Collision-induced dissociation of 4-imidazolone ions, isomers of $[b_n]^+$ derived from peptides after water loss

Ka Hang Brian Lam

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN CHEMISTRY

YORK UNIVERSITY

TORONTO, ONTARIO

April 2018

Abstract

Collision-induced dissociation spectra of protonated peptides $[(\text{aminoacid})_n + H]^+$ form, after the loss of water, nominal $[b_n]^+$ ions which are dependent on both peptide length and amino acid sequence. Protonated tetraglycine loses water predominantly from the first amide; loss from the second amide is a minor channel. The resultant $[b_4]^+$ ions have identical CID spectra, indicating interconversion prior to dissociation. Protonated pentaglycine loses water from the first and second amides in high abundances, and the resulting CID spectra of the $[b_5]^+$ ions are different, suggesting that interconversion is no longer the dominant pathway. The $[b_6]^+$ ions of hexaglycine, formed from the first three amides are in high abundance and their CID spectra are not identical. The substitution of an alanine or proline for a glycine of tetraglycine generally showed reduced water loss and the $[b_4]^+$ ions had different fragmentation pathways than those of the $[b_4]^+$ ions of tetraglycine.

Acknowledgements

I would like to thank my supervisors Alan C. Hopkinson and K.W. Michael Siu for their support, for which I am eternally grateful. Professor Hopkinson has been instrumental in teaching me about gas phase fragmentation by collision-induced dissociation. I have learned a lot about mass spectrometry as well as life. Professor K.W. Michael Siu has taught me the stringent research requirements and to become a solid researcher. His connections and massive network of former students has helped me and will surely help me in the future. Although this do not do justice for my gratitude towards them, I hope they know I am truly grateful.

Dr. Justin Kai-Chi Lau is the main research collaborator for these projects. He performed all the Density Functional Theory calculations which provided all the energies for the various structures and transition states. Without his never-ending commitment to the project and help in this project. The work would never have been finished.

There are many research associates who were also integral to my learning experience. Dr. Udo H. Verkerk, John Van Nostrand and Dr. Cheuk Kuen Isaac Lai taught me all the fundamentals of mass spectrometry, from basic usage to advanced knowledge. They taught me about the instruments, far more than what a regular user would know. Under their supervision, I was given the opportunity disassemble and reassemble everything from an orbitrap to the 2000 Qtrap. I am certain that this knowledge will always be valuable.

My roommate Master Bai-Han Backen Wu was the one who taught me a lot when I first joined the group. He paved the road for my success, I am truly grateful. I would like to extend my gratitude towards my group members Dr. Yating Wang, Dr. Olena Masui, and Nicole Chevannes-McGregor. I am also indebted to Dr. Larry C. Campbell and Dr. Yves Leblanc who were my supervisors at SCIEX. Without them I would not have gained as much insight into mass spectrometry within an industrial setting.

Finally, I would like to thank my friends and family for their constant support. Without their never-ending love, I would not be here today.

I also graciously acknowledge NSERC funding and the CREATE program as well as SCIEX. Without their generosity, I would not have been granted so many wonderful opportunities.

Figure List

Scheme List

Table of Contents

Chapter 1 Introduction

1.1 Applications of Mass Spectrometry in Bottom-up Proteomics

The development of soft ionization techniques such as electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) allowed for a variety of applications in biology including the development of proteomics.[1,2] The word proteomics was coined from merging "protein" and "genomics" in the 1990s.[3,4] The goal of proteomics is to identify and quantify all proteins of a proteome including expression, cellular localization, interactions, post-translational modifications (PTMs) and turnover as a function of time, space, and cell type.[5]

One form of proteomics is bottom-up or "shotgun" proteomics. It is where information about the protein is derived from its individual peptides.[6] A basic proteomics workflow is shown in Figure 1.1. In this method proteins are measured indirectly through peptides derived from proteolytic digestion. The most common protease used is trypsin, which cleaves proteins at the carboxyl terminal of an arginine or lysine residue, except when either residue is followed by a proline.[7] The sample is then ready for separation by liquid chromatography

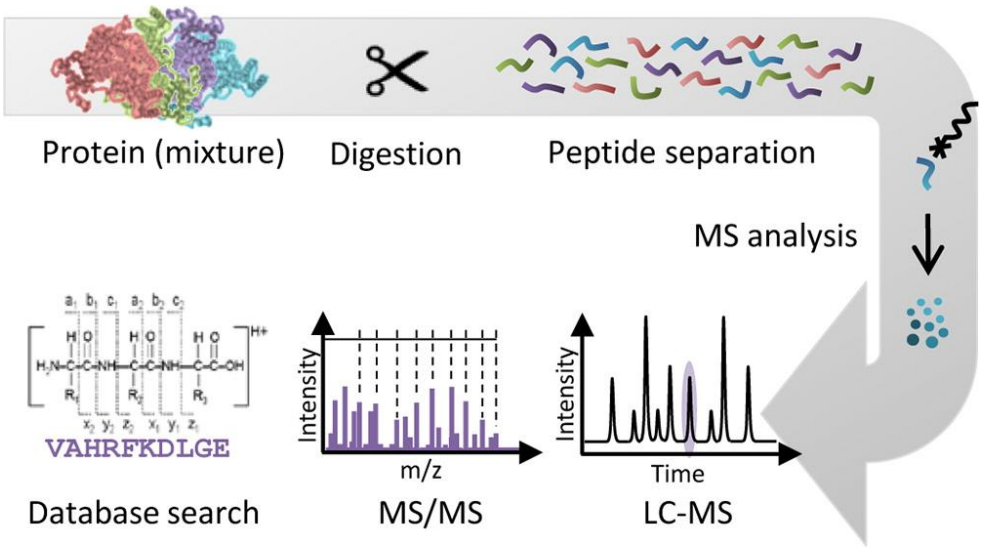


Figure 1.1 Simplified workflow overview for bottom-up proteomics.[8]

Peptide identification is then achieved by comparing the tandem mass spectra derived from collision-induced dissociation (CID) with those from a database. Peptides can either be uniquely assigned to a single protein or shared by more than one protein. The identified proteins are usually scored and grouped based on their peptides. Typically to confidently assign a protein, a unique peptide must be identified.[5]

1.2 Peptide Fragmentation

While most peptides are identified simply by matching the peak lists from the CID spectra from those generated from a protein data base, it is also possible to identify a peptide manually using what is known as *de novo* peptide sequencing. In this method the fragment ions are pieced together, like a jigsaw puzzle, to identify the peptide as a whole. This method can be time-consuming but can also yield key results because many known peptide fragmentation pathways are not incorporated into database scoring, such as the loss of imines and water.[9]

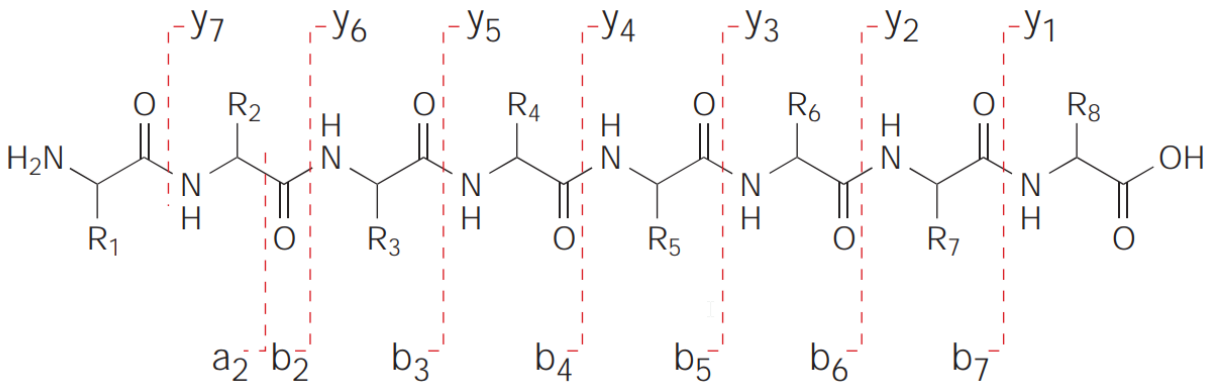
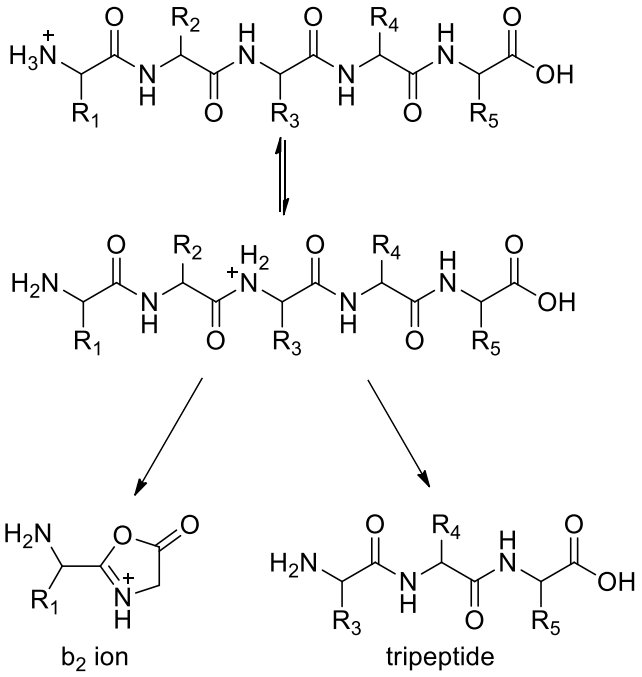
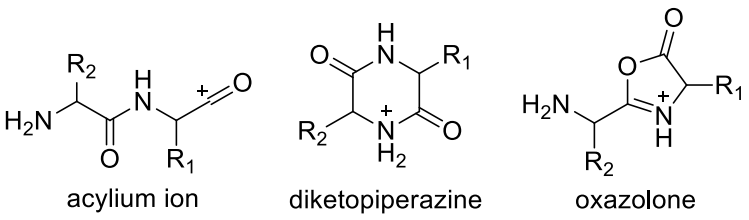


Figure 1.2 Peptide fragmentation by CID usually dominates in the cleavage of amide bonds forming $[b_n]^+$ and $[y_m+2H]^+$ ions. Subsequent fragmentations and generation of other ions are also possible. Typically, a $[b_n]^+$ ion fragments to an $[a_n]^+$ ion.[10]

Depending on the ion activation method, different fragment ions will be formed but the most common method ion activation is by means of CID. Protonated peptides tend to cleave at the amide bonds forming $[b_n]^+$ and $[y_m+2H]^+$ ions (Figure 1.2).[10] Amide bond cleavage is facilitated by protonation of the amide nitrogen, elongating the C-N bond and weakening it. Typically a preceeding amide oxygen then undergoes nucleophilic attack on the amide carbonyl carbon, cleaving the C-N bond and forming $[b_n]^+$ or $[y_m+2H]^+$ ions with the relative proton affinities dictating which ions are observed. The $[b_n]^+$ ions are believed to be protonated oxazolones while $[y_m+2H]^+$ ions are truncated peptides as seen in Scheme 1.1.[11–13]

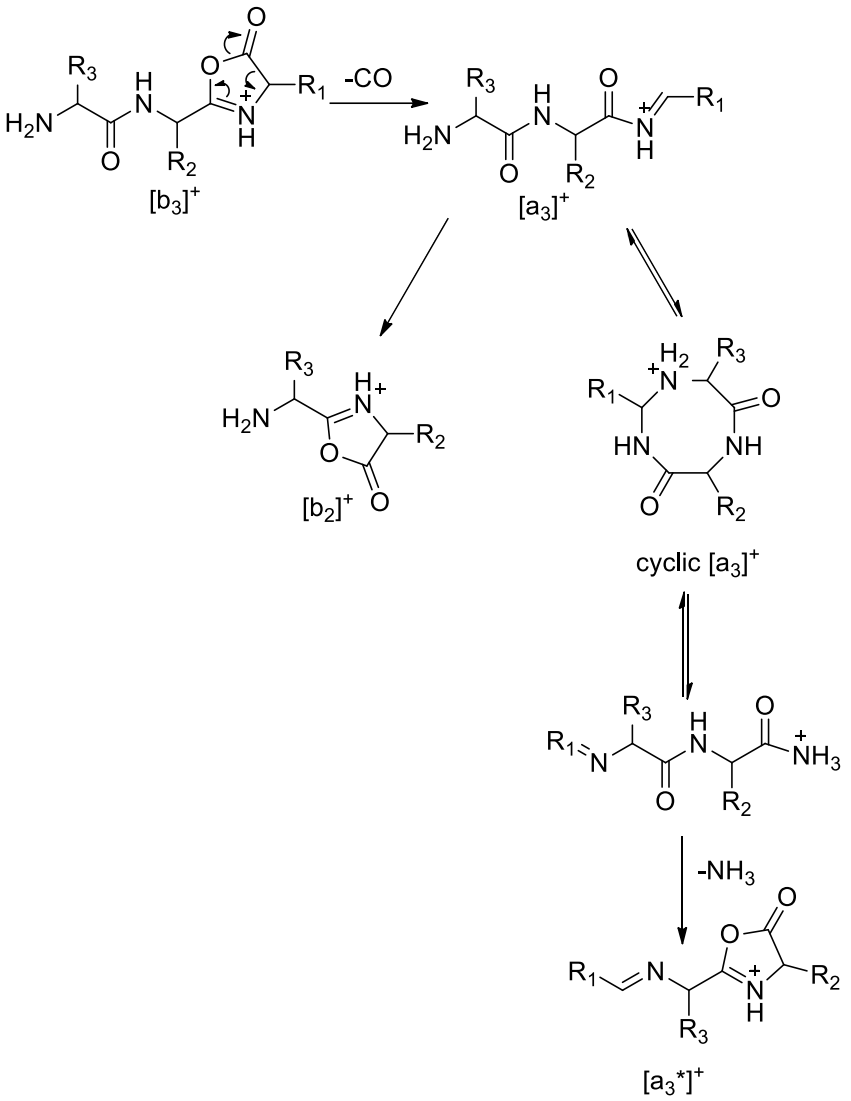


Scheme 1.1 Fragmentation mechanism for the formation of $[b_2]^+$ and $[y_3+2H]^+$ ions.



Initially the b ion was believed to be an acylium ion, formed by cleavage of the amide bond after protonation.[14–16] Acylium ions tend to lose CO easily but b ions are stable structures; consequently, the diketopiperazine and oxazolone structures were proposed as potential structures for b ions. The diketopiperazine is formed by nucleophilic attack of the N-terminal amine to an amide carbon facilitating amide bond cleavage. The oxazolone is formed by nucleophilic attack by a preceding amide oxygen facilitating amide bond cleavage. The b ions were initially believed to be diketopiperazines because nitrogen is a better nucleophile than oxygen due to its accessible lone pair of electrons and the diketopiperazine has a lower energy structure than the isomeric oxazolone.[17] However, many experimental and theoretical studies have found the b ion to be an oxazolone structure.[11,12,18–21] The amide bond possesses some double bond character which is in a *trans* configuration; this positions the amide carbonyl carbon accessible to nucleophilic attack by the preceding amide oxygen. To form the diketopiperazine, the double bond has to undergo *cis-trans* isomerization to position the carbonyl carbon accessible to nucleophilic attack by the N-terminal amine. The barrier to this *cis-trans* isomerization is the reason why the barrier to diketopiperazine function is larger than that to oxazolone formation.[22]

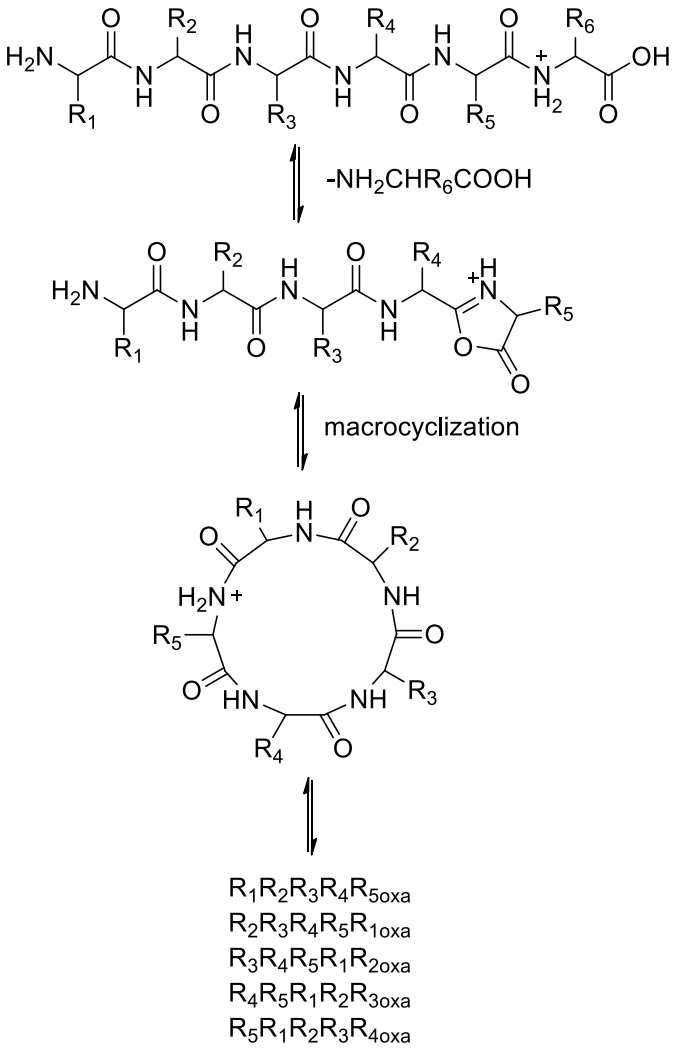
The $[b_n]^+$ and $[y_m+2H]^+$ ions are also prone to undergo further fragmentation. Typically, a $[b_n]^+$ ion can undergo ring opening and lose CO forming an $[a_n]^+$ ion (Scheme 1.2).[13] These $[a_n]^+$ ions can cyclize by nucleophilic attack by the N-terminal amine on the double bond at the C-terminus. Following a 1,3 H^+ shift, the C-N bond is cleaved reforming a linear $[a_n]^+$ ion, an imino-amide. These $[a_n]^+$ ions, most notably $[a_3]^+$, lose ammonia by nucleophilic attack by the adjacent amide oxygen, facilitating amide bond cleavage.[23] The $[a_n]^+$ ions can also lose an imine HN=CHR to form a $[b_{n-1}]^+$ ion.[24]



Scheme 1.2 Fragmentation of $[b_3]^+$ ions to form $[a_3^*]^+$ ions by loss of CO. The same mechanism can be applied to larger $[b_n]^+$ ions.

The truncated peptides, $[y_m +2H]^+$ ions, follow a fragmentation pattern similar to that of the protonated peptide. They can fragment to form smaller b and y ions (Scheme 1.1).[13] As peptide length increases the $[b_n]^+$ ions can undergo macrocyclization where the N-terminal amine undergoes nucleophilic attack of the oxazolone ring (Scheme 1.3). The resulting macrocycle can then undergo ring opening and scramble the peptide sequences such that different b ions containing the same amino acid residues, but in a different sequence, will result in identical CID spectra. [25–28] There is also a variety of other common

fragmentation pathways such as neutral losses of H_2O and NH_3.[29] Due to the variety of fragmentation pathways and the challenges they present, peptide identification could become difficult. Therefore, it is important to understand all the fragmentation pathways to better identify peptides. By understanding all the fragmentation pathways, it becomes possible to better interpret fragmentation products and hence identify a peptide.



Scheme 1.3 Macrocyclization of $[b_n]^+$ ions by nucleophilic attack of the N-terminal amine. Ring opening is achieved by formation of an oxazolone, oxa denotes the oxazolone structure.

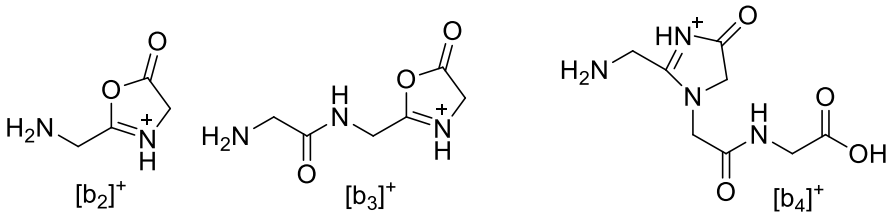
1.3 Water loss from peptides

The wider scientific community are unaware of the formation of 4-imidazolones, the products of water loss at the N-terminus. Typically, water loss is attributed to the formation of the oxazolone at the C-terminus. The purpose of this body of work is to study water loss from protonated peptides and the fragmentation mechanisms and structural features of 4-imidazolones. The work here has wide implications for proteomics, as peak lists used for peptide identification do not consider the loss of water or the formation of internal imines; utilizing the fragmentation mechanisms described here may improve peptide identification.

