP 700 Chromatography Data Station.

All peptide separations of *U. inundata* were achieved by using the VYDAC 218TP54 column. A typical elution profile of the separation of the crude extract is shown in Figure **3.4**. Preparative separations were achieved by dissolving the lyophilised peptide mixture in Milli-Q water (2000 μL) and injecting volumes of 25 μL. Up to 65 individual injections afforded workable amounts of material from each of the 34 separate fractions collected. Fractions from individual separations were collected, combined (as appropriate), concentrated and dried *in vacuo*. Each fraction was analysed by FAB MS. Purification was carried out using the above systems with the VYDAC 218TP54 column, except that a gradient of no more than 5% was used, such that the retention time of the sample was in the range 20 - 30 mins.

---

* The purchase of the 218TP510 column arrived after the completion of the separation of *U inundata*, hence all *U inundata* preparative separations were performed using the 218TP54 column (although the HPLC machines were not interchanged).

An analytical separation of the crude peptide secretion from *U. mjobergi* was achieved by using the VYDAC 218TP54 column. A typical elution profile of the crude extract is shown in Figure **4.3**, which was generated by injecting a 5 μL sample (dissolved in 1000 μL). This analytical profile served as a 'guide map' for preparative separations, which were adequately achieved by using the VYDAC 218TP510 column. The lyophilised peptide mixture was dissolved in Milli-Q water (500 μL), with injection volumes of 50 μL (approximately 6 mg/injection) affording workable amounts of material from each of the 20 separate fractions collected. Fractions from individual separations were combined (as appropriate), concentrated and dried *in vacuo*, and analysed by FAB MS. In some circumstances the fraction was a composition of two or more peptides. Where this was observed the peptide fraction was purified using the VYDAC 218TP54 column with a linear gradient changing by a maximum of 5% acetonitrile (+ 0.3% TFA) over 30 mins, such that the retention time of the peptide mixture was between 20 and 30 mins.

6.4 Automated Sequencing

The structure assignments for the peptides were confirmed by N-terminal sequencing using Edman chemistry,[328] followed by on-line (microgradient) PTH-amino acid analysis and was performed on an Applied Biosystems 470A automated protein sequencer equipped with a 900A data analysis module. Initial attempts to sequence the uperin peptides gave no more than four successful sequence cycles before all the peptide was washed from the glass fibre filter. Good results were obtained using a disc of polyvinylidene difluoride (PVDF) membrane (Immobilon P, Millipore) treated with BioPrene Plus™ (Applied Biosystems, Inc.) in ethanol, onto which the peptide (typically 300 pmol lyophilised) was absorbed from aqueous acetonitrile (90%). The disc was pierced several times with a razor blade in order to aid the flow of solvent.

6.5 Chemicals

All solvents were purified by standard methods prior to use. Milli-Q (deionised) water and Ajax HPLC grade acetonitrile 190 were used for HPLC. Heptafluorobutyric acid (HFBA) was purchased from Aldrich. Trifluoroacetic acid (TFA) spectroscopic grade was purchased from Sigma. HPLC solutions were filtered through a 0.22 μm membrane (Millipore), then sonicated prior to use. Phenylisothiocyanate (PITC) was purchased from Applied Biosystems.

The tetrapeptides Phe Pro Trp Pro (NH_2) and Phe Pro Trp Leu (NH_2) were isolated from the Australian Red Tree Frog, *Litoria rubella* (from Lake Argyle)[198, 288] and converted into the carboxylic acid as follows. The peptide (500 nmol) was dissolved in 0.1 M HCl (200 μL) and incubated at 40° C for 2 hours. The solution was lyophilised and analysed by FAB MS. Gly Ala Val X (where X = Cys, His, Ser, and Thr) were commercial samples from Chiron Mimotopes, Clayton, Victoria, Australia. Remaining tetrapeptides were commercial samples from Sigma, St. Louis, USA.

The deuterated samples were formed by the following general procedure. All procedures were performed under a blanket of high purity nitrogen. The tetrapeptide (50 μg) was dissolved in D_2O (200 μL) in a 1.5 mL polypropylene Eppindörf centrifuge tube, and lyophilised to remove H_2O, HOD, and D_2O. This procedure was repeated 3 times. Samples were then dissolved in D_2O (50 μL) and 3 μL removed for CA MIKES analyses, using d_3-glycerol as the matrix. On the basis of the full scan FAB spectra of the deuterated tetrapeptides, at least 75 - 90% of the exchangeable hydrogens were replaced.

6.6 Manual Edman Degradations

The following is an adaptation of the procedure used by Bradley and Williams.[380]

The peptide (200 nmol) was dissolved in water (40 µL) and redistilled pyridine (40 µL) in a 1.5 mL polypropylene Eppindörf centrifuge tube. A solution of phenylisothiocyanate in heptane (PITC, 5%, 40 µL, protein sequencing grade, Applied Biosystems) was added, the mixture purged with oxygen free nitrogen, capped, incubated at 40° C for 1 hour, washed with hexane : ethyl acetate (2:1 v/v, 2 x 50 µL) and lyophilised. Trifluoroacetic acid (50 µL) was added, purged with oxygen free nitrogen, capped, incubated at 40° C for 30 min and lyophilised. The residue was suspended in deionised water (50 µL) and the phenylthiocarbodiimide amino acid derivative extracted into *n*-butyl acetate (50 µL). A 2 - 5 µL aliquot of the aqueous layer was removed and subjected to FAB MS analyses. The remainder of the solution was lyophilised in readiness for the next cycle of the degradation process.

6.7 Enzymic Digests

Lys-C digests

A solution of endoprotease Lys-C (from *Lysobacter enzymogenes*, Sigma, St. Louis, MI, USA), was prepared by dissolving 3 units in Milli-Q water (150 µL). The lyophilised peptide (generally 100 nmol) was dissolved in aqueous ammonium hydrogen carbonate (0.1 M, 25 µL, pH = 8.0) and endoprotease Lys-C (2 µL) added. The solution was incubated at 40° C for 2 hours, with the products analysed directly by FAB MS and occasionally CA MIKES.

Asp N digests

A solution of HPLC purified endoprotease Asp-N (from *Pseudomonas fragi* [mutant strain], Sigma, St. Louis, MI, USA), was prepared by dissolving 2 μg in Milli-Q water (50 μL). The lyophilised peptide (generally 100 nmol) was dissolved in TRIS buffer (0.1 M, 25 μL, pH = 8.0) and endoprotease Asp-N (1 μL) added. The solution was incubated at 40° C for 18 hours, with the products analysed directly by FAB MS and occasionally CA MIKES.

Arg C digests

A solution of endoprotease Arg-C (from mouse submaxillary gland, Sigma, St. Louis, MI, USA), was prepared by dissolving 100 units in Milli-Q water (5000 μL). The lyophilised peptide (generally 100 nmol) was dissolved in phosphate buffer (KH_2PO_4:K_2HPO_4, 96:1, 0.1 M, 40 μL, pH = 8.0) and endoprotease Arg-C (2 μL) added. The solution was incubated at 40° C for 2 hours, with the products analysed directly by FAB MS and occasionally CA MIKES.

α-Chymotrypsin digests

A solution of α-chymotrypsin (type II from bovine pancreas, Sigma, St. Louis, MI, USA), was prepared by dissolving 20 units in Milli-Q water (1000 μL). The lyophilised peptide (generally 100 nmol) was dissolved in TRIS buffer (0.1 M, 40 μL, pH = 8.0) and α-chymotrypsin (2 μL) added. The solution was incubated at 40° C for 3 hours, with the products analysed directly by FAB MS and occasionally CA MIKES.

6.8 Conversion to Methyl Esters

Peptides were converted to their methyl esters by the following general procedure.

A stock solution of 2N acidic methanol was prepared by adding acetyl chloride (242 μL) to solid methanol (758 μL, frozen by immersion in liquid nitrogen) the vessel tightly stoppered and the solution gradually warmed to room temperature. A vigorous reaction occurred, the pressure was released and the vessel stoppered and stored at 5° C. The stock solution is stable for months at this temperature.

To the lyophilised peptide (50 nmol) was added acidic methanol (300 μL) and the mixture incubated at 40° C for 2 hours, then lyophilised and submitted for FAB MS analysis. No consideration was taken to prevent esterification of amide or acid functionalities of amino acid side chains.