CID fragmentation of peptides generates a variety of ions, not just $[b_n]^+$ and $[y_m+2H]^+$ ions, and not all ions formed by CID are well understood. A better understanding of the possible fragmentation pathways can assist in identification of peptides. A common pathway that is not well understood is water loss. An early example is bradykinin which was shown to have water loss as the dominant pathway. Ballard *et al.* proposed three possible sources of water loss. The first was from acidic groups including the carboxylic acid at the C-terminus and the side chains of glutamic and aspartic acid. The second was hydroxyl groups in the side chains. The third is from the peptide backbone using an oxygen from the amide bond.[29]

The pathways for water loss from polyglycines vary greatly depending on the length of the peptide backbone. Protonated diglycine loses significant amounts of water from the C-terminus in the formation of an oxazolone.[30–32] For protonated triglycine, water loss is a minor pathway and comes predominantly from the C-terminus by formation of a oxazolone.[33] However, water loss from tetraglycine was shown to come primarily from the first and second amides, using isotopic labelling and density functional theory (DFT) calculations. The structure was determined to be an 4-imidazolone using infrared multiphoton dissociation (IRMPD) spectroscopy and DFT calculations.[35,36] Subsequent work then showed that

protonated triglycine also loses water from the N-terminal amide group as well.[34] As peptide length increases, C-terminal water loss becomes a non-competitive pathway and N-terminal water loss, from the amide group, becomes more dominant for polyglycines (n>3).



In these studies glycines are predominantly used because they are the simplest amino acids and appear in high frequency in most proteins. Previous work studying water loss also used glycine systems and a follow-up study in the CID of the 4-imidazolone formed from protonated tetraglycine is discussed in Chapter 3. The $[b_4]^+$ ions formed from loss of water from the first and second residues of tetraglycine showed identical CID spectra which suggest interconversion into identical structures prior to fragmentation.

After tetrapeptides, longer polyglycines were studied. In Chapter 4 water loss from pentaglycine and hexaglycines are discussed. Water loss from the first and second amides of pentaglycine are both major pathways, while water loss from the first second and third amides of hexaglycine are major pathways. The $[b_5]^+$ and $[b_6]^+$ ions formed by water loss from the amide backbone of pentaglycine and hexaglycine do not show identical spectra, unlike the $[b_4]^+$ ions formed from tetraglycine. Alternate fragmentation pathways become competitive with the interconversion mechanism (discussed in Chapter 3) as the peptide length is increased.

Peptides typically contain a variety of amino acid residues, not just glycine. In Chapter 5, a glycine of tetraglycine is substituted by an alanine or a proline residue. Water loss is a minor channel after the substitutions. This may be due to the bulkier side chains, which sterically hinder nucleophilic attack as well

as electronic effects. Water loss is still primarily from the first and second residues, with the C-terminus being a minor channel. The CID spectra of these $[b_4]^+$ ions do not show identical CID spectra, suggesting that other fragmentation pathways are competitive with the interconversion pathway discussed in Chapter 3.

Chapter 2 Experimental Techniques

2.1 Mass spectrometers

Mass spectrometers are instruments which produce ions and separate them based on their mass to charge ratios (*m/z*). This is the ratio of the atomic mass to the fundamental charge of the ion. Numerous ionization methods and mass spectrometers exist.[39] The only ionization technique discussed herein will be electrospray ionization along with the mass analyzers: triple quadrupole, and orbitrap.

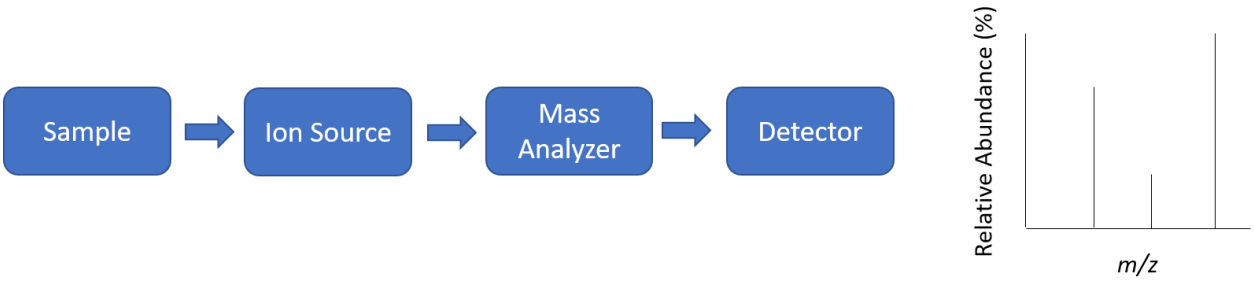


Figure 2.1 Basic schematic of a mass spectrometer

The basic components of a mass spectrometer can be seen in Figure 2.1 The inlet is designed to introduce a micro amount of sample into the ion source. The sample is then converted into gaseous ions, by electrospray ionization, a soft ionization technique. The ions are then directed into a mass analyzer to determine their *m/z* ratios.[40] There are many different mass analyzers and many companies that produce them. The two that are discussed here are the orbitrap, a high-resolution mass analyzer, from Thermofisher and a triple quadrupole from SCIEX.[41,42] When performing analysis of samples, more information can be gained by fragmenting the gas phase ion and examining the fragment ions. A soft ionization technique, like electrospray ionization, cannot sufficiently fragment the ions.[40] An ion activation method is needed to produce some fragmentation to deduce structural features. Many methods of fragmentation exist for a multitude of purposes.[43] However, in this thesis we will only discuss collision-induced dissociation (CID). A detector is then used to convert the beam of ions into an electrical signal

that can be processed and stored. A common detector is the electron multiplier. An electron multiplier can have different configurations but typically consist of a cathode and a series of dynodes, kept at higher voltages. When energetic ions strike the surfaces of the cathode, electrons are produced. These electrons are attracted to the dynodes, as they are kept at higher voltages. Each electron that contacts the dynode generates more electrons. The electrons are then attracted to the subsequent dynodes. Depending on the number of dynodes a large number of electrons are produced for every ion that strikes the cathode.[44] Sequential delivery of mass separated fragment ions then result in a CID spectrum. The CID spectrum provides information on the relative abundance and the *m/z* ratio of each ion.

2.2 Electrospray Ionization (ESI)

Soft ionization techniques received a major breakthrough in the late 20th century after the discovery of electrospray ionization (ESI) and matrix assisted laser-desorption ionization (MALDI) for the study biomolecules. Previously, only hard ionization techniques like electron impact were commonly used in mass spectrometry. In addition, ESI and MALDI are amenable to analytes with low vapour pressure, including proteins and DNAs. For their discoveries, John B. Fenn and Koichi Tanaka respectively were awarded the Nobel Prize in 2002 for Mass spectrometry applied to biological macromolecules.[1,2] Herein, only ESI will be discussed as it was the only ionization technique utilized, in this thesis work.

ESI has three conceptual steps: nebulization of the sample into electrically charged droplets, solvent evaporation, and ion transport into the mass spectrometer. The processes are depicted in Figure 2.2. The sample is pushed through a capillary using a mechanical pump and a syringe. As the sample is fed through the capillary tube, an electric field pulls the positive charge towards the capillary front. A small electrically charged droplet leaves the surface when the electrostatic repulsion is greater than the surface tension of the solvent. A small charged droplet then leaves the surface of the capillary and travels through

the surrounding gas towards the mass spectrometer due to a potential gradient.[45] Often a sheath gas (dry N_2) is directed coaxially around the capillary which allows for better nebulization.[46]

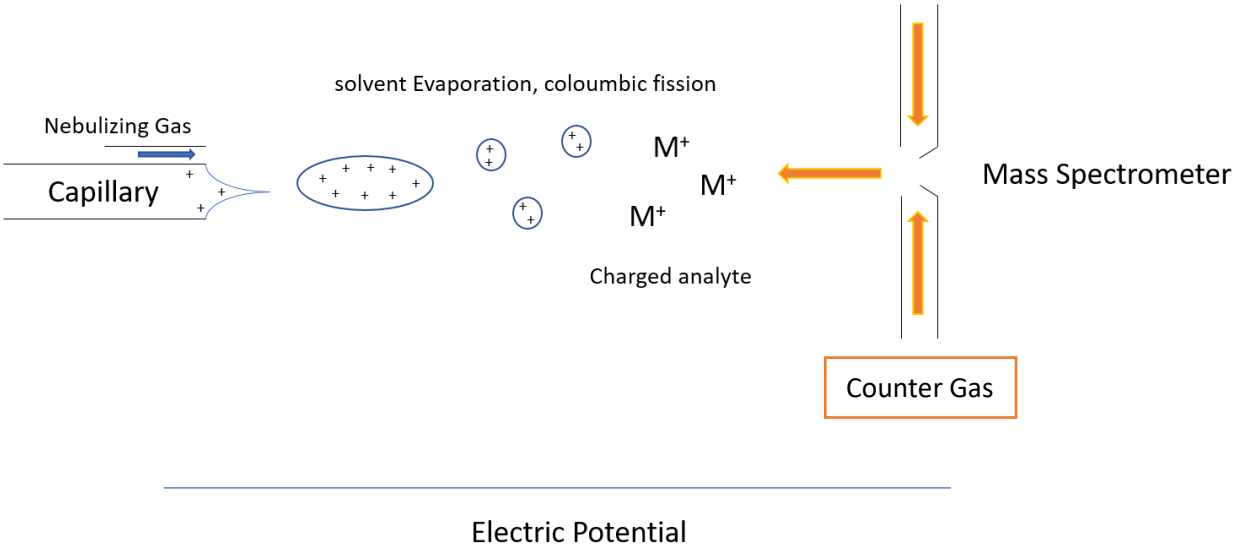


Figure 2.2 A simplified representation of the electrospray ionization process.[40]

As the droplets travel towards the inlet, solvent evaporation occurs reducing the size. This process can be accelerated using a heated capillary or heated sheath gas. As solvent evaporates, charge density increases. Droplets also undergo shear forces as they fly through dense gas at atmospheric pressure. The deformations cause the droplets to form a non-spherical shape where certain regions have higher electric fields and protrusions in the droplet. In some cases the deformations and charge density will exceed the surface tension and the droplet will fall apart, a phenomenon known as coulombic fission.[47] The stability limit where surface tension can hold the droplet together is called the Rayleigh stability limit.[48] Once the droplets undergo coulombic fission, a small jet of microdroplets leave the original parent droplet. This process of evaporation and coloumbic fission repeat until the droplet is sufficiently small to release the charge analyte into the gas phase. This process is called ion evaporation. It is also possible for the analyte to form solvation shells and form a nanodroplet. If a sample has not been desolvated before entering the

inlet, dry curtain gas desolvates the ions. The outlet for curtain gas is coaxial to the droplet direction but outward from the inlet, such that the flow of the droplets is opposite to the curtain gas flow. The ion, free from any solvation is then introduced into the mass spectrometer using the potential difference.[45]

2.3 Quadrupole Mass spectrometer

The quadrupole mass spectrometer is the most common type of mass spectrometer, due to ruggedness and low cost.[49] The quadrupole mass analyzer consists of four rods that serve as electrodes to filter and select certain *m/z* ratios at a given time. Opposite rods are connected electrically, one pair are connected to the positive terminal and the other pair to the negative terminal. A variable radio-frequency (RF) voltage is also applied to each pair of rods. A simplified schematic can be seen in Figure 2.3.[50]

When only alternating current (AC) is applied to the rods in the *xz* plane, ions will tend to converge in the center of the channel during the positive half of the AC cycle. This is because the ions are positively charged and are repelled by the positively charged rod. The opposite happens during the negative half of the AC cycle, the ions diverge from the center and move towards the rods. If the ion strikes the rods during the negative half of the cycle, the ions are neutralized. A positive direct current (DC) voltage is superimposed on the AC signal. Newtonian physics and kinetics explains that it is more difficult to deflect a heavier mass than it is to deflect a lighter one. Therefore, if the ion has a high mass and the frequency of the AC is large, the ion will not respond significantly to the AC and will be largely influenced by the DC potential. Then those ions with sufficiently high mass will remain between the rods guided by the positive DC current. The rods in the *yz* plane are maintained at a negative dc voltage. Since the ions are positively

charged, those ions that are largely affected by the DC voltage will collide with rods. However, the lighter ions that are affected strongly by the AC current will be able to pass through the field.[42,50]

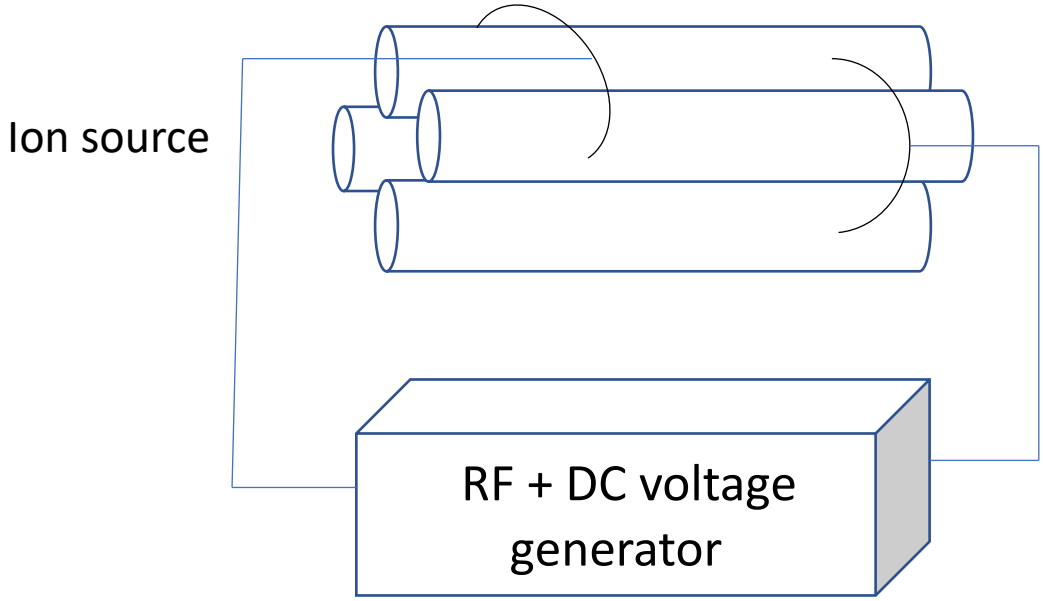


Figure 2.3 A simplified representation of a quadrupole mass spectrometer by Miller *et al.*[50]

For the ion to travel through the quadrupoles, it must have a stable trajectory for both sets of rods in the *xz* and *yz* planes. Therefore, it must be sufficiently heavy to pass through the high mass filter in the *xz* plane and the low mass filter in the *yz* plane. For one set of AC and DC voltages only a small band of ions with limited *m/z* values are transmitted. This band can be adjusted by adjusting the AC and DC voltages. Typically to scan through a mass spectrum the AC and DC voltages are increased simultaneously while maintaining a ratio of approximately 6. This allows the quadrupoles to scan through all the *m/z* ratios as seen in Figure 2.4.[50]

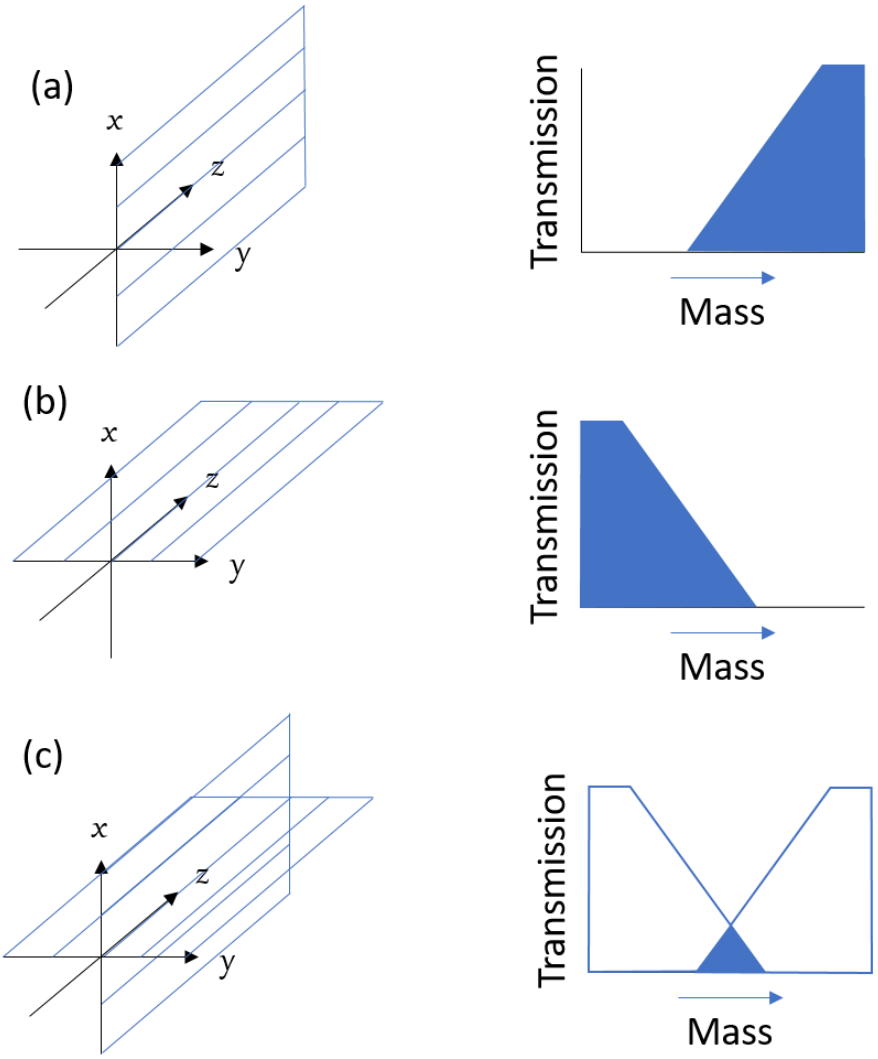


Figure 2.4 A simplified representation of the quadrupole mass filters from Miller *et al.* (a) acts as a high mass filter (b) acts as a low mass filter and (c) is when both the high mass filter and low mass filter are operational allowing us to selectively choose a m/z.[50]

For the QTRAP 4000 (SCIEX) which was used in many preliminary studies, it consists of three main quadrupoles. The first set of quadrupoles (Q1) allows us to select the precursor ion, protonated peptide. The quadrupole selects the ions by using the high and low mass filters. Then the ions are guided into (q2) for fragmentation. The second set of quadrupoles is not for ion selection but for collision-induced

dissociation, described in Chapter 2.5. The third set of quadrupoles (Q3) then mass analyzes the fragment ions for detection.[42]

2.4 Orbitrap Mass spectrometer

The Orbitrap Elite was used to record all of the final spectra for all the projects. The Orbitrap contains two main components: the dual-pressure linear ion trap and the orbitrap mass analyzer. The dual pressure linear ion trap is comprised of two quadrupoles, one at high pressure used for CID and a low pressure cell that is used for mass scanning ions. The quadrupoles function as described in chapter 2.3. There are some advantages to having a dual pressure linear ion trap including: high capture efficiency into the linear ion trap and efficiency of fragmentation. By operating at two different pressures the amount of collision gas introduced into the high pressure cell can be better controlled which improves the collisional induced dissociation as described in chapter 2.5.[7]

The distinguishing component of the Orbitrap Mass spectrometer is the Orbitrap mass analyzer which provides high resolution m/z values of up to >100,000.[41,51–53] The Orbitrap mass analyzer was used for ions that have similar *m/z* ratios where the quadrupole mass analyzer was unable to distinguish between them. A notable example is the m/z ratio for the loss of water (18.010565) compared to the loss of $^{15}NH_3$ (18.023584). The operating resolving power of a quadrupole does not have the adequate resolution to distinguish between these small mass differences.[51,52]

The Orbitrap mass analyzer is comprised of three electrodes as seen in Figure 2.5.[54] There are two outer electrodes which are concave towards the center electrode. The central electrode is in the shape of a prolate spheroid, such that the center of the electrode has the largest diameter while the edges have smaller diameters. When voltage is applied along the outer and center electrodes, a radial electric field bends the ions trajectory toward the central electrode. The ions remain in a nearly circular spiral inside the trap much like how a planet orbits the sun, the sun being the central electrode. Due to the irregular

shape of the electrodes, the ions are pushed towards the widest part of the trap, the middle, initiating harmonic axial oscillations. The outer electrodes are used as receiver plates for the image current detection of these axial oscillations. This digitized image is in the time domain which then undergoes Fourier-transform into the frequency domain, which is then converted into a m/z spectrum.[41,51,52,54]

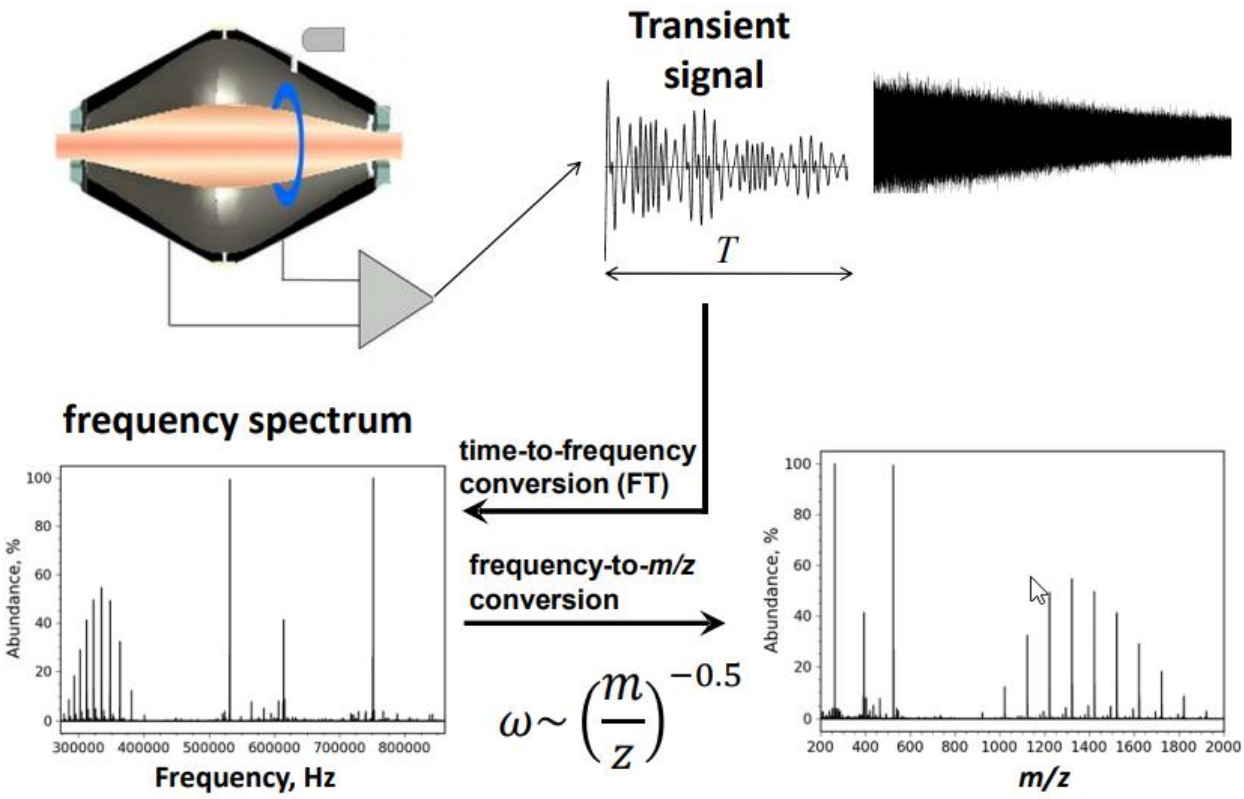


Figure 2.5 A simplified description of the process used to obtain a *m/z* spectra using an Orbitrap mass analyzer.[54]

2.5 Collision Induced Dissociation (CID)

With the emergence of soft ionization techniques, ions need to be activated for the observation of any fragmentation. As biological applications have been the dominant functions for mass spectrometry in the last decade, collision-induced dissociation (CID) has become the most common ion activation method. This is the ion activation method that is used in both the Orbitrap and the QTRAP 4000. In this method, the precursor molecule collides with an inert gas, imparting internal energy into the precursor

ion to cause bond cleavage and product ion formation.[55] The fragmentation spectrum contain a myriad of information as the relative abundances of the fragment ions reflect the relative energies of the bonds that cleave.[56]

Ions with high translational energy collide with an inert gas, N_2 in most SCIEX machines, and He in the Orbitrap, and undergo inelastic collision which converts part of the translational energy into internal energy. The ion undergoes fragmentation as the ion gains sufficient energy to overcome the bond energy, typically of the weakest bond. The total kinetic energy that can be converted into internal energy is the relative energy (E_{com}) and depends on the masses of the two objects colliding.[55]

$$E_{Com} = \left(\frac{N}{m_p + N}\right) E_{lab}$$

E_{lab} represents the ion's total kinetic energy, while m_p and N are the masses of the precursor ion and neutral gas respectively.

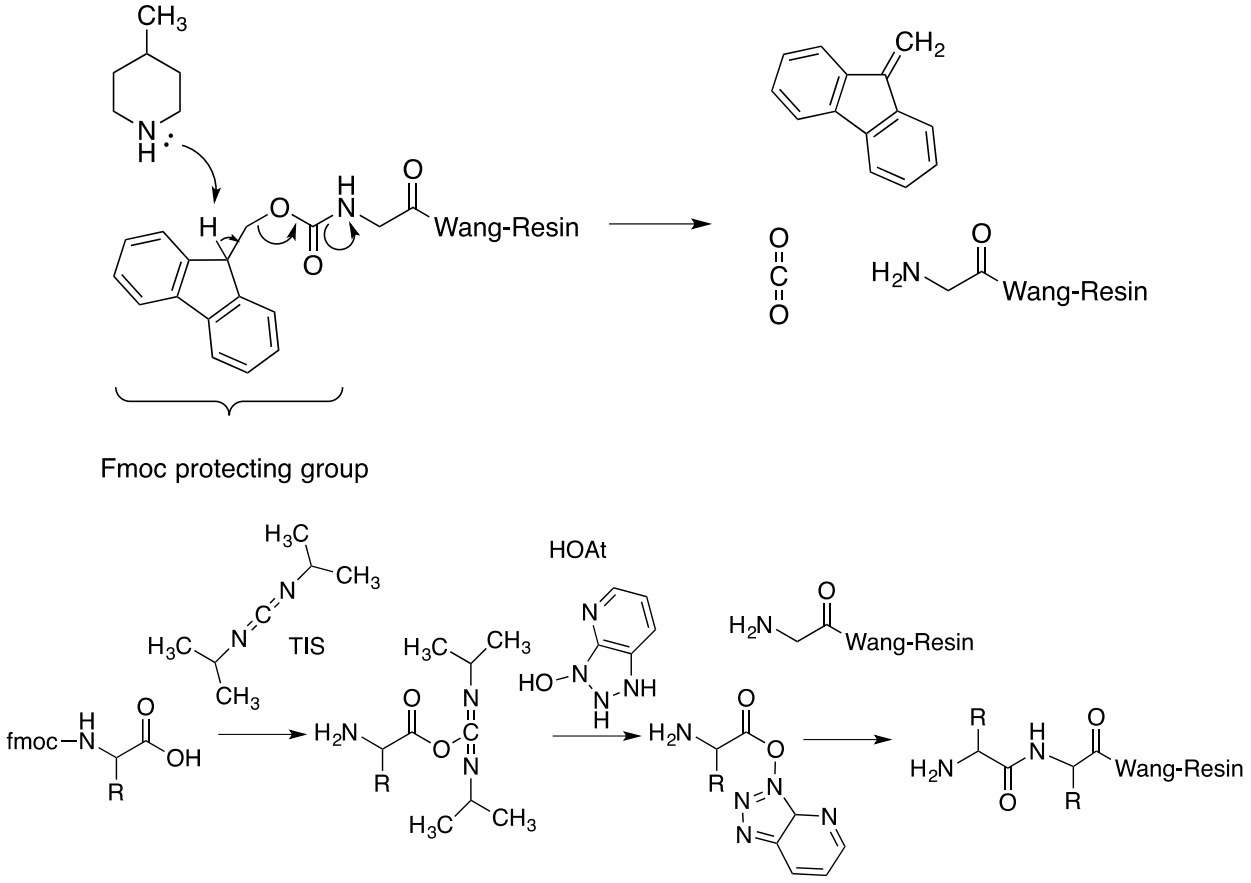
When there is sufficient energy to cause dissociation the overall CID process is assumed to occur as a two step process.

$$m_p^+ + N \rightarrow m_p^{+\neq} + N \rightarrow m_f^+ + m_n$$

In the equation described above m_f^+ is the fragment ion while m_n is the neutral ion formed because of fragmentation. The first step is the slowest step, where the ion is accelerated by the electric field to collide with the neutral gas molecule. The second step is on a much faster time scale where the translational energy is converted into internal energy for fragmentation.[55]

2.6 Solid-phase peptide synthesis (SPPS)

A variety of peptides have been synthesized according to the scheme outlined below. Isotopically labelled amino acids were incorporated using the same technique.



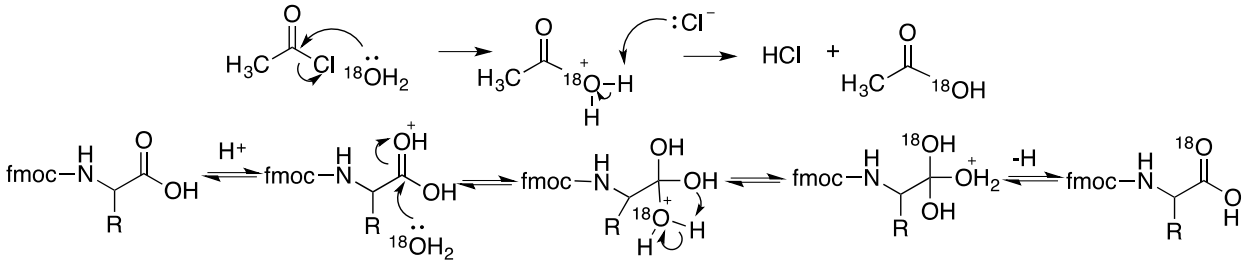
Scheme 2.1. solid-phase peptide synthesis as described by Coin *et al.*[57]

1. Fmoc-amino acid-Wang Resin (0.25 mmol) was weighed out and placed into a reaction vessel.
2. Dichloromethane (3 mL) was added to reaction vessel to swell the resin, and then removed by pressure from nitrogen.
3. 20/80 (v/v) 4-methyl-piperidine/DMF was added to the reaction vessel, to cleave the protecting group (FMOC), and allowed to bubble under nitrogen for 13 minutes. The solvent was removed using pressure from nitrogen.
4. Dimethylformamide (5 mL) was then used to rinse the reaction vessel and then removed. This was repeated 3 times.
5. Dichloromethane (5 mL) was then used to rinse the reaction vessel and then removed. This was repeated 3 times.
6. Fmoc-amino acid-OH (0.375 mmol), 1-Hydroxy-7-azabenzotriazole (0.375 mmol), N, N'-Diisopropylcarbodiimide (0.1 mL), and Dimethylformamide (3 mL) was added to the reaction vessel. The solution was allowed to bubble under nitrogen for 1.5 hours. The solvent was then removed under pressure from nitrogen.

7. Dimethylformamide (5 mL) was then used to rinse the reaction vessel and then removed. This was repeated 3 times.
8. Dichloromethane (5 mL) was then used to rinse the reaction vessel and then removed. This was repeated 3 times.
9. Steps 3 – 8 were repeated until the desired length of the peptide was reached.
10. 20/80 (v/v) 4-methyl-piperidine/DMF was added to the reaction vessel, to cleave the protecting group (FMOC), and allowed to bubble under nitrogen for 13 minutes. The solvent was removed using pressure from nitrogen.
11. Dimethylformamide (5 mL) was then used to rinse the reaction vessel and then removed. This was repeated 3 times.
12. Dichloromethane (5 mL) was then used to rinse the reaction vessel and then removed. This was repeated 3 times.
13. A cleaving solution (3 mL) of trifluoroacetic acid, dichloromethane, triisopropylsilane and water (14:4:1:1) (v/v) were added to the reaction vessel. The solution was allowed to bubble under nitrogen for 1.5 hours.
14. The resin was washed with dichloromethane twice with dichloromethane (1 mL).
15. The filtrate was collected and the solvent was allowed to evaporate using a speedvac.
16. Diethyl ether (1 mL) was then added to the powder or oil and sonicated for 1 minute. The resulting solution was allowed to sit overnight in a freezer (-16°C).
17. The diethyl ether was removed and the dry powder or oil was collected.

2.7 ^{18}O labelling of amino acids

To identify the structures of various ions and the loss of neutrals, isotopic labels were introduced to distinguish the m/z ratios of heteroatoms.



Scheme 2.2. The following protocol was adapted from Marecek *et al.* [58]

1. Oven dried glassware consisting of: a shlenk flask, a condenser and a glass nozzle were used as the reaction vessel. Figure 2.7 shows the set up.
2. Fmoc-X-OH (1 mmol), 1,4 dioxane (1 mL), $H_2{}^{18}O$(25 mmol), acetyl chloride (0.9 µL) and a magnetic stir bar were added to the shlenk flask.
3. The shlenk flask was heated to 70°C and stirred under nitrogen for 2 hours.
4. The solution was then freeze dried using liquid nitrogen.

5. The dry powder was then collected and could be used in step 6 of the solid phase peptide synthesis.

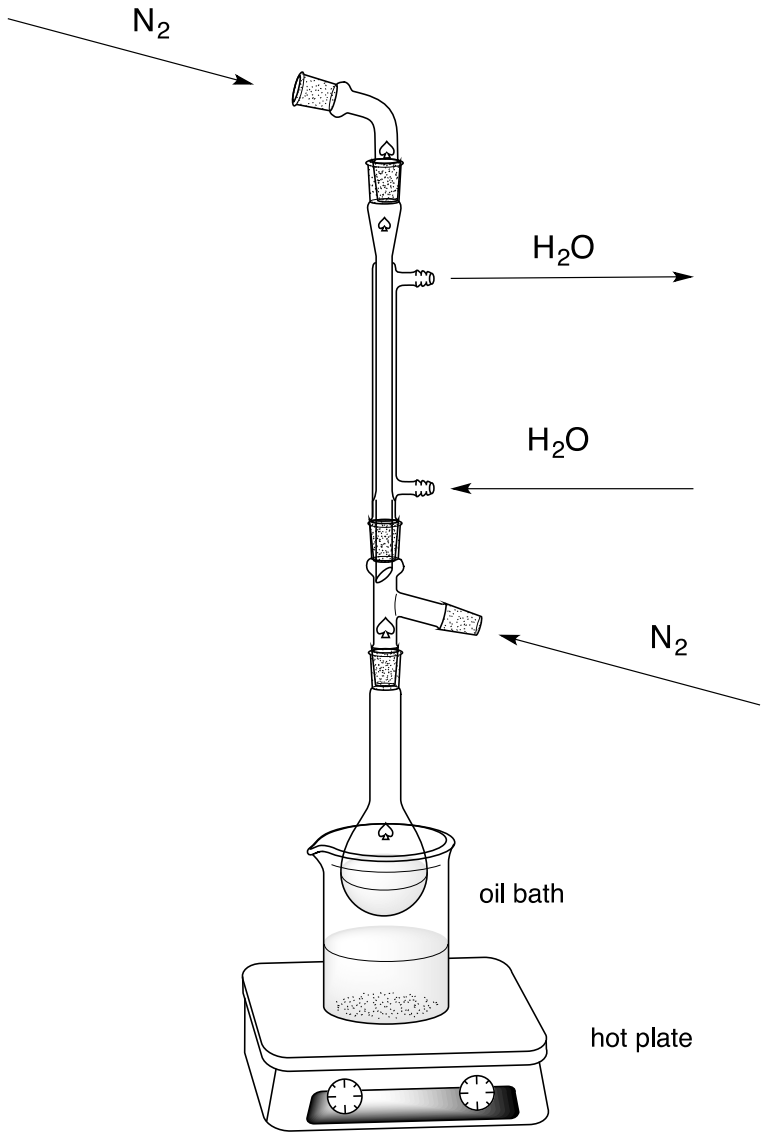


Figure 2.6 Setup for ^{18}O labelling of peptides. All fittings were sealed using grease, or a teflon sleeve. The gas line to supply the nitrogen was also connected to a bubbler to avoid overpressure.

2.8 Esterfication of amino acids

A methyl ester was introduced into the C-terminus of the peptides to determine whether fragment ions were coming from the C-terminus.

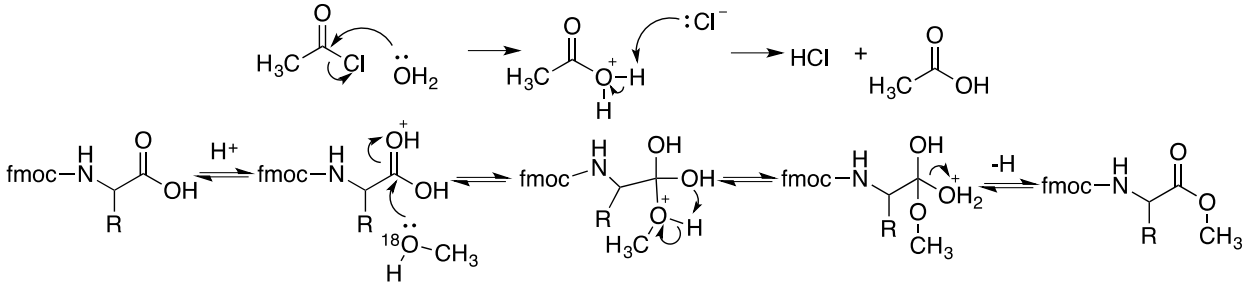


Figure 2.7 The esterification procedure was adapted from Kowalak *et al.*[59]

1. Anhydrous methanol (1 mL) and acetyl chloride (0.9 µL) was added to the dry peptide and allowed to react for 2 hours.

2.9 Sample preparation

Synthesized peptides were used without further purification. Peptides were stored in a freezer (-16°C) prior to experiments. Synthesized peptides were dissolved in acetonitrile/water (50/50) (v/v) followed by dilution to 5 µmol/L.

3.0 Symbols on CID spectra

In each spectrum the (*) denotes the precursor ion. Filled in dots (●) represents the lineage of the precursor ions and the empty circle (○) represents the ions scanned to generate the CID spectrum.

3.1 General nomenclature

For simplicity, the subscript number within the bracket represents the length of peptide from which water was initially lost. The subscript following brackets of the b ions represent the residue in which water was initially lost. For example the $[b_5]_I^+$ represents a b ion generated by loss of water from the first amide of pentaglycine.

Chapter 3 Interconversion between 4-imidazolone ions, isomers of $[b_4]^+$ derived from protonated tetraglycine

This chapter has been published as described in Appendix A: Information about the Author

3.1 Loss of water from $[GlyGlyGly + H]^+$

The loss of water from $[GlyGly + H]^+$ and $[GlyGlyGly + H]^+$ is reportedly from the carboxylic acid group[31–33,60,61] while water loss from $[GlyGlyGlyGly + H]^+$ is from an amide oxygen.[35,36] This very different behaviour prompted us to re-examine the fragmentation of $[GlyGlyGly + H]^+$. First we found that fragmentation of the ester $[GlyGlyGlyOMe + H]^+$ loses methanol and water, both in low abundance, with the second channel being the more minor (see Figure 3.1a). This encouraged us to synthesize Gly(^{18}O)GlyGly, where the N-terminal glycine was labeled with ^{18}O, and in the dissociation of $[Gly(^{18}O)GlyGly + H]^+$ we observed losses of both H_2O and $H_2{}^{18}O$, again both in low abundance and with the latter being the lower of the two (Figure 3.1b). The barrier to loss of water from the carboxylic group from M06-2X/6-311++G(d,p) calculations was found to be 128.2 kJ mol^{-1}.[33] The loss of water from the first peptide bond was higher in energy (166.7 kJ mol^{-1}) and the transition state for the latter process was very similar to that for loss of water from $[GlyGlyGlyGly + H]^+$ with the C-terminal glycine residue removed.[35]

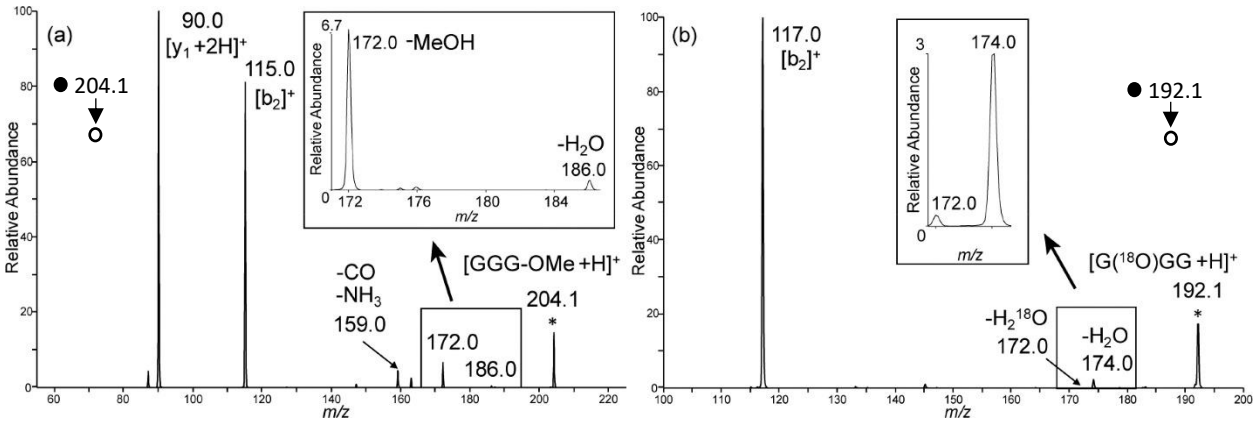


Figure 3.1. Expanded scale CID spectra of triglycine (a) after esterification and (b) with ^{18}O-labeled at the first amide oxygen. For brevity on the spectrum [GlyGlyGly-OMe + H]$^+$ is abbreviated to [GGG-OMe + H]$^+$. CID spectra (a) and (b) were collected at a normalized collision energy of 19 and 18 respectively.