6.9 Preparation of Synthetic Peptides

The uperins* selected for biological testing were synthesised commercially[381] using L-amino acids *via* the standard α-Fmoc method.[382] Each synthetic uperin was shown to be identical with the natural uperin by (i) FAB mass spectrometry, (ii) Edman sequencing, and (iii) co-elution of the synthetic and natural peptides on HPLC.

---

* For financial reasons, not all uperins could be synthesised.

6.10 Biological Testing

Uperin 1.1 was sent to the Universita Degli Studi di Roma to be tested by V. Erspamer and colleagues. The details of the bioassay are as follows:

The bioassay of uperin 1.1 was carried out in parallel with uperolein and physalaemin, the latter being the prototype of the aromatic, NK1-receptor preferring tachykinins.

Test systems were five isolated smooth muscle preparations (guinea-pig ileum, colon and urinary bladder, rat urinary bladder, rabbit terminal colon) and the rabbit blood pressure. All smooth muscle preparations were suspended in a bath of 10 mL Tyrode solution at 37° C, oxygenated by 95% O_2 and 5% CO_2. Contractions of the smooth muscles were recorded isometrically by a strain-gauge transducer (DY2, force up to 10 g) and displayed on a recording microdynamometer (Basile, Vareses, Milan). Carotid blood pressure of the anaesthetised rabbit (ethyl urethane), 1.5 g/Kg, intraperitoneally) was recorded from a carotid artery by a Trantee pressure transducer (mod. 880; Bentley, Irvine, California) connected to a microdynamometer.

The injections of the peptide were given through a polyethylene tube inserted into the jugular vein. The results are shown in Table **3.15**. The activity of physalaemin was considered equal to 100, that of the other two tachykinins was expressed as a percentage of this activity. The ranges are given in parentheses; the number of experiments in the square parentheses.

Other uperin peptides were submitted for antimicrobial activity at the Institute for Medical and Veterinary Science, Adelaide. The results are listed in Tables **3.14** and **4.14**.

APPENDIX A: USEFUL DATA REGARDING AMINO ACIDS

Amino Acid	Symbols		Residue Structure -NHCH(R)CO- R =	Monoisotopic Mass	Bull and Breese[218]	Acidities ΔH_a (kcal mol^{-1})	
						side chain[376]	amino acid[383]
Alanine	Ala	A	$-CH_3$	71.03711	610	> 390	341.2
Arginine	Arg	R	$-CH_2(CH_2)_2NHC(NH)NH_2$	156.10111	690	> 390	331.9
Asparagine	Asn	N	$-CH_2CONH_2$	114.04293	890	362	331.7
Aspartic acid	Asp	D	$-CH_2COOH$	115.02694	610	345	-
Cysteine	Cys	C	$-CH_2SH$	103.00919	360	355	332.9
Glutamic Acid	Glu	E	$-CH_2CH_2COOH$	129.04259	510	344	-
Glutamine	Gln	Q	$-CH_2CH_2CONH_2$	128.05858	970	362	331.7
Glycine	Gly	G	-H	57.02146	810	> 390	342.0
Histidine	His	H	$-CH_2(C_3H_3N_2)$	137.05891	690	353[384]	331.0
Isoleucine	Ile	I	$-CH(CH_3)CH_2CH_3$	113.08406	- 1450	> 390	338.8
Leucine	Leu	L	$-CH_2CH(CH_3)_2$	113.08406	- 1650	> 390	339.1
Lysine	Lys	K	$-CH_2(CH_2)_3NH_2$	128.09496	460	> 390	337.5
Methionine	Met	M	$-CH_2CH_2SCH_3$	131.04049	- 660	> 390	335.8
Phenylalanine	Phe	F	$-CH_2Ph$	147.06841	- 1520	380	336.5
Proline	Pro	P	$-CH_2CH_2CH_2-$	97.05276	- 170	> 390	340.3
Pyroglutamic Acid	pGlu	B	$-CH_2CH_2C(O)-$	111.03203	-	-	-
Serine	Ser	S	$-CH_2OH$	87.03203	420	375	332.7
Threonine	Thr	T	$-CH(OH)CH_3$	101.04768	290	373	332.1
Tryptophan	Trp	W	$-CH_2(C_8H_6N)$	186.07931	- 1200	352[384]	336.9
Tyrosine	Tyr	Y	$-CH_2PhOH$	163.06333	- 1430	347	356.4
Valine	Val	V	$-CH(CH_3)_2$	99.06841	- 750	> 390	339.4

REFERENCES

(1) J. J. Thomson, in "Rays of Positive Electricity and Their Application to Chemical Analysis." Longmans, London, 1913.

(2) E. Goldstein, Berl. Ber., 1886, 39, 691.

(3) W. Wein, Ann. Physik, 1898, 65, 440.

(4) J. J. Thomson, Phil. Mag., 1911, 21, 225.

(5) A. J. Dempster, Phys. Rev., 1918, 11, 316.

(6) F. W. Aston, Phil. Mag., 1919, 38, 707.

(7) H. W. Washburn, H. F. Wiley and S. M. Rock, Ind. Eng. Chem. Anal. Ed., 1943, 15, 541.

(8) A. O. C. Nier, Rev. Sci. Instr., 1940, 11, 212.

(9) R. Herzog, Z. Phys., 1934, 89, 447.

(10) J. Mattauch, R. Herzog, Z. Phys., 1934, 89, 786.

(11) M. L. Gross, E. K. Chess, P. A. Lyon, F. W. Crow, S. Evans and H. Tudge, Int. J. Mass Spectrom. Ion Phys., 1982, 42, 243.

(12) M. L. Gross, Methods Enzymol., 1990, 193, 131.

(13) J. B. Farmer, "Mass Spectrometry", C. A. McDowell Ed., McGraw - Hill Book Co. Inc., New York, New York, 1963.

(14) W. Paul and H. Steinwedel, Z. Naturforsch., 1953, 8A, 448.

(15) R. A. Yost and C. G. Enke, J. Am. Chem. Soc., 1978, 100, 2274.

(16) E. E. Ferguson, F. C. Fehsenfeld, A. L. Schmeltekopf, Adv. Atom. Mol. Phys., 1969, 5, 1.

(17) C. H. Depuy and V. M. Bierbaum, Acc. Chem. Res., 1981, 14, 146.

(18) W. C. Wiley and I. H. McLaren, Rev. Sci. Instrum., 1955, 26, 1150.

(19) G. C. Stafford Jnr., P. E. Kelly, J. E. P. Syka, W. E. Reynolds, J. F. J. Todd, Int. J. Mass Spectrom. Ion Proc., 1984, 60, 85.

(20) J. Hipple, H. Sommer and H. Thomas, Phys. Rev., 1949, 76, 1877.

(21) J. Hipple, H. Sommer and H. Thomas, Phys. Rev., 1951, 82, 697.

(22) T. A. Lehman, M. M. Bursey, "ICR Spectrometry", Wiley - Interscience, New York, New york, 1976.

(23) M. B. Comisarow and A. G. Marshall, Chem. Phys. Lett., 1974, 25, 282.

(24) D. A. Skoog and J. J. Leary, in "Principles of Instrumental Analysis", Saunders College Publishing, USA, 4th ed., 1992, Chapter 25.

(25) D. A. Skoog and J. J. Leary, in "Principles of Instrumental Analysis", Saunders College Publishing, USA, 4th ed., 1992, Chapter 26.

(26) C. M. Whitehouse, R. M. Dryer, M. Yamashita, J. B. Fenn, Anal. Chem., 1985, 57, 675.

(27) M. Mann, Org. Mass Spectrom., 1990, 25, 575.

(28) M. L. Lee and K. E. Markides, Science, 1987, 235, 1342.

(29) D. A. Skoog and J. J. Leary, in "Principles of Instrumental Analysis", Saunders College Publishing, USA, 4th ed., 1992, Chapter 27A.

(30) A. J. Dempster, Phys. Rev., 1921, 18, 415.

(31) H. Budzikiewicz, C. Djerasii and D. H. Williams, "Interpretation of Mass Spectra of Organic Compounds", Holden - Day Inc., San Francisco, California, 1964.