3.2 Loss of water from [GlyGlyGlyGly + H]$^+$

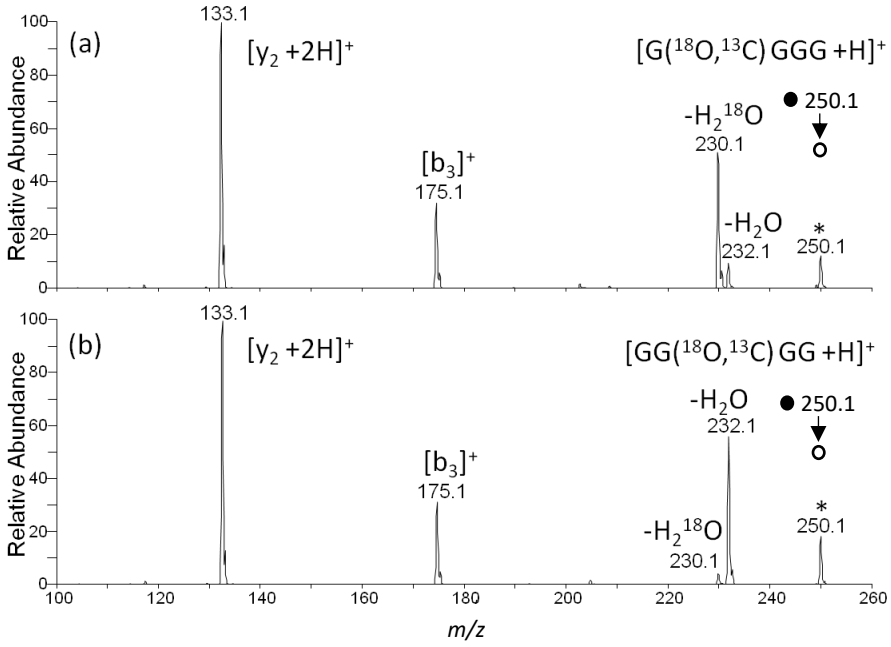
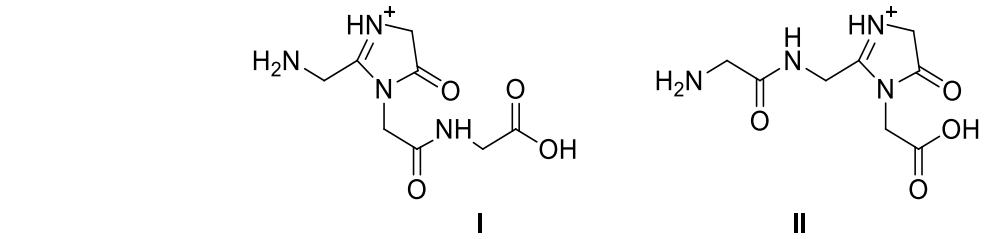


Figure 3.2. CID spectra of protonated tetraglycine with the (a) first and (b) second residue labeled by ^{18}O and $^{13}C_\alpha$. CID spectra (a) and (b) were collected at a normalized collision energy of 21.

Our current work confirmed that the water lost from [GlyGlyGlyGly + H]$^+$ comes predominantly from the first amide oxygen and that losses from the second amide and the carboxylic acid group are both minor channels.[36] We also labeled the third amide oxygen and found there is a negligible amount of water loss from this position (Figures 3.2 and 3.3). From DFT calculations, the 4-imidazolone (structure **I**) formed by

water loss from the first amide oxygen is lower in energy than when water is lost from the second amide (structure **II**) by 5.1 kJ mol^{-1}, which may explain why removal of water from the first amide is preferred over that from the second.



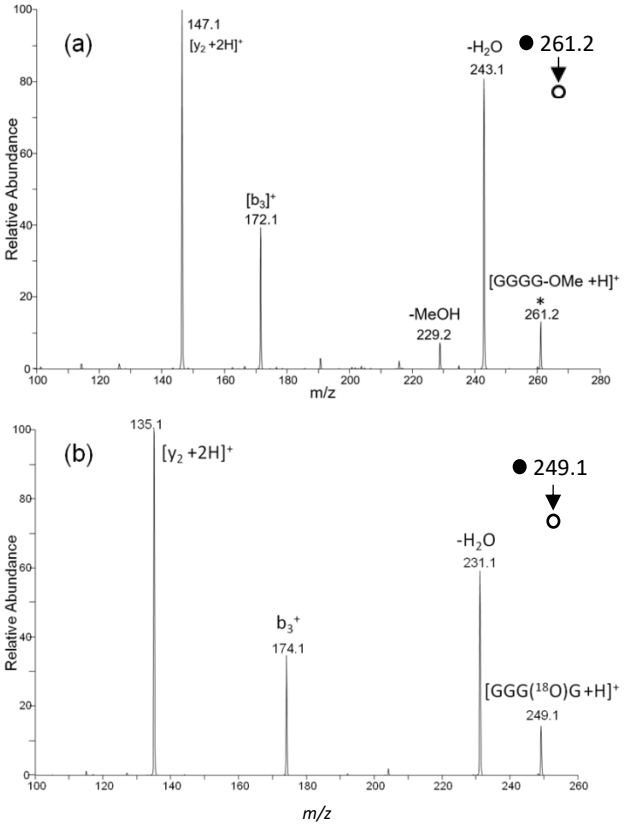


Figure 3.3. CID spectra of protonated tetraglycine (a) after esterification and (b) with the ^{18}O labeled at the third amide oxygen. CID spectra (a) and (b) were collected at a normalized collision energy of 21.

Interestingly, the CID spectra of the $[b_4]^+$ ions created by removal of water from the first and second amide groups are essentially identical (Figures 3.4(a) and 3.4(b)). The most abundant product is a $[b_4\text{-HN=CH}_2]^+$ ion at *m/z* 200, with losses of NH_3 and (CO + $HN=CH_2$) being very minor channels. The CID spectra of the $[b_4\text{-HN=CH}_2]^+$ ions (*m/z* 200) are also identical (Figures 3.4(c) and 3.4(d)) and the energy-

resolved curves for the $[b_4\text{-HN=CH}_2]^+$ ions were also similar (Figure 3.5). This suggests the two $[b_4]^+$ isomers rearrange to the same structure prior to fragmentation.

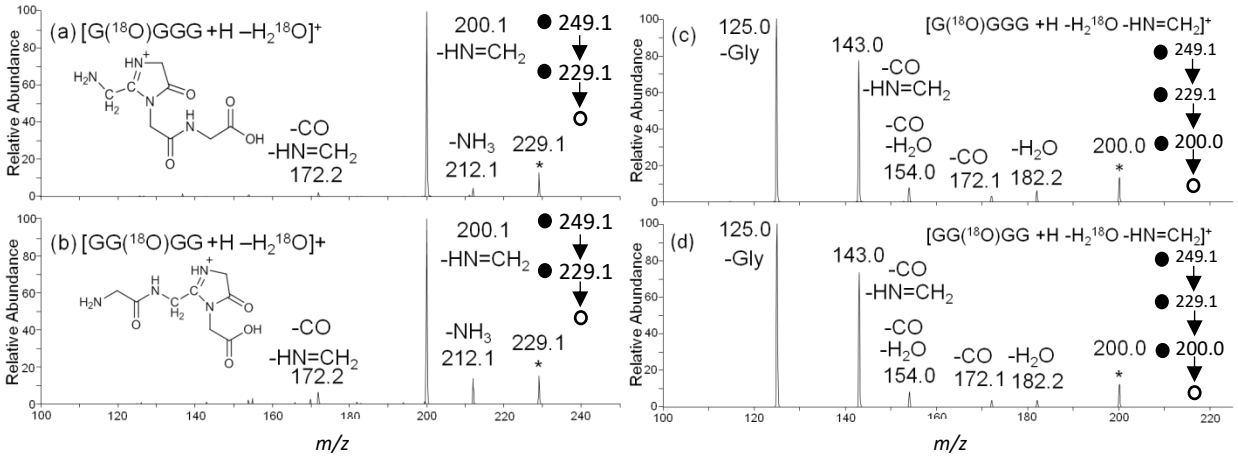


Figure 3.4. CID spectra of $[b_4]^+$ ions formed by loss of oxygen from the (a) first amide and (b) second amide. CID spectra (c) and (d) are for the dissociation of *m/z* 200 ions derived from (a) and (b) respectively. CID spectra (a) and (b) were collected at a normalized collision energy of 23 while (c) and (d) were collected at 26.

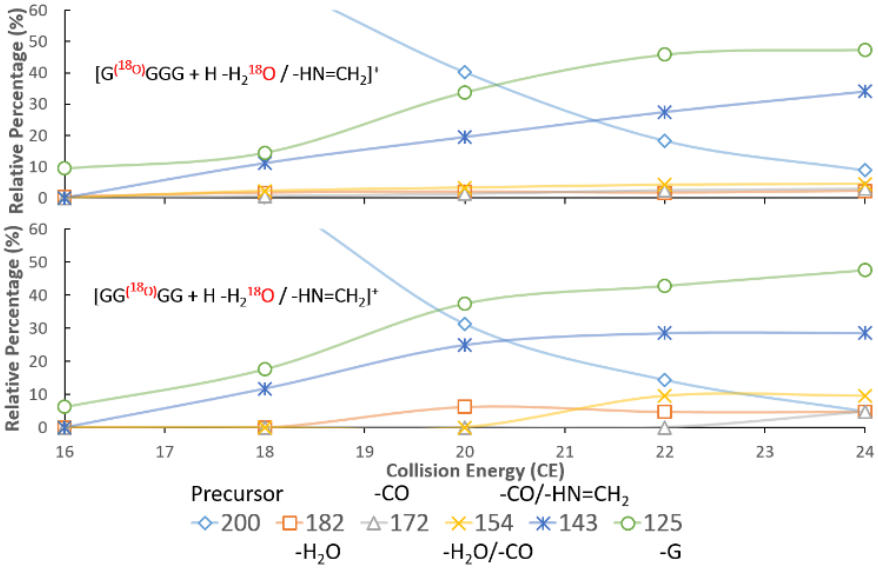


Figure 3.5. Energy-resolved curves for the ion $[b_4 - HN=CH_2]^+$ (structure **III**)

3.3 Fragmentations and structural isomerization of nominal $[b_4]^+$ ions from tetraglycine

Isotopically labelled glycine ($^{13}C_\alpha$) was used to identify which glycine residue is involved in the loss of methanimine. Comparison of Figures 3.4(a) and 3.6(a) shows that both the water and the methanimine came from the first glycine residue; similarly, comparison of Figures 3.4(b) and 3.6(b) show that both losses are from the second residue. From this we conclude that the loss of the methanimine is always from the *same* residue as the one that lost the initial water molecule.

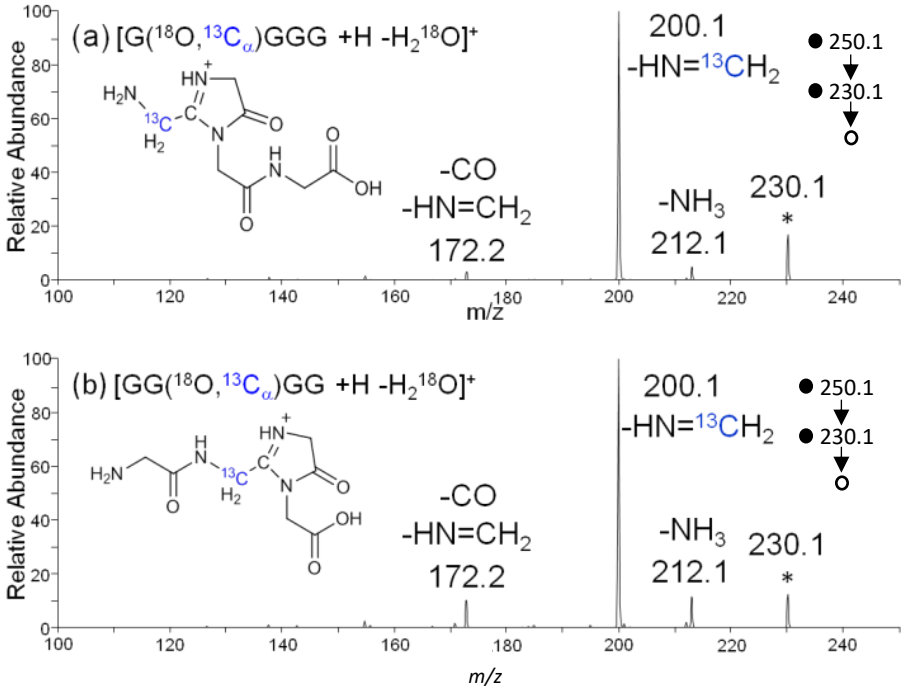
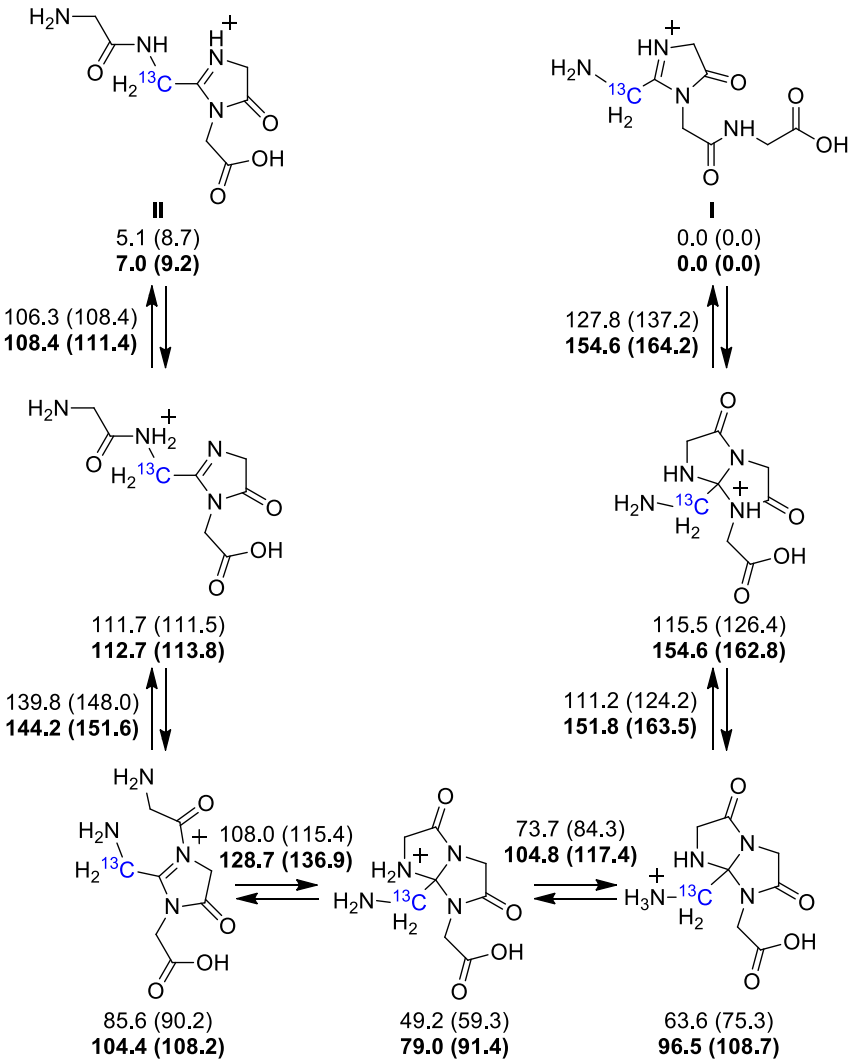


Figure 3.6. CID spectra of $[b_4]^+$ ions formed by loss of oxygen from (a) the first amide and (b) the second amide using ^{13}C labeled tetraglycines. CID spectra (a) and (b) were collected at a normalized collision energy of 23.

Loss of $HN{=}CH_2$ from the 4-imidazolone (structure **I**) formed by loss of water from the first amide is easily rationalized in terms of C_α–C bond cleavage adjacent to the five-membered ring as it is analogous to formation of an $[a_1]^+$ ion from a $[b_2]^+$ ion.[62] Loss of $HN{=}CH_2$ from the second residue requires rearrangement of structure **II** into **I**, thereby moving the second residue to the N-terminus. A possible pathway for rearrangement of **II** is given in Scheme 3.1. Proton transfer to the nitrogen of the N-terminal amide is followed by nucleophilic attack by the 4-imidazolone nitrogen on the amide carbon, resulting in

cleavage of the amide bond and transfer of the N-terminal residue onto the 4-imidazolone nitrogen. This second step has the highest barrier (139.8 kJ mol^{-1}) on the overall profile from the M06-2X/6-311++G(d,p) calculations. A subsequent attack on the newly-formed 4-imidazolone by the amino group of the NH_2 of the first residue followed by ring opening leads to structure **I** with an overall exothermicity of 5.1 kJ mol^{-1}.



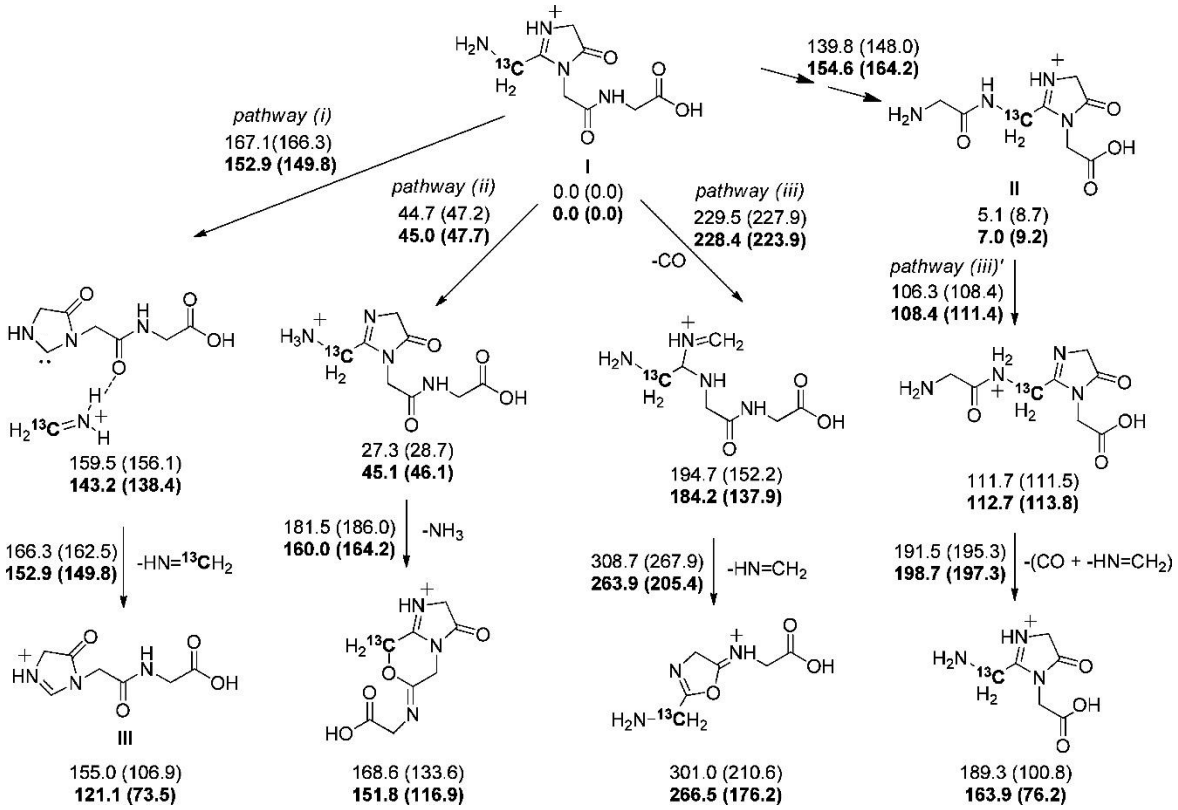
Scheme 3.1. Mechanism for the rearrangement of ion **II** into ion **I**. Enthalpies (at 0K) and (free energies at 298K), both in kJ mol-1 as calculated at the M06-2X/6-311++G(d,p) and B3LYP/6-311++G(d,p) levels. All energies are relative to structure I.

The pathways by which structure **I** dissociates or rearranges are given in Scheme 3.2.

(i) The loss of $HN{=}CH_2$ is the major pathway (*pathway i*) and occurs by direct cleavage of the exocyclic C_α–C bond to form ion **III**; the barrier to this pathway (167.1 kJ mol^{-1}) is 27.3 kJ mol^{-1} above the barrier to isomerization to **II**.

(ii) A 1,4-proton shift followed by nucleophilic attack by the carbonyl oxygen of the third amide group on the CH_2 of the first residue in structure I displaces ammonia with a barrier of 181.5 kJ mol^{-1} (*pathway ii*).

(iii) Direct loss of (CO + $HN{=}CH_2$) from **I** occurs in two steps, first loss of CO (product not observed experimentally) followed by a higher energy step in which $HN{=}CH_2$ is lost. The overall process has a much higher barrier (308.7 kJ mol^{-1}, *pathway iii*) and clearly is not competitive with the loss of NH_3 and probably is not the pathway by which CO and $HN{=}CH_2$ are lost. Rearrangement to **II** followed by cleavage of the first amide bond of **II** has a barrier of 191.5 kJ mol^{-1} (*pathway iii'*) and this makes the loss of (CO + $HN{=}CH_2$) more competitive. This pathway is initiated by a proton transfer in which the proton from the 4-imidazolone ring is transferred to the first amide nitrogen facilitating amide bond cleavage. The loss of (CO + $HN{=}CH_2$) without a ^{13}C label suggests that m/z 173 is not a subsequent loss of CO from m/z 200 as seen in Figure 3. This is consistent with what is proposed in pathway iii'.



Scheme 3.2. Mechanisms for the fragmentation and isomerization of structure I. Enthalpies (at 0K) and (free energies at 298K) both in kJ mol^{-1} as calculated at the M06-2X/6311++G(d,p) and **B3LYP/6-311++G(d,p)** levels. All energies are relative to structure **I**.

3.4 Fragmentations of structure III at *m/z* 200

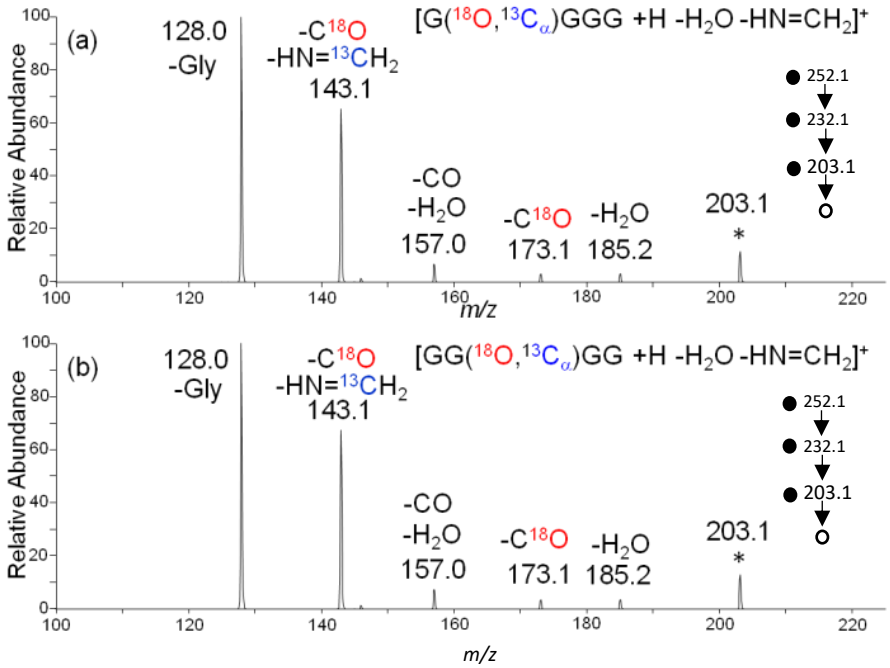


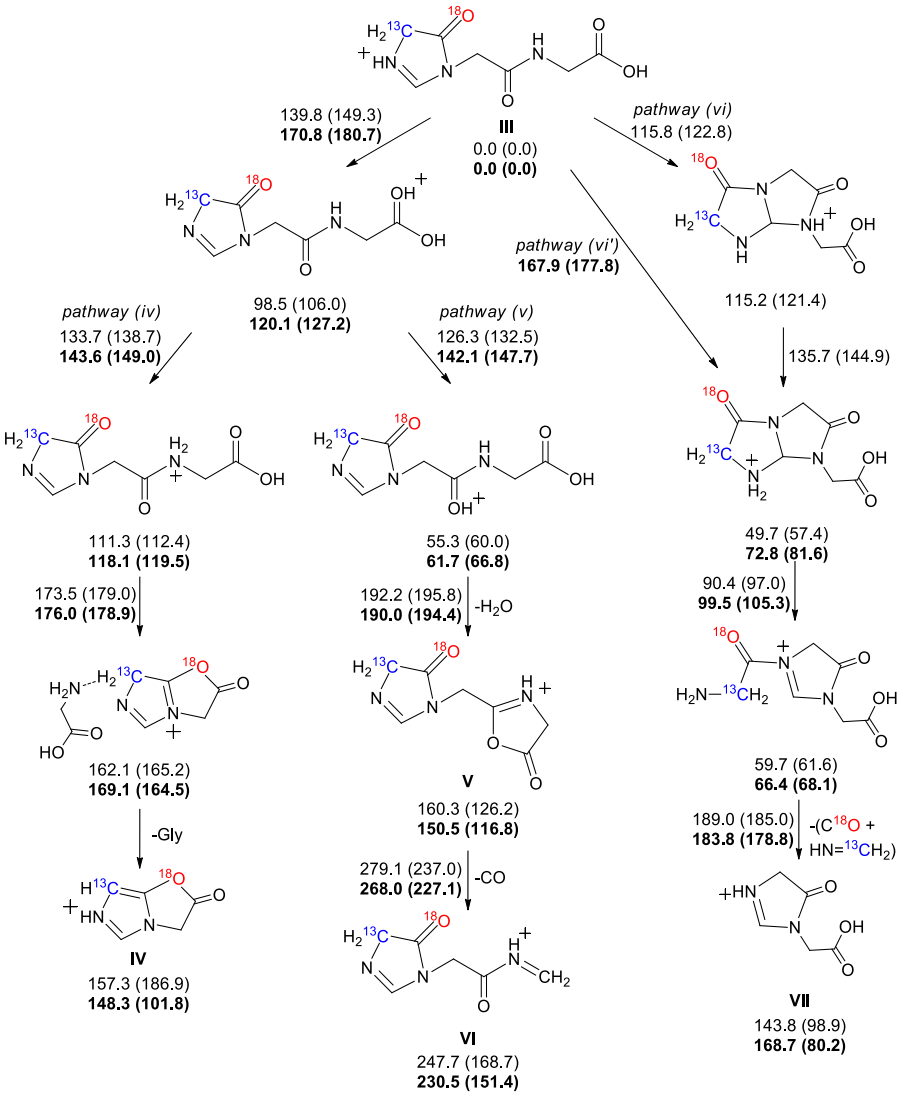
Figure 3.7. CID spectra of $[b_4 - HN{=}CH_2]^+$ (structure **III**) containing ^{18}O and $^{13}C_\alpha$ labeling in the (a) first and (b) second glycine residues. CID spectra (a) and (b) were collected at a normalized collision energy of 26.

Isotopic labeling was used to study the dissociation of $[GlyGlyGlyGly + H - H_2O - HN{=}CH_2]^+$, ion **III**. The spectra in Figures 3.7(a) and 4(b) are of ion **III** formed from **I** and **II**, respectively, where the glycine residue under consideration is labeled with ^{18}O and $^{13}C_\alpha$. Note that the ion in Figure 3.7 (a) has been formed by loss of water from the second glycine residue, structure **II**; similarly, that in Figure 3.7 (b) is the product of water loss from the first residue. Scheme 3.3 shows possible fragmentation mechanisms of this ion. The products ions that retain the labeled atoms are formed by losses of glycine and water molecules, giving structures **IV** and **V**, respectively. The loss of glycine (*pathway iv* in Scheme 3.3) is facilitated by proton migration from the 4-imidazolone ring to the amide nitrogen via protonation of the carboxy group. This is then followed by nucleophilic attack by the carbonyl oxygen of the 4-imidazolone on the carbon of the protonated amide, leading to cleavage of the amide bond and generating an ion-molecule complex (barrier = 173.5 kJ mol^{-1}). The α-hydrogen in the 4-imidazolone ring is acidic and the 1,2-proton transfer

is assisted by using the glycine molecule as a transporter to produce a highly conjugated system, structure **IV**.

Proton migration from the 4-imidazolone ring to the peptide backbone can also lead to the loss of a water molecule from the carboxylic group at the C-terminus to produce an oxazolone structure **V** (*pathway v*). Further loss of CO from the oxazolone ring produces an imine, structure **VI**.

The labeled atoms are lost as ($C^{18}O$ + $HN=^{13}CH_2$ in Figure 4) to give an ion at *m/z* 143. This is facilitated by nucleophilic attack by the amide nitrogen at the C-terminus on the carbocation in the 4-imidazolone ring to generate a bicyclic structure. Proton migration to the nitrogen of the original 4-imidazolone ring weakens the N−C_α bond within the ring and cleavage of this bond followed by breaking an amide bond results in the loss of (CO + $HN=CH_2$) containing the two isotopically labeled atoms from the N-terminus giving structure **VII**.



Scheme 3.3. Mechanisms for the fragmentation of [GGGG +H -H_2O - $HN=CH_2$]$^+$ (structure **III**) illustrated using the products in Figure 3.7. Enthalpies (at 0K) and (free energies at 298K) both in kcal mol^{-1} as calculated at the M06-2X/6-311++G(d,p) and **B3LYP/6-311++G(d,p)** levels. All energies are relative to structure **III**.

3.5 Intra-molecular H/D exchange

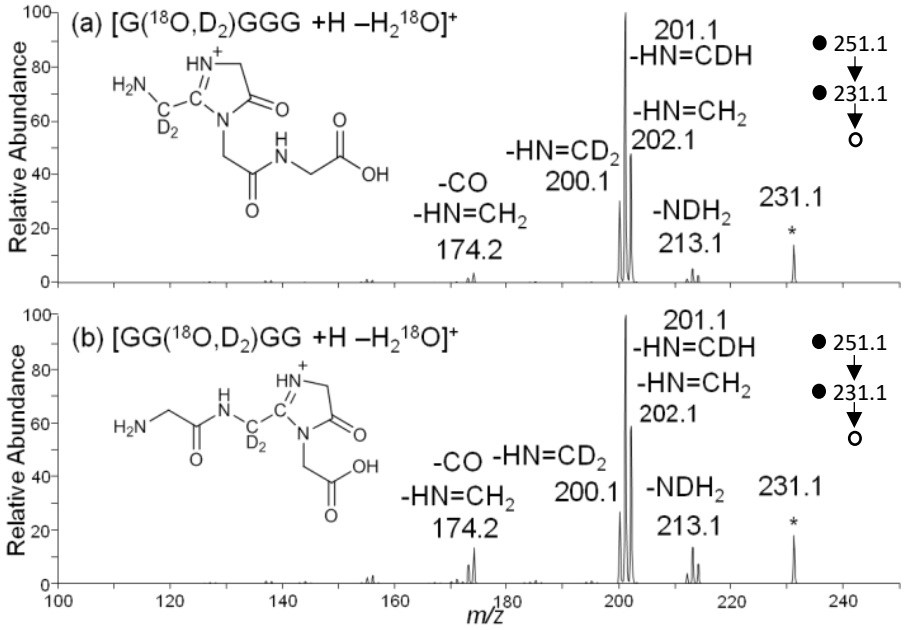


Figure 3.8. CID spectra of $[b_4]^+$ ions after water loss from (a) the first and (b) the second amide oxygens using ^{18}O-labeling and glycine (2,2-D_2) in the same residue. CID spectra (a) and (b) were collected at a normalized collision energy of 23.

While searching for possible mechanisms for water loss using isotopically labelled glycine (2,2 D_2), scrambling of the hydrogens at the α-carbon was observed. Figure 3.8 shows that the α-hydrogens from the same residue that lost the initial water molecule in $[b_4]^+$ ions can move to other basic sites. The energy barrier for α-hydrogen scrambling in structure **I** is 110.2 kJ mol^{-1} (Scheme 3.4), lower than the barriers to conversion to **II** and to dissociation. The acidity of the α-CH_2 is attributed to the proximity to the positive charge that is formally delocalized between the two nitrogen atoms in the 4-imidazolone ring of structure **I**. The energy-resolved curves for the deuterated $[b_4]^+$ in Figure 3.9 shows that loss ofmethanimine with one deuterium is the most dominant at all collision energies.

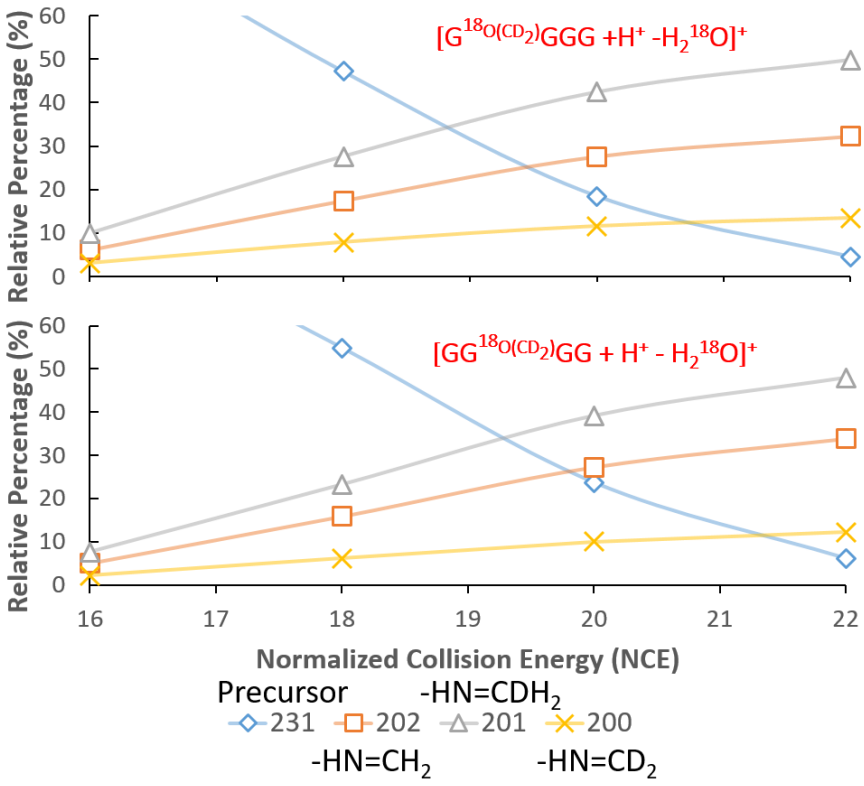
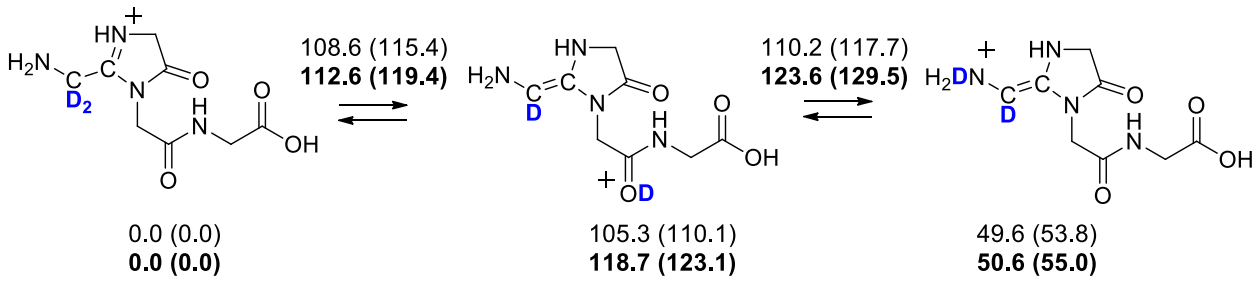


Figure 3.9. Energy-resolved curves for $[b_4]^+$ ions formed by loss of water from the first and second positions using glycine (2,2 D_2) in the residue that was dehydrated.



Scheme 3.4. Mechanism for α-hydrogen atom scrambling in structure **I**. Enthalpies (at 0K) and (free energies at 298K) both in kcal mol^{-1} as calculated at the M06-2X/6-311++G(d,p) and **B3LYP/6-311++G(d,p)** levels. All energies are relative to structure **I**.

Conclusion

The nominal $[b_4]^+$ ion formed by loss of water from the second amide oxygen of tetraglycine, protonated 4-imidazolone ion **II**, rearranges to the lower energy isomer **I** prior to dissociation. A consequence of this rearrangement is that the loss of methanimine from the $[b_4]^+$ ions derived from $[GlyGlyGlyGly + H]^+$ is always from the *same* residue from which water has initially been lost. The fragmentation pattern of $[GlyGlyGlyGly + H - H_2O - HN{=}CH_2]^+$, ion **III**, follows several pathways that are common in the fragmentation of protonated peptides. Finally, the hydrogen atoms on the α-carbon of the residue from which water is lost undergo extensive scrambling in the 4-imidazolone structure at lower energies than those required for rearrangement or dissociation.

Chapter 4 Fragmentation of protonated pentaglycine and hexaglycine after water loss, isomers of $[b_5]^+$

4.1 Loss of water from $[Gly_5 + H]^+$

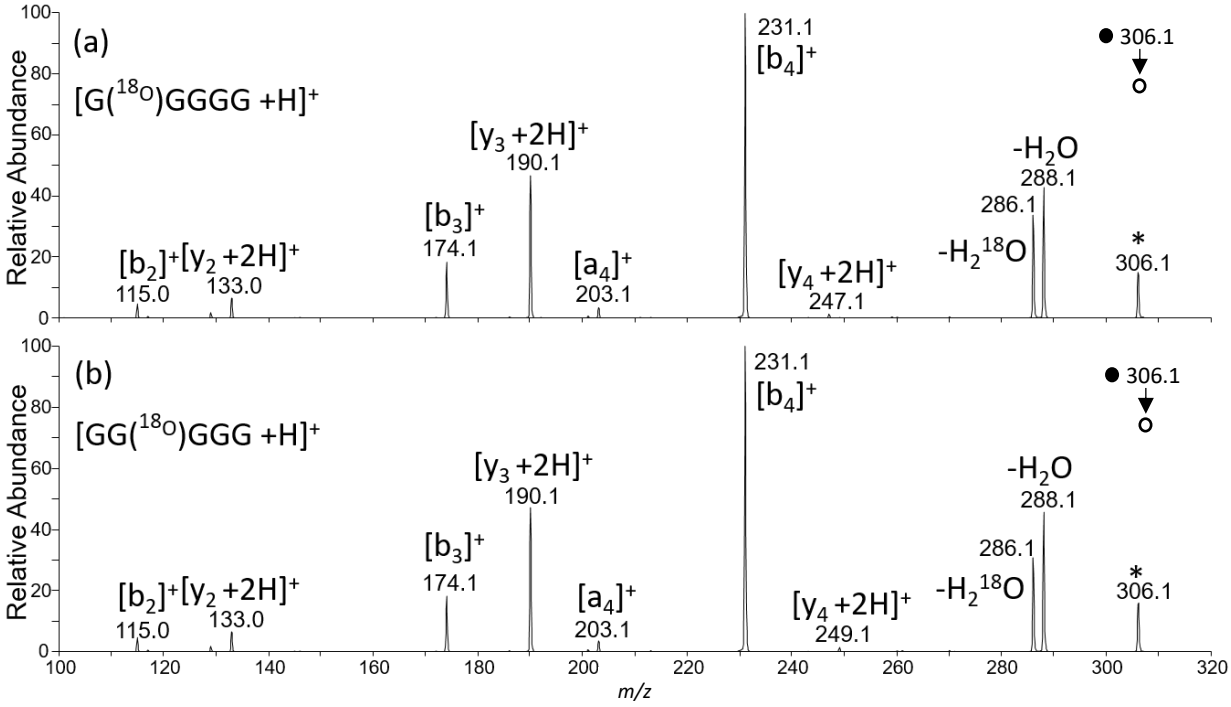


Figure 4.1 Spectra of $[b_5]^+$ ions formed by loss of oxygen (a) from the first amide and (b) from the second amide. CID spectra (a) and (b) were collected at a normalized collision energy of 20.