(32) M. S. B. Munsen and F. H. Field, J. Am. Chem. Soc., 1966, 88, 2621.

(33) K. R. Jennings, Int. J. Mass Spectrom. Ion Phys., 1968, 1, 227.

(34) H. V. Winkler and H. D. Beckey, Biochem. Biophys. Res. Commun., 1972, 46, 391.

(35) A. Benninghoven and W. K. Sichtermann, Anal. Chem., 1978, 50, 1180.

(36) R. J. Day, S. E. Unger and R. G. Cooks, Anal. Chem., 1980, 52, 557A.

(37) D. Briggs, A. Brown and J. C. Vickermann, "Handbook of Static Ion Secondary Mass Spectrometry", J. Wiley & Sons Inc., Chichester, U.K., 1989.

(38) M. S. Barber, R. S. Bordoli, R. Sedgwick and A. N. Tyler, J. Chem. Soc., Chem. Comm., 1981, 325.

(39) M. S. Barber, R. S. Bordoli, R. Sedgwick and A. N. Tyler, Nature, 1981, 293, 270.

(40) W. Aberth, K. M. Straub, A. L. Burlingame, Anal. Chem., 1982, 54, 2029.

(41) R. M. Caprioli, T. Fan and J. S. Cottrell, Anal. Chem., 1986, 58, 2949.

(42) P. G. Kistemaker, G. J. Q. van der Peryl, J. Haverkamp, "Soft Ionisation Biological Mass Spectrometry", H. E. Morris (Ed.), Heyden, London, 1981, p120-136.

(43) M. Karas, U. Bahr, A. Ingendoh and F. Hillenkamp, Angew. Chem., Int. Ed. Engl., 1989, 28, 760.

(44) D. F. Torgerson, R. P. Skowronski R. D. Macfarlane, Biochem. Biophys. Res. Commun., 1974, 60, 616.

(45) R. D. Macfarlane, "Soft Ionisation Biological Mass Spectrometry", H. E. Morris (Ed.), Heyden, London, 1981, p110-119.

(46) C. R. Blakely and M. L. Vestal, Anal. Chem., 1983, 55, 750.

(47) A. O. Nier, Rev. Sci. Instrum., 1947, 18, 398.

(48) C. E. Melton, "Mass Spectrometry of Organic Ions", F. W. McLafferty (Ed.), Academic Press Inc., New York, New York, 1963, Chapter 4.

(49) J. Wilson, Mass Spectrom. Specialist Rep., Chemical Soc., London, 1971.

(50) T. D. Mark, Int. J. Mass Spectrom. Ion Phys., 1982, 45, 125.

(51) T. D. Mark, G. H. Dunn, "Electron Impact Ionization." Springer - Verlag, New York, 1985.

(52) H. M. Rosenstock, K. Draxl, B. W. Steiner, J. T. Herron, J. Phys. Chem. Ref. Data 6, Suppl. 1, 1977, p783.

(53) R. C. Dougherty, C. R. Weisenberger, J. Am. Chem. Soc., 1968, 90, 6570.

(54) J. D. Dillard, Chem. Rev., 1973, 73, 589.

(55) J. H. Bowie, Mass Spectrom. Specialist Rep., Chemical Soc., 1975, 3, 287.

(56) J. H. Bowie and B. D. Williams, in "Int. Rev. Sci. Phys. Chem. Ser. 2", A. Maccoll (Ed.), Butterworths, London, 1975, vol 5, p 89.

(57) A. G. Harrison, "Chemical Ionisation Mass Spectrometry", CRC Press, Florida, 1983.

(58) H. Budzikiewicz, Mass Spectrom. Rev., 1986, 5, 345.

(59) J. H. J. Dawson, A. M. Kaandorp and N. M. M. Nibbering, Org. Mass Spectrom., 1977, 12, 330.

(60) H.-F. Grutzmavher, B. Grotemeyer, Org Mass Spectrom., 1984, 19, 135.

(61) S. T. Graul and R. R. Squires, Mass Spectrom. Rev., 1988, 7, 273.

(62) J. E. Bartmess, "The 1987 Gas Phase Acidity Scale", Department of Chemistry, University of Tennesee, 1987.

(63) J. B. Westmore and M. M. Alauddin, Mass Spectrom. Rev., 1986, 5, 381.

(64) A. Noest and N. M. M. Nibbering, J. Am. Chem. Soc., 1980, 102, 6427.

(65) A. L. C. Smit and F. H. Field, J. Am. Chem. Soc., 1977, 99, 6471.

(66) K. R. Jennings, Mass Spectrom. Specialist Rep, Chem. Soc. London, 1979, 5, 203.

(67) G. Caldwell and J. E. Bartmess, Org. Mass Spectrom., 1982, 17, 456.

(68) P. J. Chantry, J. Phys. Chem., 1969, 51, 3369.

(69) T. O. Tiernan, C. Chang, C. Cheng, Envir. Health Perspect., 1980, 47, 36.

(70) A. G. Harrison and R. J. Cotter, Methods Enzymol., 1990, 193, 3.

(71) H. I. Schiff, D. K. Bohme, Int. J. Mass Spectrom. Ion Phys., 1975, 16, 167.

(72) J. H. Bowie, Adv. Mass Spectrom., 1986, 10A, 553.

(73) G. Paul and P. Kebarle, J. Am. Chem. Soc., 1989, 111, 464.

(74) S. A. Sullivan, C. H. DePuy and R. Damrauer, J. Am. Chem. Soc., 1981, 103, 480.

(75) J. C. Kleingeld and N. M. M. Nibbering, Int. J. Mass Spectrom. Ion Phys., 1983, 49, 311.

(76) C. H. DePuy, R. Damrauer, J. H. Bowie J. C. Sheldon, Acc. Chem. Res., 1987, 20, 127.

(77) R. C. Lum and J. J. Grabowski, J. Am. Chem. Soc., 1988, 110, 8568.

(78) C. H. DePuy, V. M. Bierbaum, L. A. Flippin, J. J. Grabowski, G. K. Hing, R. J. Schmitt and S. A. Sullivan, J. Am. Chem. Soc., 1979, 101, 6443.

(79) R. J. Waugh, R. N. Hayes, P. C. H. Eichinger, K. M. Downard, J. H. Bowie, J. Am. Chem. Soc., 1990, 112, 2537.

(80) M. Barber, R. S. Bordoli, G. J. Elliot, R. D. Sedgwick, A. N. Tyler, Anal. Chem., 1982, 54, 645A.

(81) A. Benninghoven, W. Sichtermann, Org. Mass Spectrom., 1977, 12, 595.

(82) R. D. Macfarlane, D. F. Torgerson, Science, 1976, 191, 920.

(83) P. Hakansson, I. Kamensky and B. Sundqvist, Nucl. Instrum. Meth., 1982, 198, 43.

(84) F. Hillenkamp, Int. J. Mass Spectrom. Ion Phys., 1982, 45, 305.

(85) W. Aberth and A. L. Burlingame, Anal. Chem., 1984, 56, 2915.

(86) S. A. Martin, C. E. Costello and K. Biemann, Anal. Chem., 1982, 54, 2362.

(87) M. Barber, R. S. Bordoli, G. V. Garner, D. B. Gordon, R. D. Sedgwick, L. W. Tetler and A. N. Tyler, Biochem. J., 1981, 197, 401.

(88) M. Barber, R. S. Bordoli, R. D. Sedgwick and A. N. Tyler, Biomed. Mass Spectrom., 1981, 8, 492.

(89) M. Barber, R. S. Bordoli, G. J. Elliot, R. D. Sedgwick, A. N. Tyler, J. Chem. Soc. Faraday Trans., 1983, 679, 1249.

(90) S. S. Wong and F. W. Rollgen, Nucl. Instrum. Methods Phys. Res., 1986, B14, 436.

(91) S. Naylor, G. Moneti and S. Guyan, Biomed. Environ. Mass Spectrom., 1988, 17, 393.

(92) W. V. Ligon, S. B. Dorn, Int. J. Mass Spectrom. Ion Proc., 1984, 61, 113.

(93) W. V. Ligon, S. B. Dorn, Int. J. Mass Spectrom. Ion Proc., 1985, 63, 315.

(94) R. G. Finke, M. W. Droege, J. C. Cook K. S. Suslick, J. Am. Chem. Soc., 1984, 106, 5750.

(95) J. L. Whitten, M. H. Schaffer, M. O'Shea, J. C. Cook, M. E. Hemling and K. L. Rinehart Jr., Biochem. Biophys. Res. Commun., 1984, 124, 350.

(96) E. de Pauw, Methods Enzymol., 1990, 193, 201.

(97) C. Fenselau and R. J. Cotter, Chem. Rev., 1987, 87, 501.

(98) R. G. Cooks, K. L. Busch, Int. J. Mass Spectrom. Ion Phys., 1983, 53, 111.

(99) J. A. Sunner, R. Kulatunga and P. Kebarle, Anal. Chem., 1986, 58, 1312.

(100) P. Roepstorff and W. J. Richter, Int. J. Mass Spectrom. Ion Proc., 1992, 118/119, 789.

(101) K. B. Tomer, Mass Spectrom. Rev., 1989, 8, 445.

(102) K. L. Busch, G. L. Glish and S. A. McLuckey, "Mass Spectrometry, Mass Spectrometry: Techniques and Applications of Tandem Mass Spectrometry", VCH Publishers, New York, New York, 1988.

(103) D. F. Hunt, "Tandem Mass Spectrometry", F. W. McLafferty (Ed.), J. Wiley & Sons Inc., New York, New York, 1983, Chapter 5.

(104) K. T. Bainbridge and E. B. Jordan, Phys. Rev., 1936, 50, 282.

(105) E. G. Johnson and A. O. Nier, Phys. Rev., 1953, 91, 10.

(106) J. R. Trainer and P. J. Derrick, "Sectors and Tandem Sectors", in "Mass Spectrometry in the Biological Sciences: A Tutorial", M. L. Gross (Ed.), Kluwer Acad., USA, NATO ASI Series C, 353, 3, 1992.

(107) K. R. Jennings, Methods Enzymol., 1990, 193, 295.

(108) D. H. Williams and I. Howe, "Principles of Organic Mass Spectrometry", McGraw-Hill, London, U.K., 1972, 9.

(109) F. W. McLafferty, P. F. Bente, R. Kornfeld, S.-C. Tsai and I. Howe, J. Am. Chem. Soc., 1973, 95, 2120.

(110) M. B. Stringer, D. J. Underwood, J H. Bowie, J. L. Holmes, A. A. Mommers, J. E. Szulejko, Can. J. Chem., 1986, 64, 76.

(111) J. L. Holmes and J. K. Terlouw, Org. Mass Spectrom., 1980, 15, 383.

(112) A. Fraefel and L. Seibl, Mass Spectrom. Rev., 1985, 4, 151.

(113) H. M. Rosenstock, M. B. Wallenstein, A. L. Wahrhaflig, H. Eyring, Proc. Nat. Acad. Sci. USA, 1953, 38, 667.

(114) R. G. Cooks, J. H. Beynon, R. M. Caprioli, G. Lester, "Metastable Ions", Elsiever, Amsterdam, 1973.

(115) V. G. Instruments Model ZAB 2HF, Wythenshawe, Manchester, M23 9LE, UK.

(116) R. P. Morgan, J. H. Beynon, R. H. Bateman and B. N. Green, Int. J. Mass Spectrom. Ion Phys., 1978, 28, 171.

(117) J. H. Beynon, R. G. Cooks, J. W. Amy, W. E. Baitinger, T. Y. Ridley, Anal. Chem., 1973, 45, 1023A.

(118) V. Erspamer, P. Melchiorri, Pure and Applied Chemistry, 1973, 35, 464.

(119) R. C. Stebbins and N. W. Cohen, "A Natural History of Amphibians", Princeton University Press, Princeton, New Jersey, 1995.

(120) M. J. Tyler, personal communication.

(121) K. P. Oakley and H. M. Muir-Wood, "The Succession of Life through Geological Time", London, 1962.

(122) J. A. Tihen, Amer. Zool., 1965, 5, 263.

(123) A. S. Romer, The Vertebrate Body, Saunders, 2nd edn., Philadelphia, 1955.

(124) W. E. Duellman, Sci. Am., 1992, 267, 80.

(125) M. Neuwirth, J. W. Daly, C. W. Myers and L. W. Tice, Tissue Cell, 1979, 11, 755.

(126) J. W. Daly, C. W. Myers and N. Whittaker, Toxicon, 1987, 25, 1023.

(127) V. Erspamer, Archs. Biochem. Biophys., 1959, 82, 431.

(128) M. Roseghini, V. Erspamer and R. Endean, Comp. Biochem. Physiol., 1976, 54C, 31.

(129) V. Erspamer, Basic Appl. Histochem., 1981, 25, 3.

(130) V. Erspamer, Comp. Biochem. Physiol., 1984, 79C, 1.

(131) M. Roseghini, V. Erspamer, G. F. Erspamer, J. M. Cei, Comp. Biochem. Physiol., 1986, 85C, 139.

(132) R. Llinás and W. Precht, "Frog Neurobiology: A Handbook", Springer-Verlag, Berlin, Heidelberg, New York, 1976, p976.

(133) L. A. Blaylock, R. Ruibal and K. Platt-Aloia, Copeia, 1976, 283.

(134) H. Jensen, J. Am. Chem. Soc., 1935, 57, 1765.

(135) G. J. Dockray and C. R. Hopkins, J. Cell Biol., 1975, 64, 724.

(136) M. Ravazzola, D. Brown, J. Leppäluotto, L. Orci, Life Sci., 1979, 25, 1331.

(137) B. E. Flucher, C. Lenglachner-Bachinger, K. Pohlhammer, H. Adam and C. Mollay, J. Cell Biol., 1986, 103, 2299.

(138) J. L. Bolaffi and I. M. D. Jackson, Cell Tissue Res., 1979, 202, 505.

(139) F. Flury, Arch. Exp. Path. Pharmak., 1917, 81, 319.

(140) S. Yoshie, T. Iwanaga and T. Fujita, Cell Tissue Res., 1985, 239, 25.

(141) M. G. Giovannini, L. Poulter, B. W. Gibson, D. H. Williams, Biochem. J., 1987, 243, 113.

(142) M. J. Tyler, "Toxic Plants and animals. A guide for Australia", J. Covacevich, P. Davies and J. Pearn (Eds.), Queensland Museum, Brisbane, 1987, 329.

(143) S. Iwamuro, D. Kuwagaki and S. Kikuyama, Zool. Sci., 1991, 8, 743.

(144) T. Seki, S. Kikuyama and N. Yanaihara, Cell Tissue Res., 1989, 258, 483.

(145) F. Hauser, E.-M. Gertzen, W. Hoffmann, Expl. Cell Res., 1990, 189, 157.

(146) M. Dickerson, "The Frog Book", Doubleday, Page & Co.: New York, 1906.

(147) I. Barteczko and H. Kuziemski, Zool. Poloniae, 1970, 20, 189.

(148) M. J. Tyler, D. J. M. Stone and J. H. Bowie, J. Pharmacological Methods, 1992, 28, 199.

(149) N. C. Bols, M. M. Roberson, P. L. Haywood-Reid, S. H. Barondes and R. F. Cerra, J. Cell Biol., 1986, 102, 492.

(150) J. G. Frazer, "The Golden Bough", Macmillan: London, 1963.

(151) L. H. Lazarus and M. Attila, Progress in Neurobiology, 1993, 41, 473.

(152) O. Goldsmith, Nat. Hist., 1790.

(153) J. W Daly, J. Caceres, R. W. Moni, F. Gusovsky, M. Moos Jr., K. B. Seamon, K. Milton and C. W. Myers, Proc. Natl. Acad. Sci. USA, 1992, 89, 10960.

(154) T. H. Johnston, Trans. Roy. Soc. S.A., 1943, 67, 244.

(155) S. Grenard, "Medical Herpetology", Reptile and Amphibian Magazine, Pottsville, PA, USA, 1994.

(156) M. J. Tyler, Australian Natural History, 1995, 24(12), 46.

(157) H. Weiland, W. Konz and H. Mittasch, Justus Liebigs Annalen Chem., 1934, 513, 1.