Water loss from $[Gly_5 + H]^+$ is a major dissociation pathway at all collision energies and ^{18}O-labelling showed that the loss is predominantly from the first and second amides with similar abundances (Figures 4.1 and Appendix B.1 and B.2). ^{18}O-labelling at the third and fourth amides revealed minimal water loss from these locations (Appendix B.3 and B.4). By comparison, protonated tetraglycine only had substantial water loss from the first amide group.[23]

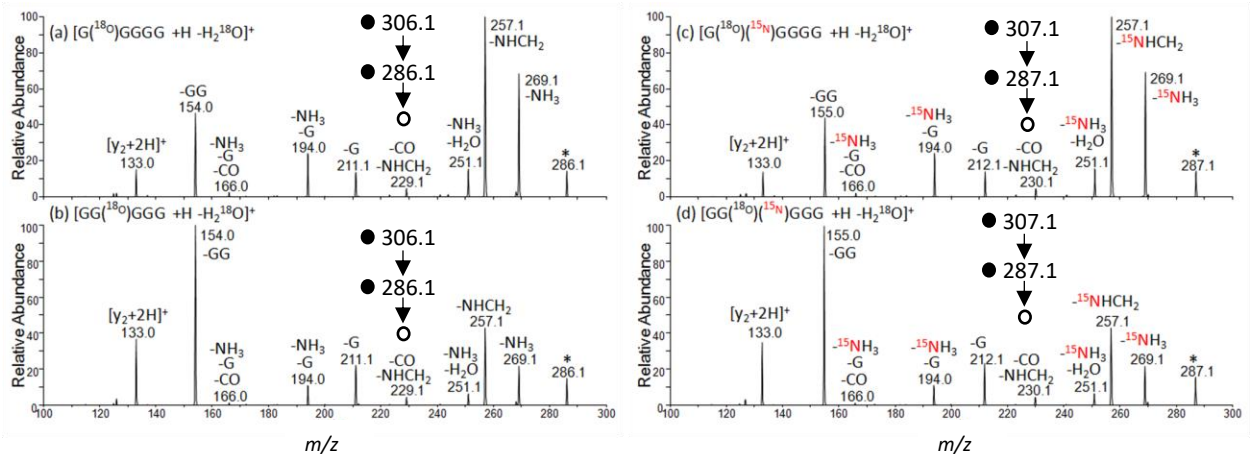


Figure 4.2. CID spectra of (a) $[b_5]^+_{I}$ and (b) $[b_5]^+_{II}$ ions. CID spectra (c) and (d) are for the $[b_5]^+_{I}$ and $[b_5]^+_{II}$ ions containing ^{18}O- and ^{15}N-labelling in the first and second residue, respectively.

The CID spectra of the $[b_5]^+_{I}$ and $[b_5]^+_{II}$ ions have the same product ions but are not identical, based upon the relative abundances of the product ions (Figure 4.2) and their very different energy-resolved curves (Figure 4.3). In this respect the fragmentation behaviours of these two $[b_5]^+$ ions differ from those of the $[b_4]^+$ ions of protonated tetraglycine, where the $[b_4]^+_{I}$ and $[b_4]^+_{II}$ ions had identical CID spectra, the result of interconversion prior to dissociation.[23]

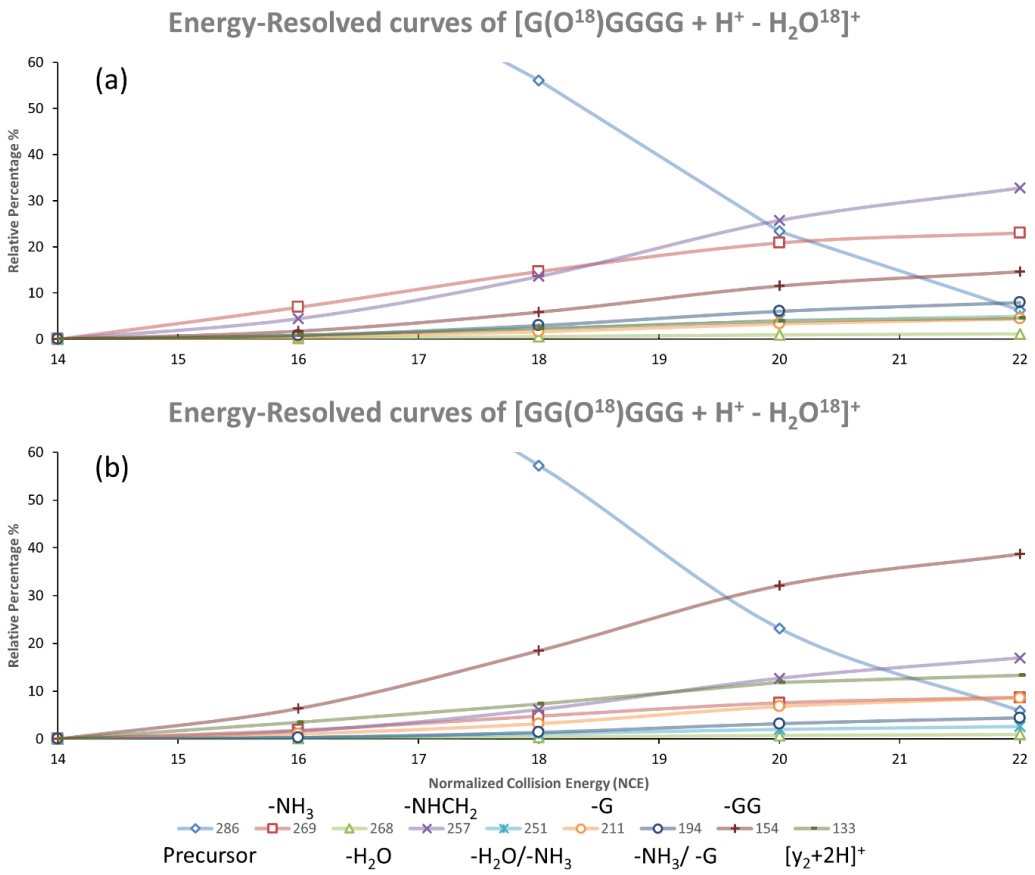


Figure 4.3. Energy-resolved diagrams for (a) $[b_5]^+_I$ and (b) $[b_5]^+_{II}$ ions.

For the $[b_5]^+_I$ ion energy-resolved curves (Figure 4.3(a)) show that NH_3 loss is the dominant pathway at the lowest collision energy giving an ion at m/z 269 that subsequently loses H_2O and glycine (ions at m/z 251 and 194 respectively). At higher energies loss of $HN{=}CH_2$ becomes the dominant pathway. Cleavage of the third amide bond to give the complementary ions at m/z 154 (loss of GlyGly) or 133 ($[y_2 + H]^+$) is the third most abundant pathway.

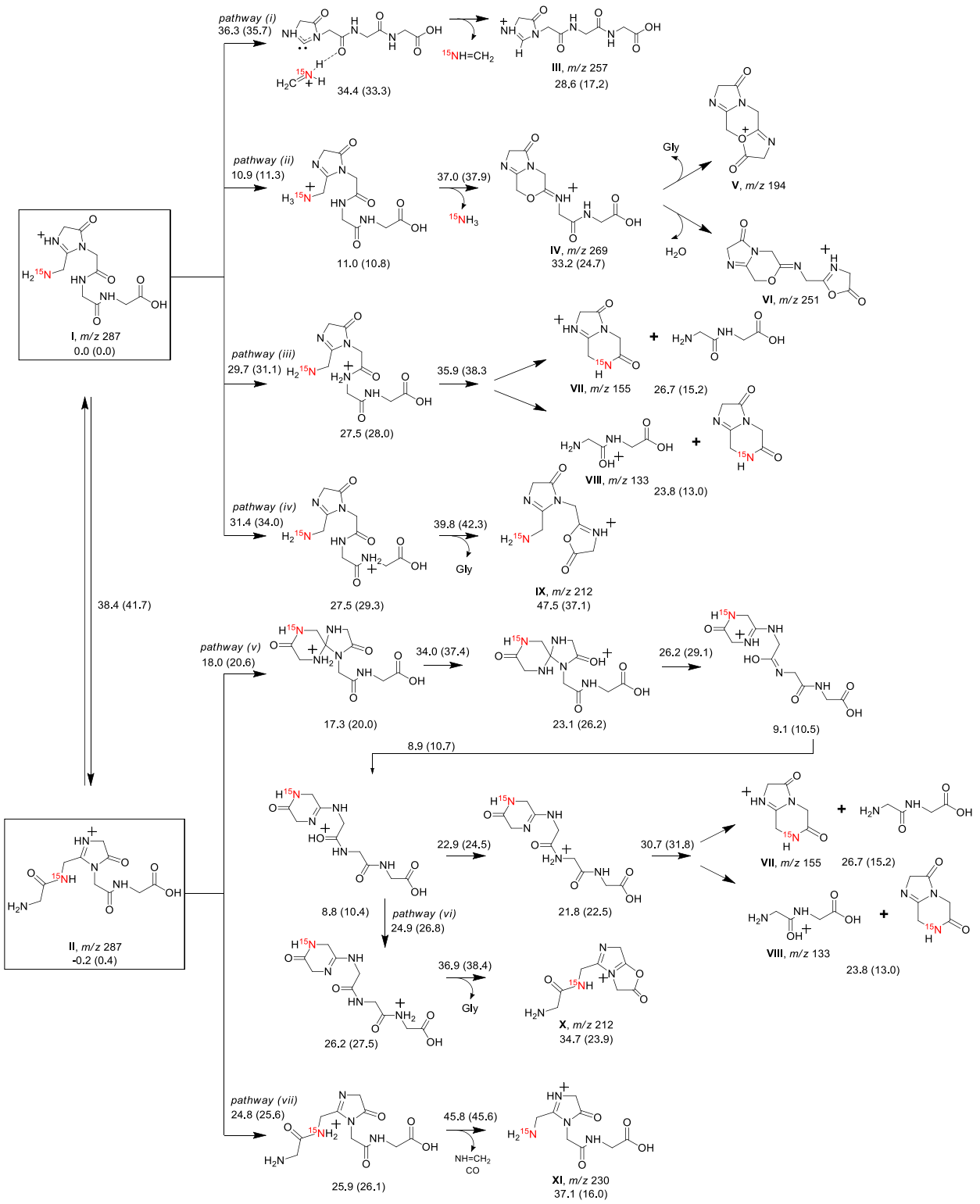
In contrast, for the $[b_5]^+_{II}$ ion cleavage of the third amide bond to give either an ion at m/z 154 or 133, $[y_2 + 2H]^+$ is the dominant pathway at all collision energies. The onset for the loss of $HN{=}CH_2$ (m/z 257) is at higher energies and this pathway becomes the second most abundant at higher collision energies. Loss of NH_3 (m/z 269) is a minor dissociation channel for this ion. From the differences in the energy-resolved

curves (Figure 4.3(b)) it appears that the two ions formed by losses of water from the first two amides of $[Gly_5 + H]^+$ dissociate more readily than interconvert; in this respect they differ from the $[b_4]^+_I$ and $[b_4]^+_{II}$ ions of tetraglycine.

Comparison of the CID spectrum of the $[b_5]^+_I$ ions with that of the $[b_4]^+_I$ from protonated tetraglycine showed some similarities, losses of $HN{=}CH_2$, NH_3, and $(CO + HN{=}CH_2)$ although the latter are minor fragmentation pathways for the $[b_4]^+_I$ ion. There are additional product ions in the spectrum of the $[b_5]^+_I$ ions, the results of cleaving the third amide bond and of losses of a glycine molecule (m/z 211), (glycine + NH_3, m/z 194) and (H_2O + NH_3, m/z 251). For both $[b_5]^+_I$ and $[b_4]^+_I$ ions, loss of methanimine was the dominant pathway at higher collision energies. For the $[b_5]^+_{II}$ ions, loss of methanimine also occurred, but was a relatively minor pathway. Instead amide bond cleavage resulting in the loss of GlyGly from the C-terminus and the formation of the $[y_2 + H]^+$ ion was the dominant pathway. There was no equivalent pathway observed for the $[b_4]^+_{II}$ ions. Isotopically labelled ^{15}N was used to determine from which residue of the $[b_5]^+_{II}$ ion the methanimine originated (Figure 4.2(d)) and, as was found for $[b_4]^+_{II}$ ions, the methanimine came from the residue that initially lost water. Similarly, the ^{15}N-labelling experiment also revealed the loss of NH_3 is always from the same residue from which water has been lost. This suggests that there is certain extent of interconversion of the $[b_5]^+_{II}$ ions into $[b_5]^+_I$ ions but, unlike for the $[b_4]^+$ ions, it is not the dominant pathway by which the $[b_5]^+_{II}$ ions dissociate.

The lowest-energy structure for the $[b_5]^+_I$ is an imidazolone, structure **I** (Scheme 4.1). The loss of methanimine from structure **I** is by the cleavage of the C_α–C bond (pathway (i)) resulting the formation of structure **III** (m/z 257). This mechanism is analogous to the formation of $[a_1]^+$ ion directly from a $[b_2]^+$ ion.[39] The barrier to this process is 36.9 kcal mol^{-1} and is very similar to that calculated for the same loss from the $[b_4]^+_I$ ion.[23] On pathway (ii) a 1,4-proton transfer from the imidazolone ring to the N-terminal amine group facilitates loss of NH_3 to produce structure **IV** (m/z 269) and the barrier to this process is 37.0 kcal mol^{-1}, essentially the same as for methanimine loss. The analogous process for the $[b_4]^+_I$ ion was slightly

higher in energy and the abundance of the ($[b_4]^+_I - NH_3$) ion was lower. Structure **IV** can further lose a glycine or water molecule from the C-terminus to produce structure **V** (m/z 194) and **VI** (m/z 251), respectively. Loss of GlyGly from structure **I** is initiated by proton transfer to the third amide nitrogen followed by nucleophilic attack by the second carbonyl oxygen of the imidazolone ring on the carbonyl carbon of the third amide group (pathway (iii)). The barrier to this step is 35.9 kcal mol^{-1}. If the proton is transferred to the C-terminus fragment, $[y_2 + 2H]^+$ ion is produced. Loss of a single glycine molecule (pathway (iv)) is facilitated by proton transferred to the fourth amide nitrogen, thus weakening the amide bond. Cleavage of the amide bond via nucleophilic attack by the adjacent carbonyl oxygen leading to the lost of glycine (structure **IX**, m/z 212).



Scheme 4.1. Mechanisms for the fragmentations of $[b_5]^+_{\mathbf{I}}$ and $[b_5]^+_{\mathbf{II}}$ ions. Enthalpies (at 0K) and (free energies at 298K), both in kcal mol^{-1} as calculated at the B3LYP/6-311++G(d,p) level. All energies are relative to structure **I**.

The barrier to conversion of structure **I** into **II** is 38.4 kcal mol^{-1}, again comparable to the barriers to dissociation of **I.** Detailed mechanism for the isomerization is summarized in Appendix B.7. Loss of GlyGly from structure **II** may also occur by conversion into structure **I** and then following pathway (iii). Alternatively, it can also occur by a lower-energy route initiated by nucleophilic attack by the N-terminal amine on the imidazolone ring, pathway (v). The barrier to the loss of GlyGly by this mechanism has a barrier of 34.0 kcal mol^{-1}, lower than the barrier to interconversion by 4.4 kcal mol^{-1}. This lower barrier for the loss of GlyGly from structure **II** explains why loss of GlyGly is the dominant product in the dissociation of the $[b_5]^+_{II}$ at all collision energies (Figure 4.3(b)). Due to the multiple collisions within the CID cell, some ions will possess more energy and are able to undergo conversion into $[b_5]^+_I$. This then leads to formation of the same additional products, ions at m/z 269, 257, 251 and 194, as formed from structure **I** but in diminished abundances. It is noted that the loss of a single glycine molecule can also occur from structure **II** directly (via pathway (vi)). The energy barrier (36.9 kcal mol^{-1}) for this process is slightly lower than that of the isomerisation (38.4 kcal mol^{-1}). Thus, it is more likely that the loss of glycine molecule is from $[b_5]^+_{II}$ directly. Finally, loss of (CO + $HN=CH_2$) is observed in very low abundance in the spectra of both the $[b_5]^+_I$ and $[b_5]^+_{II}$ ions. ^{15}N-labelling experiments established that this combined loss involves a different $HN=CH_2$ molecule from the formation of ion **III** (Figures 4.2(c) and (d)) and this is most easily achieved by cleavage of the first amide bond of structure **II** (pathway (vii)).

4.2 Loss of water from $[Gly_6 + H]^+$, protonated hexaglycine

As found in the dissociations of both protonated tetraglycine and protonated pentaglycine, water loss from $[Gly_6 + H]^+$ is a major fragmentation pathway, second only in abundance to the loss of a glycine molecule. CID of ^{18}O-labelled protonated hexaglycine revealed that water loss is predominantly from the first, second and third amides, with similar abundances for each $[b_6]^+$ ion (Appendix B.8, 9 and 10). The energy-resolved curves also revealed that losses of water from the first, second and third amide groups

are always major pathways at all collision energies (Appendix B.11, 12 and 13). Isotopic ^{18}O-labelling at the fourth and fifth amides established that there are minimal water losses from these positions (CID spectra in Appendix B.14 and B.15 and energy-resolved curves in Appendix B.16 and B.17). This behaviour is similar to the fragmentations of protonated pentaglycine, where there are significant water losses from both the first and second amide groups, and of protonated tetraglycine, where water loss is mainly from the first amide. Apparently, as peptide length is increased, water loss occurs from more amides at the N-terminus but the two amides at the C-terminal end lose minimal amounts of water.

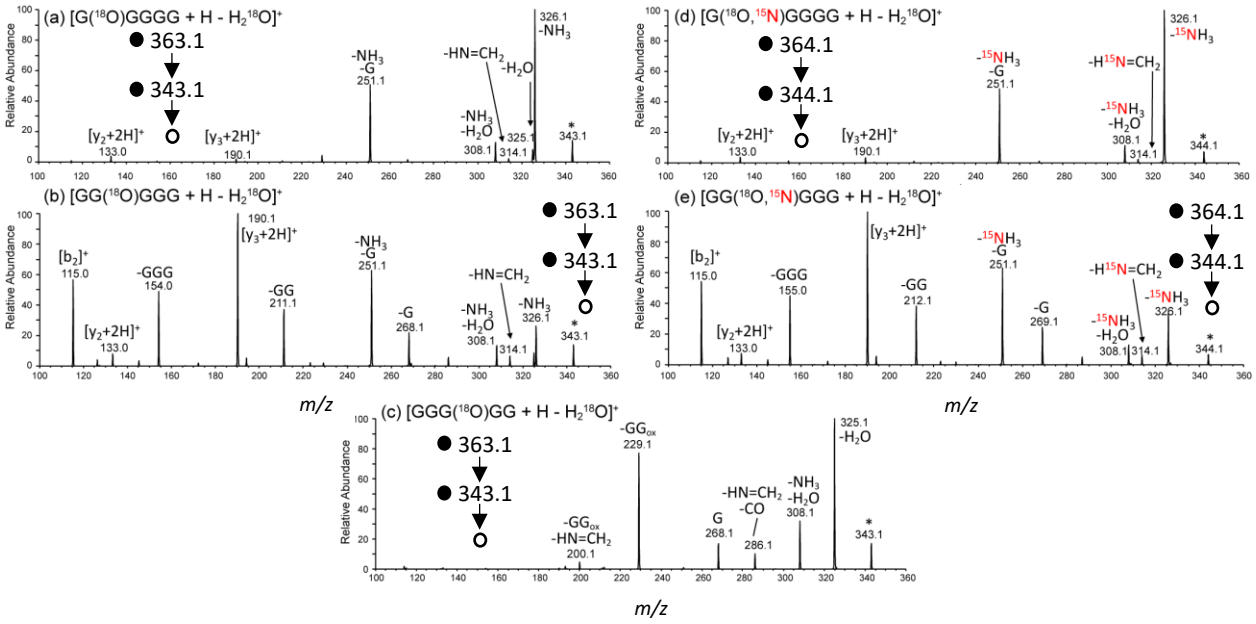


Figure 4.4. CID spectra of the (a) $[b_6]^+_I$, (b) $[b_6]^+_{II}$, and (c) $[b_6]^+_{III}$ ions. Spectra (d) and (e) have both ^{18}O- and ^{15}N-labelling in the first and second residue, respectively.

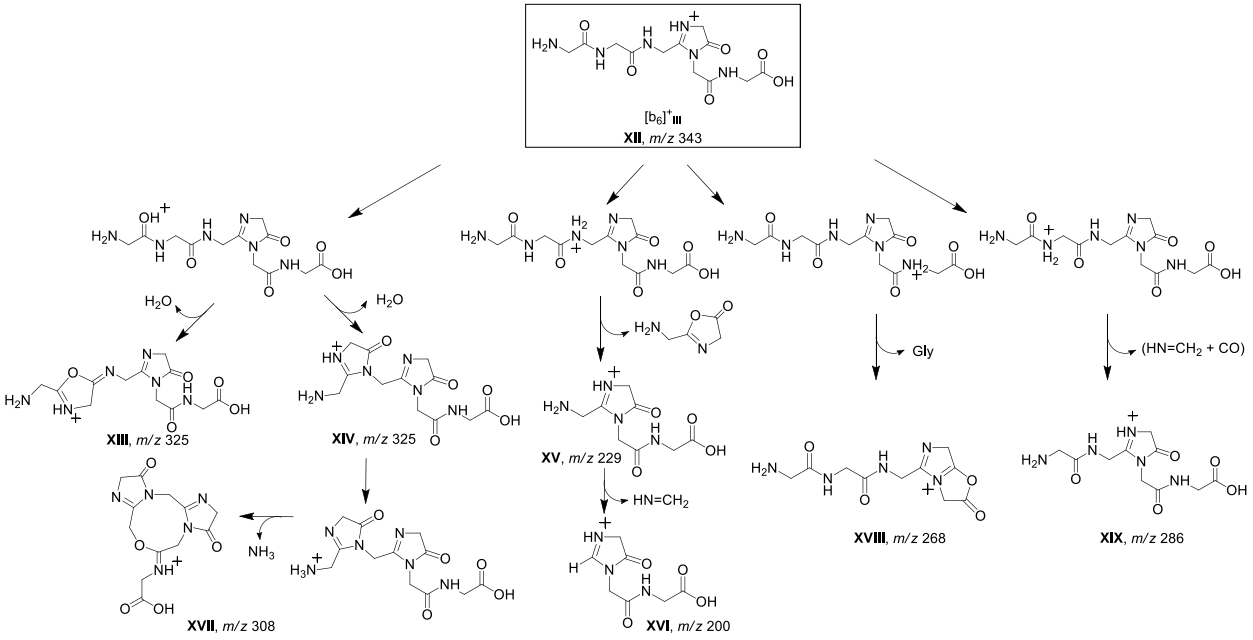
The CID spectra of the $[b_6]^+$ ions formed by water loss from the first, second and third amide groups of protonated hexaglycine are not identical, as shown by the different product ions (Figures 4.4) and their very different energy-resolved curves (Appendix B.18, 19 and 20). The dominant fragmentation pathway of the $[b_6]^+_I$ ion, at all collision energies, is the loss of ammonia (m/z 326) followed by losses of (NH_3 + glycine, m/z 251) and ($NH_3 + H_2O$, m/z 308); formations of $[y_3 + 2H]^+$ and $[y_2 + 2H]^+$ (m/z 190 and 133) are very minor dissociation channels (Figures 4.3a). The fragmentation of this $[b_6]^+_I$ ion is somewhat similar

to that of the $[b_5]^+_I$ of protonated pentaglycine in that fragmentation at the N-terminus dominates over amide bond cleavage at the C-terminus. However, the major difference is that there is little evidence of cleavage of the C_α–C bond at the N-terminus (loss of methanimine, m/z 314) that is the dominant pathway by which both the $[b_5]^+_I$ ion of protonated pentaglycine and the $[b_4]^+_I$ ion of protonated tetraglycine fragment. The CID spectrum of the $[b_5]^+_I$ ion showed NH_3 loss as a major pathway, but not the dominant one (Figure 4.2(a)), whereas the spectrum of the $[b_4]^+_I$ ion had the loss of NH_3 as only a minor channel.[23] Apparently, as the peptide length is increased, loss of ammonia becomes a more dominant pathway for the $[b_n]^+_I$ ions. ^{15}N-labelling established that the ammonia loss from the $[b_6]^+_I$ ion is from the first residue, the residue from which water was initially lost (Figure 4.2 (d)), reminiscent of the behaviour of the $[b_5]^+_I$ ion from pentaglycine .

The CID spectrum of the $[b_6]^+_I$ ion can be largely explained by initial formation of an imidazolone and the fragmentation pathways deduced for the $[b_5]^+_I$ ion, can be used to explain the loss of ammonia, as seen in pathway (ii) in Scheme 1. Ammonia loss is also followed by loss of a glycine molecule by formation of an oxazolone similar to structure **V**. Formation of $[y_2 + 2H]^+$ and $[y_3 + 2H]^+$ can be rationalized by amide bond cleavage, similar to pathways (iii) and (iv) in Scheme 4.1.

The CID spectra of the $[b_5]^+_{II}$ and $[b_6]^+_{II}$ ions are similar, again suggesting analogous fragmentation pathways (Figures 4.2(b) and 4.4(b)). The dominant pathway for the $[b_6]^+_{II}$ ion is cleavage of the third amide bond resulting in formation of an ion at m/z 154 that has lost GlyGlyGly along with generating the complementary ion $[y_3 + 2H]^+$ (m/z 190); this is analogous to formation of the m/z 154 ion (loss of GlyGly) and the complementary $[y_2 + 2H]^+$ ion from the $[b_5]^+_{II}$ ion of protonated pentaglycine. Loss of ammonia is a more minor channel in the dissociation of $[b_6]^+_{II}$ than for the $[b_6]^+_I$ ion, with ^{15}N-labelling again showing that the ammonia lost is from the same residue from which the water was initially lost (Figure 4.4(e)). This is consistent with initial formation of an imidazolone at the second amide followed by conversion to the $[b_6]^+_I$ ion probably by via the same mechanism as found for $[b_4]^+_{II} \rightarrow [b_4]^+_I$ Scheme 3.1 in Chapter 3. Amide

bond cleavage near the C-terminus is dominant over the interconversion mechanism for both $[b_5]^+_{II}$ and the $[b_6]^+_{II}$ ions. One major difference between the CID spectra of the $[b_6]^+_{II}$ and $[b_5]^+_{II}$ ions is that the former shows only a hint of a loss of $HN=CH_2$, whereas this pathway is the second most dominant for $[b_5]^+_{II}$. A possible explanation may be that as ammonia loss becomes more prevalent as the peptide chain length is increased, then loss of $HN=CH_2$ becomes less competitive. To conclude, the fragmentation pathways of the $[b_6]^+_{II}$ ion follow similar fragmentation pathways of $[b_5]^+_{II}$ as described in Scheme 4.1.



Scheme 4.2. Proposed mechanisms for the fragmentations of the $[b_6]^+_{III}$ ions.

The CID spectrum of the $[b_6]^+_{III}$ ions formed by loss of water of the third amide of protonated hexaglycine (Figure 4.4(c)) is very different from those of the isomeric $[b_6]^+_{I}$ ions and $[b_6]^+_{II}$ ions, indicating that there is no interconversion prior to dissociation. The major products in the dissociation of $[b_6]^+_{III}$ in decreasing order of abundance are:

An ion at m/z 325, corresponding to the loss of water. The $[b_6]^+_{III}$ ion derived from GlyGlyGly(^{18}O)GlyGlyGly-OMe had an almost identical CID spectrum (Figure 4.5) and exhibited no loss of methanol, indicating that the loss of water is from an amide group. The water lost probably comes from

one of the first two amide groups, forming either an oxazole (structure **XIII** in Scheme 4.2) or a second imidazolone (structure **XIV).** Based on calculations for the protonated tetraglycine, the imidazolone route forming **XIV** is likely the preferred one.[22]

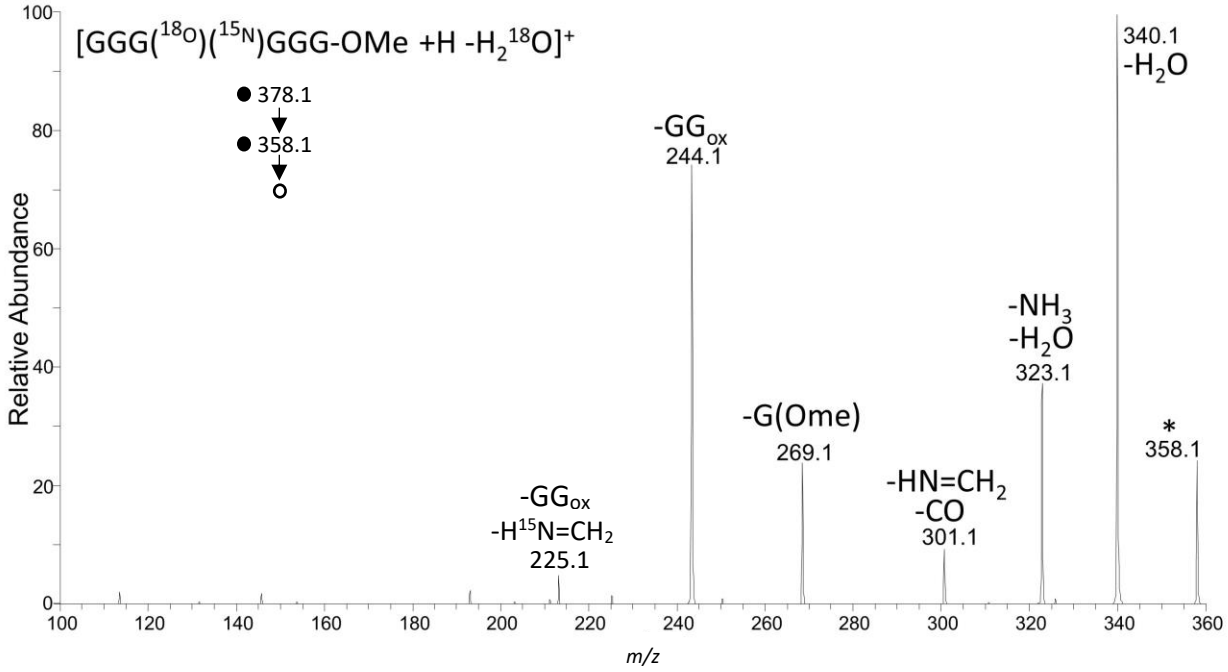


Figure 4.5. CID spectrum of the $[b_6]^+_{III}$ ion after esterification.

An ion at m/z 229 indicating the loss of diglycine residue in the form of an oxazolone ($GlyGly_{ox}$, 114 Da) from the N-terminus (Scheme 4.2). The high abundance of this ion suggests that the structure of the $[b_6]^+_{III}$ ion has the first two residues intact and that the loss of water in forming this ion must have involved nucleophilic attack on the third carbonyl oxygen by a heteroatom on a residue located closer to the C-terminus. Loss of $GlyGly_{ox}$ then is initiated by a 1,4-proton transfer from the imidazolone ring to the second amide nitrogen followed by nucleophilic attack by the first amide oxygen to the amide carbon thereby facilitating amide bond cleavage. The resulting ion, structure **XV,** has the same structure as the $[b_4]^+_I$ ion formed by loss of water from the first amide of protonated tetraglycine.[23] The ion at m/z 200 (structure

XVI), present in low abundance in the CID spectrum of the $[b_6]^+_{III}$ ion, is the product of the dominant fragmentation pathway for the $[b_4]^+_I$ ion and provides further evidence for the imidazolone structure **XV**.

An ion at m/z 308 corresponding to the loss of (H_2O + NH_3). As there is no detectable loss of only NH_3 and MS^4 on the m/z 325 ion gave predominantly loss of NH_3 (Figure SX), it is probable that the m/z 308 ion is the product of sequential losses of water followed by ammonia. The energy-resolved curves for the $[b_6]^+_{III}$ ion support this analysis (Appendix B.20). After the loss of water (structure **XIV**), NH_3 loss is initiated by a 1,4-proton transfer from the imidazolone to the N-terminal amine. This is followed by nucleophilic attack by the fifth amide oxygen to the C_α of the first residue to form tricyclic ring structures **XVII**.

An ion at m/z 268 is the result of a loss of a glycine molecule (75 Da) from the C-terminus and it is further confirmed by the CID spectrum of the ester (Figure 4.5) in which loss of Gly-OMe was observed. This fragmentation is initiated by protonation on the fifth amide nitrogen, thus weakening the C–N bond, followed by nucleophilic attack by the amide oxygen of the imidazolone ring forming a bicyclic structure, structure **XVIII**.

An ion at m/z 286 is the $[b_5]^+_{II}$ **ion**, (structure **II)**. To form this ion there is a loss of (CO + $HN{=}CH_2$) from the N-terminus, initiated by proton transfer to the first amide nitrogen followed by amide bond cleavage.

Fragmentation of an oxazolone structure $[b_n]^+$ ion.

Another kind of $[b_5]^+$ isomer was formed by the loss of glycine methyl ester from a hexapeptide derivative $[GlyGlyGly(^{18}O,^{15}N)GlyGlyGly\text{-}OMe + H]^+$. This results in a $[b_5]^+$ ion possessing the conventional oxazolone structure at the C-terminus, namely $[b_5]^+_V$. The base peak in the CID spectrum of this ion is for the loss of water, with the ratio of the abundances of the product ions at m/z 271 ($-H_2O$) and 269 ($-H_2{}^{18}O$) being exactly 4:1 (Figure 4.6a). Similarly, the ratio of the abundances ions of the $[a_5]^+$ at m/z 261 (-CO) and 259 ($-C^{18}O$) is 4:1. These ratios indicate that there is complete sequence scrambling of the $[b_5]^+_V$ ion prior to

dissociation.[45,46] The ion at m/z 244 is the $[a_5^*]^+$ ion, corresponding to the product of loss of (CO + NH_3) while that at m/z 241 corresponds to the loss of ($C^{18}O$ + $^{15}NH_3$); the loss of both labelled atoms together is consistent with the accepted mechanism by which $[a_n^*]^+$ ions are formed.[29,30,47,48] The ions at m/z 243 (losses of H_2O + CO) and 241 (losses of H_2O + $C^{18}O$) are also associated with the loss of CO after water loss and are either from the oxazolone at the C-terminus or from the imidazolone at the N-terminus. The ion at m/z 242 corresponds to water loss from the first amide group followed by loss of $NHCH_2$, which is consistent with fragmentation of 4-imidazolones formed at the first residue.[23]

The $[b_6]^+_{VI}$ ion, created by loss of methanol in the CID of [GlyGlyGly(^{18}O,^{15}N)GlyGlyGly-OMe + H]$^+$, loses water as the dominant fragmentation pathway, with the loss of a second water being the most next abundant (Figure 4.6b). The ratio of the abundances of the $([b_6]_{vi} - H_2O)^+$ ions for H_2O and $H_2{}^{18}O$ is a perfect 5:1, again showing complete sequence scrambling in the $[b_6]^+_{VI}$ ion. There was no product ion corresponding to a loss of CO (28 Da) and the ion at m/z 286, the product of the combined loss of (CO + $HN=CH_2$), was less than 2% indicating that the dissociation pathway expected for a C-terminal oxazolone, $[b_6]^+ \rightarrow [a_6]^+ \rightarrow [b_5]^+$ did not occur. Apparently as the peptide length increases from a $[b_5]^+_V$ to a $[b_6]^+_{VI}$, CO loss is no longer a competitive pathway for the oxazolones and water loss is the only pathway.

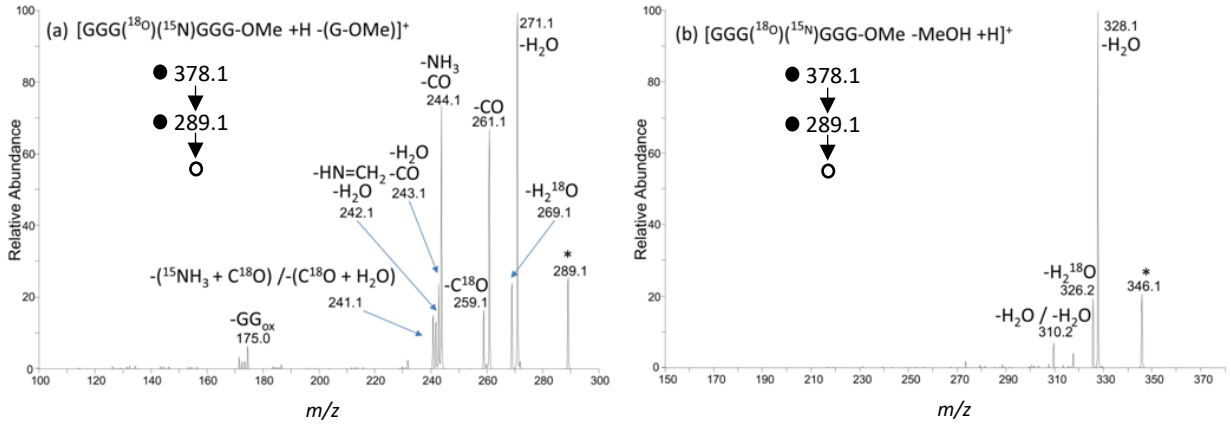


Figure 4.6. CID spectra of the (a) $[b_5]^+_V$ and (b) $[b_6]^+_{VI}$ ions with both ^{18}O- and ^{15}N-labelling on the third residue and esterification at the C-terminus.

Protonated polyglycines greater than 6 residues

Ions $[Gly_n + H]^+$ (n = 7-9), all lose a water molecule at low collision energies and this pathway becomes more dominant as the peptide length is increased. For $[Gly_7 + H]^+$ cleavage at the fifth amide bond gives the most abundant product (the $[b_5]^+$ ion), but the $[b_6 - H_2O]^+$ and $[b_7]^+$ ions, both products formed by water losses from an amide, are the second and fourth most abundant . Furthermore, the spectrum of $[Gly_7\text{-OMe} + H]^+$ had essentially the same products as that of $[Gly_7 + H]^+$ with the $[b_7]^+$ ion retaining the ester group, thereby providing evidence that in forming the $[b_7]^+$ ion water was lost from an amide group. For $[Gly_8 + H]^+$ the $[b_5]^+$ ion is again the most abundant product, followed by three ions, $[b_6 - H_2O]^+$, $[b_7 - H_2O]^+$ and $[b_8]^+$, that all involve water loss from an amide. Similarly in the spectrum of $[Gly_9 + H]^+$ four of the five most abundant ions , $[b_9]^+$, $[b_7 - H_2O]^+$, $[b_6 - H_2O]^+$, $[b_5]^+$ and $[b_8 - H_2O]^+$, all involve the loss of a water molecule.

Larger $[b_n]^+$ ions

We have not used labelling to examine the numerous possible isomers of the $[b_n]^+$ ions (n = 7-9). Nevertheless, the CID spectra of these ions are revealing. In the CID spectrum of the $[b_7]^+$ ion from Gly_7, loss of water is the base peak, followed by losses of 114 Da ($GlyGly_{ox}$), 75 Da (Gly), 171 Da ($GlyGlyGly_{ox}$) and 35 Da ($NH_3 + H_2O$). Again there was no loss of only CO and the abundance of the $[b_6]^+$ ion was very low, less than 5%. The spectrum of the $[b_7]^+$ ion from $[Gly_7\text{-OMe} + H]^+$ was very similar, again showing loss of water and not methanol, showing that the water loss was not from the carboxy group. Losses of $GlyGlyGly_{ox}$ and $GlyGly_{ox}$ are from the N-terminus and the loss of Gly is from the C-terminus indicating that the initial water loss in forming the $[b_7]^+$ ion must come predominantly from one of the central residues, probably the fifth amide.

The major product ions in the dissociation of the $[b_8]^+$ ions derived from $[Gly_8 + H]^+$, in order of decreasing abundance, are due to losses of 75 Da (Gly), 18 Da (H_2O), 132 Da (GlyGly), 35 Da ($NH_3 + H_2O$)

and 171 Da ($GlyGlyGly_{ox}$) , two from the C-terminus and one from the N-terminus, again suggesting that one of the central residues (4^{th}, 5^{th} or 6^{th}) is the origin of the water molecule that was lost.

For the $[b_9]^+$ ion derived from $[Gly_9 + H]^+$ the major dissociation products, in order of decreasing abundance, correspond to losses of 132 Da (GlyGly), 75 Da (Gly), 18 Da (H_2O) and 189 Da (GlyGlyGly), all losses from the C-terminus of the peptide except that involving water. We note that there are losses of GlyGlyGly from the $[b_9]^+$ ion, GlyGly from $[b_8]^+$ and only Gly from $[b_7]^+$. These all suggest that water loss is from the same central residues in all the $[b_n]^+$ ions.

The tendency for the loss of water from $[b_n]^+$ ions of $[Ala_n + H]^+$ (n =3-15) increases with chain length and has been noted previously in ions derived from polyalanine.[29] These $[b_n]^+$ ions also lost CO and ($CO + NH_3$), products not found in the dissociations of $[b_n]^+$ ions derived directly from $[Gly_n + H]^+$ ions (n = 5-9). This difference is probably because some of the $[b_n]^+$ ions derived from polyalanine have oxazolone structures, whereas those from $[Gly_n + H]^+$ are imidazolones. Note that the $[b_5]^+_{\mathbf{v}}$ ion derived from $[Gly_6 + H]^+$ is an oxazolone and this was found to lose both H_2O and CO, with the former being in the higher abundance (Figure 4.6a).

4.3 Conclusion

^{18}O-labelling of the amide oxygens revealed that loss of water is predominantly from the first and second amides of protonated pentaglycine. The resulting $[b_5]^+_{\mathbf{I}}$ and $[b_5]^+_{\mathbf{II}}$ ions did not have identical CID spectra; for the $[b_5]^+_{\mathbf{I}}$ ion, the loss of methanimine was the major pathway while loss of GlyGly was the dominant pathway for the $[b_5]^+_{\mathbf{II}}$ ion. Theoretical calculations suggested that the $[b_5]^+_{\mathbf{I}}$ and $[b_5]^+_{\mathbf{II}}$ ions can interconvert but at low collision energies the preferred pathway for the $[b_5]^+_{\mathbf{II}}$ ion is loss of GlyGly via *pathway (v)*.