(158) V. Erspamer, Annu. Rev. Pharmacol., 1971, 11, 327.

(159) V. Erspamer, L. Negri, G. Falconieri Erspamer and R. Endean, Naunyn-Schmiedeberg's Arch. Pharmacology, 1975, 289, 41.

(160) G. Bertaccini, Pharmacol. Rev., 1976, 28 (2), 127.

(161) V. Erspamer, P. Melchiorri, "Growth Hormone and other Biologically Active Peptides", A. Pecile and E. E. Muller (Eds.), Amsterdam: Excerpta Medica, 1980, 185.

(162) V. Erspamer and P. Melchiorri, Trends Pharmacol. Sci., 1980, 1, 391.

(163) V. Mutt, Peptides, 1981, 2 (Suppl. 2), 3.

(164) V. Erspamer, P. Melchiorri, M. Broccardo, G. F. Erspamer and P. Falaschi, Peptides, 1981, 2 (Suppl 2), 7.

(165) T. Nakajima, Trends Pharmacol. Sci., 1981, 2, 202.

(166) G. G. Habermehl, "Venomous Animals and their Toxins", Springer-Verlag, Berlin, 113, 1981.

(167) V. Erspamer, G. Falconieri Erspamer, G. Mazzanti and R. Endean, Comp. Biochem. Biophys., 1984, 79C, 99.

(168) V. Erspamer, G. Falconieri Erspamer, P. Melchiorri and G. Mazzanti, Neuropharmacology, 1985, 24, 783.

(169) P. Melchiorri, Peptides, 1985, 6 (Suppl 3), 3.

(170) V. Erspamer, G. F. Erspamer and J. M. Cei, Comp. Biochem. Physiol., 1986, 85C, 125.

(171) T. Renda, L. D`Este, A. Fasolo, L. H. Lazarus, F. Minniti and V. Erspamer, Arch. Histol.C ytol., 1989, 52 (Suppl.), 317.

(172) M. Roseghini, G. Falconieri Erspamer, C. Severini and M. Simmaco, Comp. Biochem. Physiol., 1989, 94C, 455.

(173) V. Erspamer, P. Melchiorri, G. Falconieri Erspamer, L. Negri, R. Corsi, C. Severini, D. Barra, M. Simmaco and G. Kriel, Proc. Natl. Acad. Sci. USA, 1989, 86, 5188.

(174) C. L. Bevins and M. Zasloff, Annu. Rev. Biochem., 1990, 59, 395.

(175) V. Erspamer, Int. J. dev Neurosci., 1992, 10, 3.

(176) K. S. Moore, C. L. Bevins, M. M. Brasseur, N. Tomassini, K. Turner, H. Eck and M. Zasloff, J. Biol. Chem., 1991, 266, 19851.

(177) A. C. Andersen, M.-C. Tonon, G. Pelletier, J. M. Conlon, A. Fasolo and H. Vaudry, Int. Rev. Cytology, 1992, 138, 89.

(178) A. C. Andersen, M.-C. Tonon, G. Pelletier, J. M. Conlon, A. Fasolo and H. Vaudry, Int. Rev. Cytology, 1992, 138, 315.

(179) M. Otsuka and K. Yoshioka, Physiol. Rev., 1993, 73(2), 229.

(180) M. Zasloff, Proc. Natl. Acad. Sci. USA, 1987, 84, 5449.

(181) M. Zasloff, B. Martin and H.-C. Chen, Proc. Natl. Acad. Sci. USA, 1988, 85, 910.

(182) J. M. Cei, V. Erspamer and M. Roseghini, Syst. Zool., 1967, 16, 328.

(183) J. M. Cei, V. Erspamer and M. Roseghini, Syst. Zool., 1968, 17, 232.

(184) J. M. Cei, Peptides, 1985, 6, 13.

(185) M. J. Tyler, Alytes, 1991, 9, 43.

(186) M. J. Tyler, Aust. Nat. Hist., 1991, 23(8), 618.

(187) T. Nakajima, T. Yasuhara, V. Erspamer, G. F. Erspamer, L. Negri and R. Endean, Chem. Pharm. Bull., 1980, 28, 689.

(188) V. Erspamer, P. Melchiorri, T. Nakajima, T. Yasuhara and R. Endean, Experientia, 1979, 35, 1132.

(189) A. Anastasi, V. Erspamer and R. Endean, Experientia, 1975, 31, 510.

(190) A. Csordás and H. Michl, Monatshefte f. Chem., 1970, 101, 182.

(191) T. Yasuhara, O. Ishikawa, T. Nakajima, K. Araki, S. Tochibana, Chem. Pharm. Bull., 1979, 27(2), 486.

(192) D. J. M. Stone, R. J. Waugh, J. H. Bowie, J. C. Wallace and M. J. Tyler, J. Chem. Soc. Perkin Trans. 1, 1992, 3173.

(193) A. Anastasi, V. Erspamer and R. Endean, Archs. Biochem. Biophys., 1968, 125, 57.

(194) A. M. Bradford, M. J. Raftery, J. H. Bowie, J. C. Wallace and M. J. Tyler, Aust. J. Chem., 1993, 46, 1235.

(195) R. Horikawa, D. S. Parker, P. L. Herring and J. J. Pisano, Fed. Proc., 1972, 44, 695.

(196) V. Erspamer, G. Bertaccini and J. M. Cei, Experientia, 1962, 18, 562.

(197) S. T. Steinborner, C. Gao, M. J. Raftery, R. J. Waugh, T. Blumenthal, J. H. Bowie, J. C. Wallace and M. J. Tyler, Aust. J. Chem., 1994, 47, 2099.

(198) K. Araki, S. Tachibana, M. Uchiyama, T. Nakajima, T. Yasuhara, Chem. Pharmac. Bull., 1973, 21, 2801.

(199) B. W. Gibson, D. Tang, R. Mandrell, M. Kelly and E. R. Spindell, J. Biol. Chem., 1991, 266, 23103.

(200) E. T. Kaiser and F. J. Kézdy, Proc. Natl. Acad. Sci. USA, 1983, 80, 1137.

(201) E. T. Kaiser and F. J. Kézdy, Science, 1984, 223, 249.

(202) A. Brack, Biosystems, 1977, 9, 99.

(203) M. Mutter, Angew. Chem. Int. Ed. Engl., 1985, 24, 639.

(204) D. G. Osterman and E. T. Kaiser, J. Cell. Biochem., 1985, 29, 57.

(205) D. G. Osterman, R. Mora, F. J. Kézdy, E. T. Kaiser and S. C. Meredith, J. Am. Chem. Soc., 1984, 106, 6845.

(206) W. F. DeGrado and J. D. Lear, J. Am. Chem. Soc., 1985, 107, 7684.

(207) E. T. Kaiser and F. J. Kézdy, Ann. Rev. Biophys. Chem., 1987, 16, 561.

(208) D. Fukushima, J. P. Kupferberg, S. Yokoyama, D. J. Kroon, E. T. Kaiser and F. J. Kézdy, J. Am. Chem. Soc., 1979, 101, 3703.

(209) W. F. DeGrado, F. J. Kézdy, E. T. Kaiser, J. Am. Chem. Soc., 1981, 103, 679.

(210) D. Fukushima, S. Yokoyama, F. J. Kézdy and E. T. Kaiser, Proc. Natl. Acad. Sci. USA, 1981, 78, 2732.

(211) D. J. Kroon, J. P. Kupferberg, E. T. Kaiser and F. J. Kézdy, J. Am. Chem. Soc., 1978, 100, 5975.

(212) S. Y. M. Lau, A. K. Taneja, R. S. Hodges, J. Biol. Chem., 1984, 259, 13253.

(213) J. A. Carver and J. H. Bowie, unpublished results.

(214) P. Kanellis, A. Y. Romans, B. J. Johnson, H. Kercret, R. Chiovetti Jr., T. M. Allen and J. P. Segrest, J. Biol. Chem., 1980, 225, 11464.

(215) P. Dunnil, Biophys. J., 1968, 8, 865.

(216) M. Schiffer and A. B. Edmundson, Biophys. J., 1967, 7, 121.

(217) Y. Nozaki and C. Tanford, J. Biol. Chem., 1971, 246, 2211.

(218) H. B. Bull and K. Breese, Arch. Biochem. Biophys., 1974, 161, 665.

(219) J. P. Segrest and R. J. Feldman, J. Mol. Biol., 1974, 87, 853.

(220) C. Cothia, J. Mol. Biol., 1976, 105, 1.

(221) G. von Heijne and C. Blomberg, Eur. J. Biochem., 1979, 97, 175.

(222) J. Janin, Nature, 1979, 277, 491.

(223) R. Wolfenden, L. Andersson, P. M. Cullis and C. C. B. Southgate, Biochemistry, 1981, 20, 849.

(224) P. Argos, J. K. M. Rao and P. A. Hargrave, Eur. J. Biochem., 1982, 128, 565.

(225) D. Eisenberg, R. M. Weiss, T. C. Terwilliger, W. Wilcox, Faraday Symp. Chem. Soc., 1982, 17, 109.

(226) J. Kyte and R. F. Doolittle, J. Mol. Biol., 1982, 157, 105.

(227) R. M. Sweet and D. Eisenberg, J. Mol. Biol., 1983, 171, 479.

(228) H. V. Westerhoff, D. Juretic, R. W. Hendler, M. Zasloff, Proc. Natl. Acad. Sci. USA, 1989, 86, 6597.

(229) B. Christensen, J. Fink, R. B. Merrifield, D. Mauzerall, Proc. Natl. Acad. Sci. USA, 1988, 85, 5072.

(230) H. Duclohier, G. Molle and G. Spach, Biophys. J., 1989, 56, 1017.

(231) D. Wade, A. Boman, B. Wåhlin, C. M. Drain, D. Andreu, H. G. Boman and R. B. Merrifield, Proc. Natl. Acad. Sci. USA, 1990, 87, 4761.

(232) J. H. Bowie, unpublished results.

(233) W. L. Maloy, U. Prasad Kari, Biopolymers (Peptide Science), 1995, 37, 105.

(234) B. W. Gibson, L. Poulter and D. H. Williams, Peptides, 1985, 6, 23.

(235) J. P. Glusker, M. Lewis and M. Rossi, "Crystal Structure Analysis for Chemists and Biologists", V. C. H., New York, 1994.

(236) H. J. Dyson and P. E. Wright, Ann. Rev. Biophys. Biochem. Acta., 1991, 20, 519.

(237) M. P. Williamson and J. P. Waltho, Chem. Soc. Rev., 1992, 21, 227.

(238) A. M. Maxam and W. Gilbert, Proc. Natl. Acad. Sci USA, 1977, 74, 560.

(239) F. Sanger, A. R. Coulson, B. G. Barrel, A. J. H. Smith, B. A. Roe, J. Mol. Biol., 1980, 143, 161.

(240) R. M. Hewick, M. W. Hunkapiller, L. E. Wood and W. J. Dryer, J. Biol. Chem., 1981, 256, 7990.

(241) K. Biemann, Anal. Chem., 1986, 58, 1288A.

(242) R. Uy and F. Wold, Science, 1977, 198, 890.

(243) F. Wold, Annu. Rev. Biochem., 1981, 50, 783.

(244) "Post-translational Modifications" in "Methods in Enzymology", F. Wold and K. Moldave Eds., Academic Press, Orlando, Florida, 1984, 106, 107.

(245) R. Kornfeld and S. Kornfeld, Annu. Rev. Biochem., 1985, 54, 631.

(246) K. Verner and G. Schatz, Science, 1988, 241, 1307.

(247) G. Allen, "Sequencing of Proteins and Peptides", in "Laboratory, Techniques in Biochemistry and Molecular Biology", T. S. Work and R. H. Burdon, (Eds.), Elsevier/North-Holland Biomedical Press, Amsterdam, The Netherlands, 1981, 9.

(248) S. Tsunasawa and F. Sakiyama, Methods Enzymol., 1984, 106, 165.

(249) B. W. Gibson and K. Biemann, Proc. Natl. Acad. Sci. USA, 1984, 81, 1956.

(250) S. L. Ramsay, S. T. Steinborner, R. J. Waugh, S. Dua, J. H. Bowie, Rapid Commun. Mass Spectrom., 1995, 9, 1241.

(251) P. Schlack, W. Kumpf, Hoppe-Seylers Z. Physiol. Chem., 1926, 154, 125.

(252) F. E. Dwulet and F. R. N. Gurd, Int. J. Peptide Protein Res., 1979, 13, 122.

(253) L. Meuth, D. E. Harris, F. E. Dwulet, M. L. Crowl-Powers, F. R. N. Gurd, Biochemistry, 1982, 21, 3751.

(254) B. T. Chait, R. Wang, R. C. Beavis, S. B. H. Kent, Science, 1993, 262, 89.

(255) M. Bartlet-Jones, W. A. Jeffrey, H. F. Hansen and D. J. C. Pappin, Rapid Commun. Mass Spectrom., 1994, 8, 737.

(256) R. M. Caprioli, Mass Spectrom. Rev., 1987, 6, 237.

(257) K. Biemann, in "Methods in Protein Sequence Analysis", M. Elzinga, (Ed.), Humana Press, Clifton, N.J., 1982, 279.

(258) H. R. Morris, M. Panico and G. W. Taylor, Biochem. Biophys. Res. Commun., 1983, 117, 299.

(259) H. R. Morris and F. M. Greer, TIBTECH, 1988, 6, 140.

(260) R. S. Johnson and K. Biemann, Biochemistry, 1987, 26, 1209.

(261) K. Biemann, Biomed. Envir. Mass Spectrom., 1988, 16, 99.

(262) R. S. Johnson, S. A. Martin, K. Biemann, Int. J. Mass Spectrom. Ion Proc., 1988, 86, 137.

(263) P. T. M. Kenny, K. Nomoto and R. Orlando, Rapid Commun. Mass Spectrom., 1992, 6, 95.

(264) T. Yalcin, I. G. Csizmadia, M. R. Peterson and A. G. Harrison, J. Am. Soc. Mass Spectrom., 1996, 7, 233.

(265) T. Yalcin, C. Khouw, I. G. Csizmadia, M. R. Peterson and A. G. Harrison, J. Am. Soc. Mass Spectrom., 1995, 6, 1165.

(266) P. Roepstorff and J. Fohlman, Biomed. Mass Spectom., 1984, 11, 601.

(267) K. Biemann, Methods Enzymol., 1990, 193, 455.

(268) R. S. Johnson, S. A. Martin, K. Biemann, J. T. Stults, J. T. Watson, Anal. Chem., 1987, 59, 2621.

(269) N. L. Squire, S. Beronová and C. Wesdemiotis, J. Mass Spectrom., 1995, 30, 1429.

(270) N. Goldberg and H. Schwarz, Acc. Chem. Res., 1994, 27, 347.

(271) K. D. Ballard, S. J. Gaskell, Int. J. Mass Spectrom. Ion Proc., 1991, 111, 173.

(272) K. Biemann, Methods Enzymol., 1990, 193, 886.

(273) H. A. Scoble and S. A Martin, Methods Enzymol., 1990, 193, 519.

(274) G. Bertaccini, Arch. für Pharmakologie, 1971, 269 (2-4), 139.

(275) A. Anastasi, V. Erspamer and R. Endean, Experientia, 1975, 31, 394.

(276) T. Braibanti, G. Bertaccini, F. Uva, Ann. Radiol. Diagnost., 1968, 41, 516.

(277) G. Bertaccini, "Pharmacology and Clinical use of Caerulein", in "Symposium on Gastrin and its Antagonists", J. Borsy and G. Mozsik (Eds.), Akadémiai Kiadò, Budapest, 1973, 47.

(278) A. Anastasi, G. Bertaccini, V. Erspamer, M. Impicciatore, M. Roseghini, J. M. Cei and G. De Caro, Br. J. Pharmac., 1970, 38, 221.

(279) M. J. Tyler, "Australian Frogs", Australian Natural History, Reed, Sydney, 1994.

(280) D. J. M. Stone, J. H. Bowie, M. J. Tyler, J. C. Wallace, J. Chem. Soc. Chem. Commun., 1992, 1224.

(281) D. J. M. Stone, R. J. Waugh, J. H. Bowie, J. C. Wallace and M. J. Tyler, J. Chem. Research (S), 1993, 138.

(282) R. J. Waugh, D. J. M. Stone, J. H. Bowie, J. C. Wallace and M. J. Tyler, J. Chem. Research (S), 1993, 139.

(283) R. J. Waugh, D. J. M. Stone, J. H. Bowie, J. C. Wallace and M. J. Tyler, J. Chem. Soc., Perkin Trans. 1, 1993, 573.

(284) R. J. Waugh, S. T. Steinborner, J. H. Bowie, J. C. Wallace, M. J. Tyler, P. Hu and M. L. Gross, Aust. J. Chem., 1995, 48, 1981.

(285) S. T. Steinborner, P. A. Wabnitz, R. J. Waugh, J. H. Bowie, C. Gao, M. J. Tyler and J. C. Wallace, Aust. J. Chem., 1996, 49, 92.

(286) M. J. Raftery, R. J. Waugh, J. H. Bowie, J. C. Wallace and M. J. Tyler, J. Peptide Science, 1996, 2(2), 117.

(287) S. T. Steinborner, P. A. Wabnitz, J. H. Bowie and M. J. Tyler, Rapid Commun. Mass Spectrom., 1996, 10, 92.

(288) S. T. Steinborner, R. J. Waugh, J. H. Bowie, J. C. Wallace, M. J. Tyler and S. L. Ramsay, in press

(289) M. J. Raftery, A. M. Bradford, J. H. Bowie, J. C. Wallace, M. J. Tyler, Aust. J. Chem., 1993, 46, 833.

(290) L. Bernadi, G. Bosisio, F. Chillemi, G. de Caro, R. de Castiglione, V. Erspamer, A. Glässer and O. Goffredo, Experentia (Basel), 1964, 20, 306.

(291) G. de Caro and L. Farruggia, Arch. Int. Pharmacodyn. Thér., 1966, 160, 44.

(292) G. de Caro, Arch. Int. Pharmacodyn. Thér., 1966, 162, 437.

(293) G. de Caro, L. Farrugia, E. Minardi, L. Novarini, Naunyn-Schmiedebergs Arch. Pharmakol. Exp. Pathol., 1966, 254, 194.

(294) C. B. Fregnan, A. H. Glässer, C. Beretta, Eur. J. Pharmacol., 1968, 4, 421.

(295) J. Bergmann, M. Bienert, H. Niedrich, B. Mehlis, P. Oehme, Experentia (Basel), 1974, 30, 401.

(296) J. Bergmann, P. Oehme, M. Bienert and H. Niedrich, Experentia (Basel), 1974, 30, 1315.

(297) J. Bergmann, P. Oehme, M. Bienert, H. Niedrich, Acta Biol. Med. Germ., 1975, 34, 475.

(298) H. Niedrich, M. Bienert, B. Mehlis, J. Bergmann, P. Oehme, Acta Biol. Med. Germ., 1975, 34, 483.

(299) P. Ormas, G. Pompa, C. Beretta, R. Faustini, Folia Veter. Lat., 1977, 7, 252.

(300) M. Impicciatore, G. Maraimi and G. Bertaccini, Naunyn-Schmiedebergs Arch. Pharmakol. Exp. Pathol., 1973, 279, 127.

(301) G. B. Bietti, P. Capra, G. de Caro, Bericht Ophthalm. Ges., 1974, 73, 399.

(302) G. de Caro and M. Cordella, Ann. Oftalm. Clin. Ocul., 1976, 91, 933.

(303) A. G. E. Pearse, Nature (London), 1976, 262, 92.

(304) T. Höckfelt, A. Ljungdahl, L. Terenius, R. Elde, G. Nilsson, Proc. Natl. Acad. Sci. U.S.A., 1977, 74, 3081.

(305) T. Höckfelt, O. Johansson, A. Ljungdahl, J. M. Lundberg, M. Schultzberg, Nature (London), 1980, 284, 515.

(306) A. I. Basbaum and H. L. Fields, Annu. Rev. Neurosci., 1984, 7, 309.

(307) A. S. Dutta, in "Comprehensive Medicinal Chemistry. The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds." J. C. Emmett (Ed.), Pergamon Press, 1990, 1001.

(308) S. Konishi and M. Otsuka, Brain Res., 1974, 65, 397.

(309) R. A. Nicoll, J. Pharmacol. Exp. Ther., 1978, 207, 817.

(310) T. M. Materson and A. Anderson, The Complete Works of William Shakespeare, New York: World Syndicate, 1926.

(311) J. E. Gray, Ann. Mag. Nat. Hist., 1841, 7, 86.

(312) H. W. Parker, Novit Zool., 1940, 42, 1.

(313) A. R. Main, "Frogs of southern Western Australia", West. Aust. Nat. Club, Perth Handb. No. 8.

(314) J. A. Moore, Bull. Am. Mus. Nat. Hist., 1961, 121, 155.

(315) A. Loveridge, Occas. Pap. Boston Soc. Nat. Hist., 1933, 8, 89.

(316) I. R. Straughan, Ph. D. Thesis, University of Queensland, 1966.

(317) M. J. Littlejohn, in "Australian Inland Waters and their Fauna", A. H. Weatherley (Ed.), Australian National University Press, Canberra, Chapter 6, 1967.

(318) J. D. Lynch, Univ. Kans. Mus. Nat. Hist. Misc. Publ., 1971, 5, 1.

(319) W. R. Heyer and D. S. Liem, Smithson. Contrib., 1976, 233, 1.

(320) M. Simmaco, C. Severini, D. de Baise, D. Barra, F. Bossa, J. D. Roberts, P. Melchiorri and V. Erspamer, Peptides,, 1990 11, 299.

(321) M. J. Tyler, M. Davies, A. A. Martin, Aust. J. Zool., Sup. Ser., 1981, 79, 1.

(322) M. J. Tyler, "Encyclopedia of Australian Animals: Frogs", The National Photographic Index of Australian Wildlife, The Australian Museum, R. Strahan (Ed.), Angus and Robertson, 1992.

(323) M. J. Raftery, J. H. Bowie and M. J. Tyler,unpublished observations.

(324) R. de Castiglione, F. Faoro, G. Perseo, S. Piana, Int. J. Pept. Protein Res., 1981, 17, 263.

(325) P. C. Montecucchi, R. de Castiglione, S. Piana, L. Gozzini, V. Erspamer, Int. J. Pept. Protein Res., 1981, 17, 275.

(326) P. C. Montecucchi, R. de Castiglione, V. Erspamer, Int. J. Pept. Protein Res., 1981, 17, 316.

(327) G. Kreil, D. Barra, M. Simmaco, V. Erspamer, G. F. Erspamer, L. Negri, G. Severini, R. Corsi and P. Melchiorri, Eur. J. Pharm., 1989, 162, 123.

(328) M. W. Hunkapiller, R. M. Hewick, W. J. Drewer, L. E. Hood, Methods Enzymol., 1983, 91, 399.

(329) K. Biemann and S. A. Martin, Mass Spectrom. Revs., 1987, 6, 1.

(330) J. Jacob and M. Zasloff, "Therapeutic Activity of the Magainins", H. G. Boman, J. Marsh and J. A. Goode (Eds.), in Antimicrobial Peptides, Ciba Symposium 186, J. Wiley and Sons, Chichester, New York, Brisbane, Toronto and Singapore, 1995, p 197-223.

(331) W. R. Pearson, D. J. Lipman, Proc. Natl. Acad. Sci. USA,, 1988 85, 2444.

(332) J. M. Park, J.-E. Jung and B. J. Lee, Biochem. Biophys. Res. Commun., 1994, 205, 948.

(333) Instituto di Farmacologia Medica, Universita Delgi Studi di Roma, 'La Sapienza', 00185, Roma, Italy.

(334) L. G. Andersson, "Results of Dr. E. Mjöberg's Swedish Scientific Expeditions to Australia 1910-1913". IV. Batrachians. K. Svenska Vetenskapsakad. Handl., 1913, 52(4), 1.

(335) A. M. Bradford, M. J. Raftery, J. H. Bowie, M. J. Tyler, J. C. Wallace, G. W. Adams and C. Severini, Aust. J. Chem., 1996, 49, 475.

(336) V. Erspamer, in "Amphibian Biology", H. Heatwole (Ed.), Surrey Beatty and Sons, N.S.W., Australia, 1994, p178-350.

(337) Sir J. J. Thomson, in a discussion on "Isotopes". Proc. Roy. Soc., 1921, 99, 87.

(338) J. H. Bowie, Mass Spectrom. Rev., 1990, 9, 349.

(339) M. J. Raftery, J. H. Bowie, Int. J. Mass Spectrom. Ion Proc., 1988, 85, 167.

(340) G. J. Currie, J. H. Bowie, R. A. Massy-Westropp, G. W. Adams, J. Chem. Soc. Perkin Trans. 2, 1988, 403.

(341) P. C. H. Eichinger and J. H. Bowie, Int. J. Mass Spectrom. Ion Proc., 1991, 110, 123.

(342) M. Eckersley, J. H. Bowie, R. N. Hayes, Int. J. Mass Spectrom. Ion Proc., 1989, 93, 199.

(343) P. C. H. Eichinger, J. H. Bowie and R. N. Hayes, J. Am. Chem. Soc., 1989, 111, 4224.

(344) P. C. H. Eichinger and J. H. Bowie, J. Org. Chem., 1986, 51, 5078.

(345) P. C. H. Eichinger, J. H. Bowie, R. N. Hayes, J. Org. Chem., 1987, 52, 5224.

(346) R. J. Waugh, P. C. H. Eichinger, J. H. Bowie and R. N. Hayes, Int. J. Mass Spectrom. Ion Proc., 1990, 96, 347.

(347) J. Adams, Mass Spectrom. Rev., 1990, 9, 141.

(348) M. L. Gross, Int. J. Mass Spectrom. Ion Proc., 1992, 118/119, 137.

(349) N. J. Jensen, K. B. Tomer, M. L. Gross, J. Am. Chem. Soc., 1985, 107, 1863.

(350) M. M. Cordero and C. Westemiotis, Anal. Chem., 1994, 66, 861.

(351) J. H. Bowie, "Mass Spectrometry Specialist Reports", Chemical Society, London, 1984, 7, 151; 1985, 8, 161; 1987, 9, 172; 1989, 10, 145.

(352) J. H. Bowie, in 'Topics in Mass Spectrometry', vol 1, "Experimental Mass Spectrometry", D. H. Russel (Ed.), Plenum Publ. Corp., New York & London, 1994, p 1 - 38.

(353) W. Kulik and W. Heerma, Biomed. Envir. Mass Spectrom., 1988, 17, 173.

(354) W. Kulik and W. Heerma, Biomed. Envir. Mass Spectrom., 1989, 18, 910.

(355) W. Kulik and W. Heerma, Biomed. Envir. Mass Spectrom., 1991, 20, 553.

(356) H. Tsunematsu, H. Hanazono, K. Horie, T. Fukuda, M. Yamamoto, Org. Mass Spectrom., 1994, 2, 197.

(357) D. Voigt and J. Schmidt, Biomed. Mass Spectrom., 1978, 5, 44.

(358) C. V. Bradley, I. Howe and J. H. Beynon, J. Chem. Soc. Chem. Commun., 1980, 562.

(359) C. V. Bradley, I. Howe, J. H. Beynon, Biomed. Mass Spectrom., 1981, 8, 85.

(360) A. M. Buko, L. R. Phillips and B. A. Fraser, Biomed. Mass Spectrom., 1983, 10, 387.

(361) I. Katakuse, D. M. Desiderio, Int. J. Mass Spectrom. Ion Proc., 1983, 54, 1.

(362) W. Kulik and W. Heerma, Biomed. Envir. Mass Spectrom., 1986, 15, 419.

(363) M. Eckersley, J. H. Bowie and R. N. Hayes, Org. Mass Spectrom., 1989, 24, 597.

(364) R. J. Waugh, M. Eckersley, J. H. Bowie and R. N. Hayes, Int. J. Mass Spectrom. Ion Proc., 1990, 98, 135.

(365) R. J. Waugh, J. H. Bowie and R. N. Hayes, Int. J. Mass Spectrom. Ion Proc., 1991, 107, 333.

(366) R. J. Waugh, J. H. Bowie, R. N. Hayes, Org. Mass Spectrom., 1991, 26, 250.

(367) R. J. Waugh, J. H. Bowie and M. L. Gross, Aust. J. Chem., 1993, 46, 693.

(368) R. J. Waugh, J. H. Bowie, M. L. Gross, Rapid Commun. Mass Spectrom., 1993, 7, 623.

(369) R. J. Waugh, J. H. Bowie, D. Vollmer, M. L. Gross, Int. J. Mass Spectrom. Ion Proc., 1994, 133, 165.

(370) R. J. Waugh, J. H. Bowie, Rapid Commun. Mass Spectrom., 1994, 8, 169.

(371) M. B. Stringer, J. H. Bowie, P. C. H. Eichinger, G. J. Currie, J. Chem. Soc., Perkin Trans. 2, 1987, 387.

(372) J. J. Grabowski and X. Cheng, J. Am. Chem. Soc., 1989, 111, 3106.

(373) R. A. O'Hair, S. Gronert, C. H. DePuy and J. H. Bowie, J. Am. Chem. Soc., 1989, 111, 3105.

(374) M. Fujio, R. T. McIver and R. W. Taft, J. Am. Chem. Soc., 1981, 103, 4017.

(375) E. M. Marzluff, S. Campbell, M. T. Rodgers and J. L. Beauchamp, J. Am. Chem. Soc., 1994, 116, 7787.

(376) S. G. Lias, J. E. Bartmess, J. F. Liebman, J. L. Holmes, R. D. Levin and W. G. Mallard, Gas Phase Ion Neutral Thermochem. J. Phys. Chem. Ref. Data, 17, Suppl. 1, 1988.

(377) S. T. Steinborner and J. H. Bowie, submitted for publication.

(378) D. H. Williams, C. V. Bradley, S. Santikarn and G. Bojesen, Biochem. J., 1982, 201, 105.

(379) P. C. Burgers, J. L. Holmes, J. E. Szulejko, A. A. Mommers, J. K. Terlouw, Org. Mass Spectrom., 1983, 18, 254.

(380) C. V. Bradley, D. H. Williams and M. R. Hanley, Biochem. Biophys. Res. Commun., 1982, 104, 1223.

(381) Chiron Mimotopes, Clayton, Victoria, Australia.

(382) N. J. Maeji, A. M. Bray, R. M. Valerio and W. Wang, Peptide Research, 1995, 8, 33.

(383) M. J. Locke, R. T. McIver Jr., J. Am. Chem. Soc, 1983., 105, 4226.

(384) M. Moet-Ner, J. Am. Chem. Soc., 1988, 110, 3071.

PUBLICATIONS

A.M. Bradford, R.J. Waugh and J.H. Bowie, "Characterization of Underivatized Tetrapeptides by Negative-ion fast-atom Bombardment Mass Spectrometry", *Rapid Commun. Mass Spectrom.*, 1995, **9**, 677. Erratum 1995, **11**, 1083.

A.M. Bradford, M.J. Raftery, J.H. Bowie, M.J. Tyler, J.C. Wallace, G.W. Adams, C. Severini, "Novel Uperin Peptides from the Dorsal Glands of the Australian Floodplain Toadlet *Uperoleia inundata*", *Aust. J. Chem.*, 1996, **49**, 475.

A.M. Bradford, J.H. Bowie, M.J. Tyler and J.C. Wallace, "New Antibiotic Uperin Petides from the Dorsal Glands of the Australian Toadlet *Uperoleia mjobergi*", *in press*.

Bradford, A.M., Waugh, R.J. & Bowie, J.H. (1995) Characterization of underivatized tetrapeptides by negative-ion fast-atom bombardment mass spectrometry.
Rapid Communications in Mass Spectrometry, v. 9(8), pp. 677-685

Bradford, A.M., Raftery, M.J., Bowie, J.H., Tyler, M.J., Wallace, J.C., Adams, G.W. & Severini, C. (1996) Novel uperin peptides from the dorsal glands of the Australian floodplain toadlet *Uperoleia inundata.*
Australian Journal of Chemistry, v. 49(4), pp. 475-484