^{18}O-labelling of the amide oxygens revealed that loss of water is predominantly from the first, second and third amides of protonated hexaglycine. The resulting $[b_6]^+$ ions did not have identical CID spectra and there is little evidence for interconversion prior to dissociation. Ammonia loss is a major pathway for both $[b_5]^+_{\mathbf{I}}$ and $[b_6]^+_{\mathbf{I}}$ but loss of methanimine is the dominant channel for $[b_5]^+_{\mathbf{I}}$ and is quite minimal for $[b_6]^+_{\mathbf{I}}$. The $[b_6]^+_{\mathbf{III}}$ ion has the simplest CID spectrum of these isomers and shows mainly water loss.

The $[b_5]^+_{\mathbf{V}}$ and $[b_6]^+_{\mathbf{VI}}$ ions, which are oxazolones showed different fragmentation pathways from those of the imidazolones. These ions also exhibited perfectly sequence scrambling in their CID spectra, further supporting that the structures are oxazolones. Apparently as peptide length is increased, the loss of CO from $[b_n]^+_{\mathbf{n}}$ ions is less competitive and water loss becomes the preferred pathway.

The structures of the $[b_n]^+$ ions formed by water loss from $[G_n + H]^+$ (n = 7-9) were not determined. However, the losses of $[y_m + 2H]^+$ in their CID spectra suggested the formation of imidazolones on the amide backbone rather than an oxazolone at the C-terminus. The losses of the $[y_m + 2H]^+$ from the C-terminus and $(G_k)_{ox}$ at the N-terminus suggest that the loss of water is from the central residues. The losses of various $(G_k)_{ox}$ and $[y_m + 2H]^+$ ions suggest that several amide oxygens can be sources for water loss.

Chapter 5 Collision-induced dissociation of protonated tetraglycine derivatives with an alanine or proline substitution after water loss, isomers of $[b_4]^+$

5.1 Loss of water from $[XGlyXGly + H]^+$; X = alanine or glycine

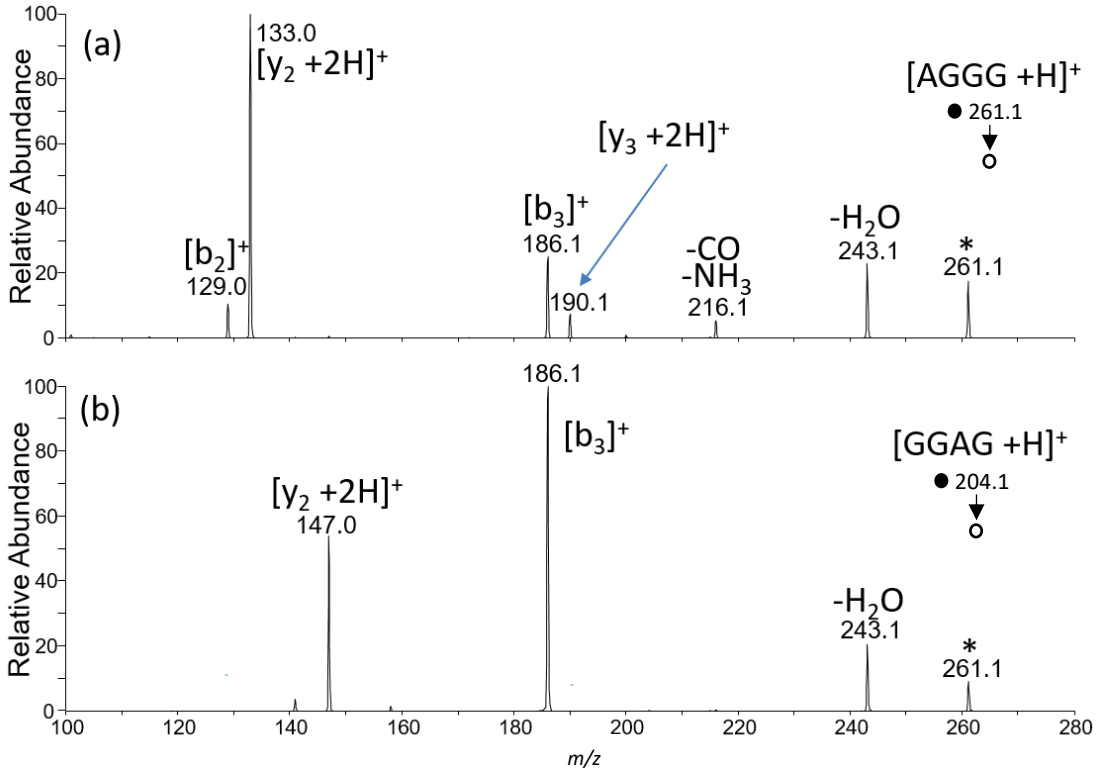


Figure 5.1 CID spectra for protonated (a) AGGG and (b) GGAG. CID spectra (a) and (b) were collected at a normalized collision energy of 20.

To determine the effect of substituents on the loss of water, alanine was substituted for a glycine of GlyGlyGlyGly. Water loss from the first amide is the dominant pathway from $[GlyGlyGlyGly + H]^+$.[34–36] In the CID of $[AlaGlyGlyGly + H]^+$ and $[GlyGlyAlaGly + H]^+$ the abundance of water loss was approximately 50% of that observed for $[GlyGlyGlyGly + H]^+$, at various collision energies. The abundance of water loss from $[AlaGlyAlaGly + H]^+$ was a minor pathway, as seen in Appendix B.22. Water loss from $[GlyGlyGlyGly + H]^+$ is predominantly from the first amide oxygen; assuming that the pathway for water loss is still predominantly from the first amide oxygen after an alanine substitution, then the α-methyl group appears

to deter water loss. Since peptides with a methyl group in either the receiving group (the first amide) or the nucleophile (the third amide) had similar abundances in the loss of water, the bulkier side chain in either position has a similar effect. Energy-resolved curves in Figure 5.2 show that water loss from $[AlaGlyGlyGly + H]^+$ and $[GlyGlyAlaGly + H]^+$ were similar at all collision energies. This suggests that there is an equal impact by introducing a substituent in either the first or third position. Furthermore, adding a methyl substituent in *both* the receiving group and the nucleophile positions as in $[AlaGlyAlaGly + H]^+$, further reduced the abundance of water loss (Appendix B.22). This is consistent with the idea that a bulkier side chain in the nucleophile and receiving group positions have an impact on water loss.

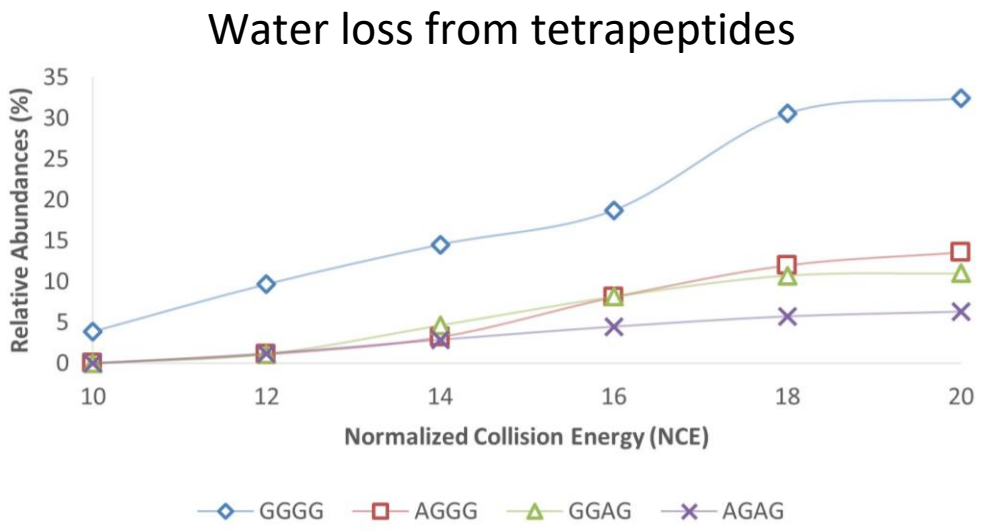


Figure 5.2 Energy-resolved curves comparing the water loss from tetrapeptides.

A methyl substituent should have both electronic and steric effects. Introducing the methyl group into the receiving group, $[AlaGlyGlyGly + H]^+$, would increase electron density since the methyl group is a weak electron donor making the α-carbon less susceptible to nucleophilic attack. A methyl group is also bulkier than a proton, suggesting that both steric and electronic effects play a role in reducing nucleophilic attack when introduced into the receiving group. Introducing a methyl substituent into the nucleophile group, $[GlyGlyAlaGly + H]^+$, would increase nucleophilicity by weakly increasing electron density. Steric

effects are also increased by introducing a methyl substituent. The abundance of water loss from both $[AlaGlyGlyGly + H]^+$ and $[GlyGlyAlaGly + H]^+$ were similar which suggests that increasing the nucleophilicity of the attacking group may be less significant than increasing the steric effects.

5.2 Loss of water from $[AlaGlyGlyGly + H]^+$

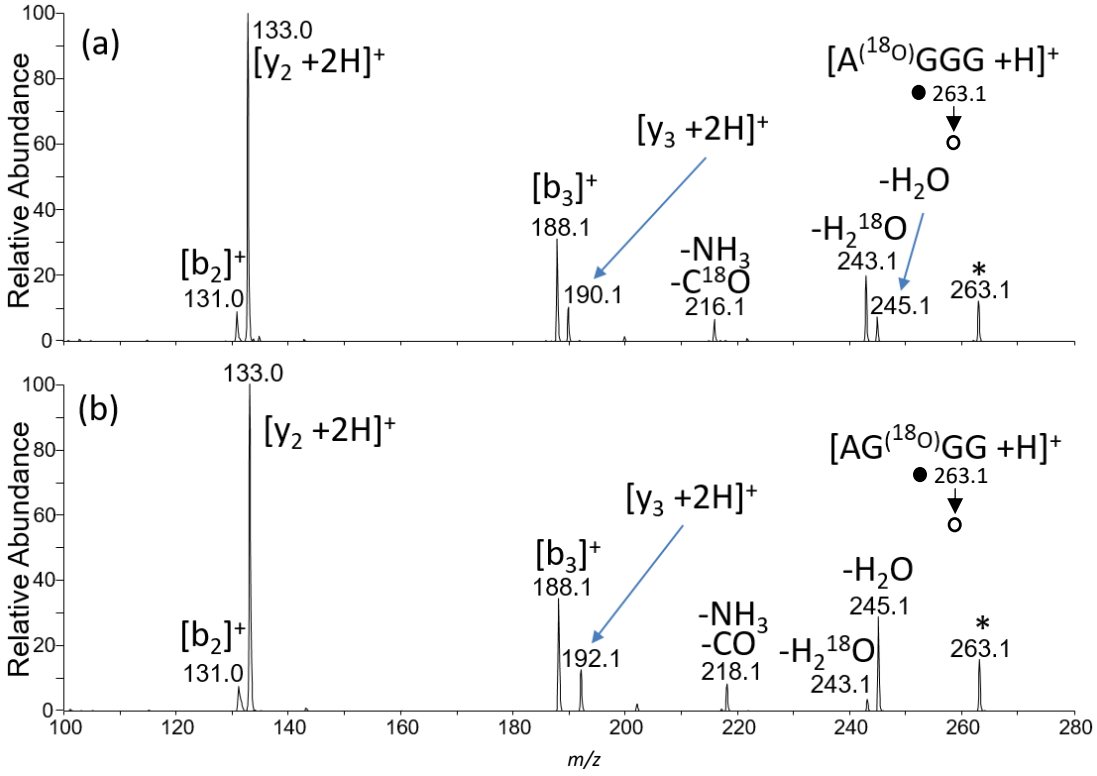
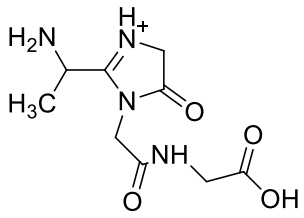


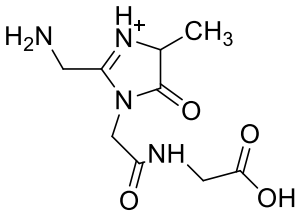
Figure 5.3 CID spectra of protonated AlaGlyGlyGly (a) with the ^{18}O labeled at the first amide oxygen (b) with the ^{18}O labeled at the second amide oxygen. CID spectra (a) and (b) were collected at a normalized collision energy of 20.

Isotopic labelling with ^{18}O showed that water loss is predominantly from the first amide. Loss from the second amide is a minor channel, as found for $[GlyGlyGlyGly + H]^+$ (Figure 5.3). The CID of $[AlaGlyGlyGly + H - H_2O]^+$ ions produced by water loss from the first and second residues follow similar fragmentation patterns as those from the analogous ions derived from $[GlyGlyGlyGly + H - H_2O]^+$. Dissociation of the ion formed by the loss of water from the first amide, the alanine residue, resulted in the loss of ethanimine as the major fragmentation pathway (Figure 5.4 (a)). Dissociation of the ion formed by loss of water from

the second amide, a glycine residue, resulted in the loss of methanimine as the major fragmentation pathway (Figure 5.4 (b)). These fragmentation pathways, loss of methanimine and ethanimine, are analogous to those found for the equivalent ions derived from [GlyGlyGlyGly + H]$^+$; this suggests the formation of an imidazolone and interconversion prior to fragmentation. Loss of ethanimine is directly from Structure **I-A** (Figure 5.4 (a)) and loss of methanimine as shown in (Figure 5.4 (b)) suggests the formation of the imidazolone followed by conversion into structure (**I-B)** following the mechanism found for the equivalent ion from GlyGlyGlyGly.



I-A



I-B

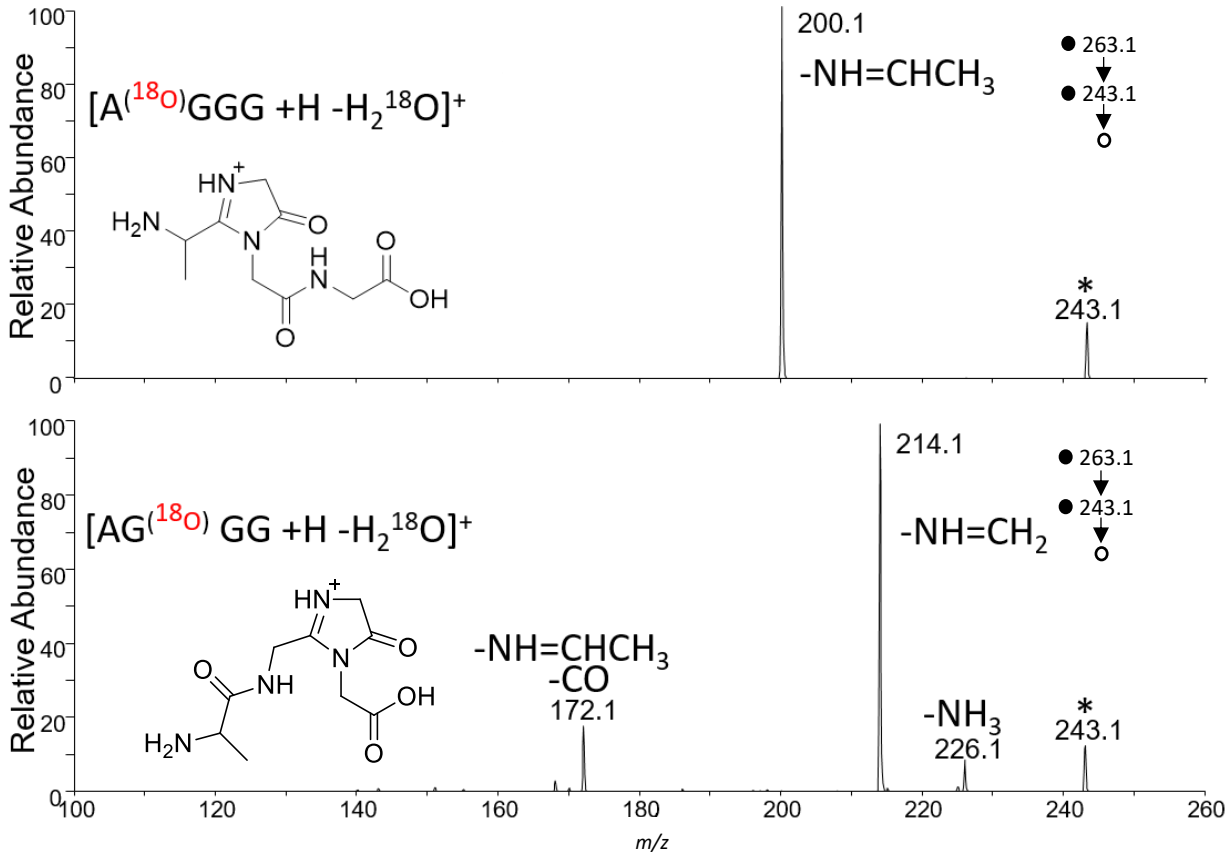


Figure 5.4 CID spectra of the nominal $[b_4]^+$ ions of AlaGlyGlyGly formed by loss of labelled oxygen from the (a) first amide and (b) second amide. CID spectra (a) and (b) were collected at a normalized collision energy of 17.

Isotopic labelling was then used to confirm the interconversion mechanism, identifying from which residue the methanimine came. ^{13}C labelling on the α-carbon was used and identified that the loss of methanimine was from the second glycine residue (Figure 5.5). This provides further evidence of interconversion prior to fragmentation, as observed previously for $[GlyGlyGlyGly + H - H_2O]^+$ (Scheme 5.1)

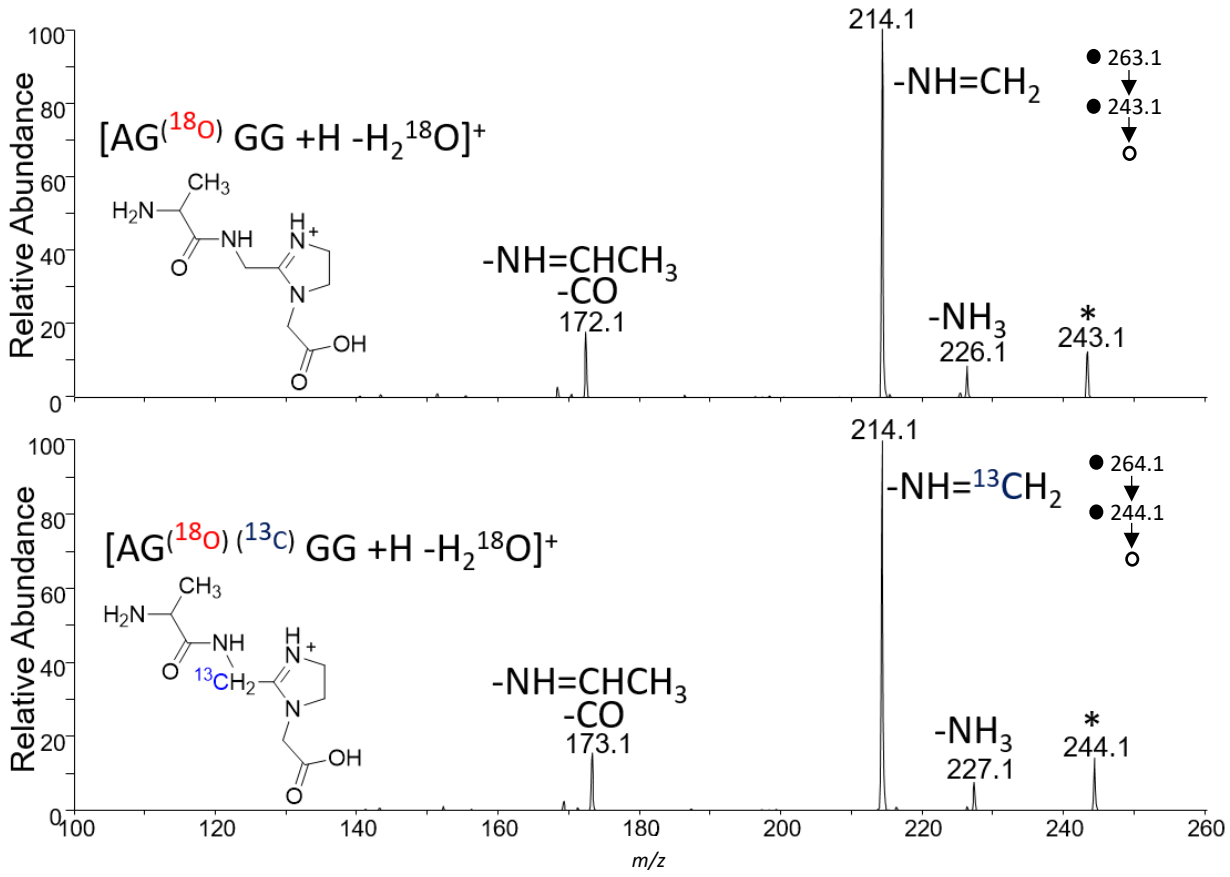
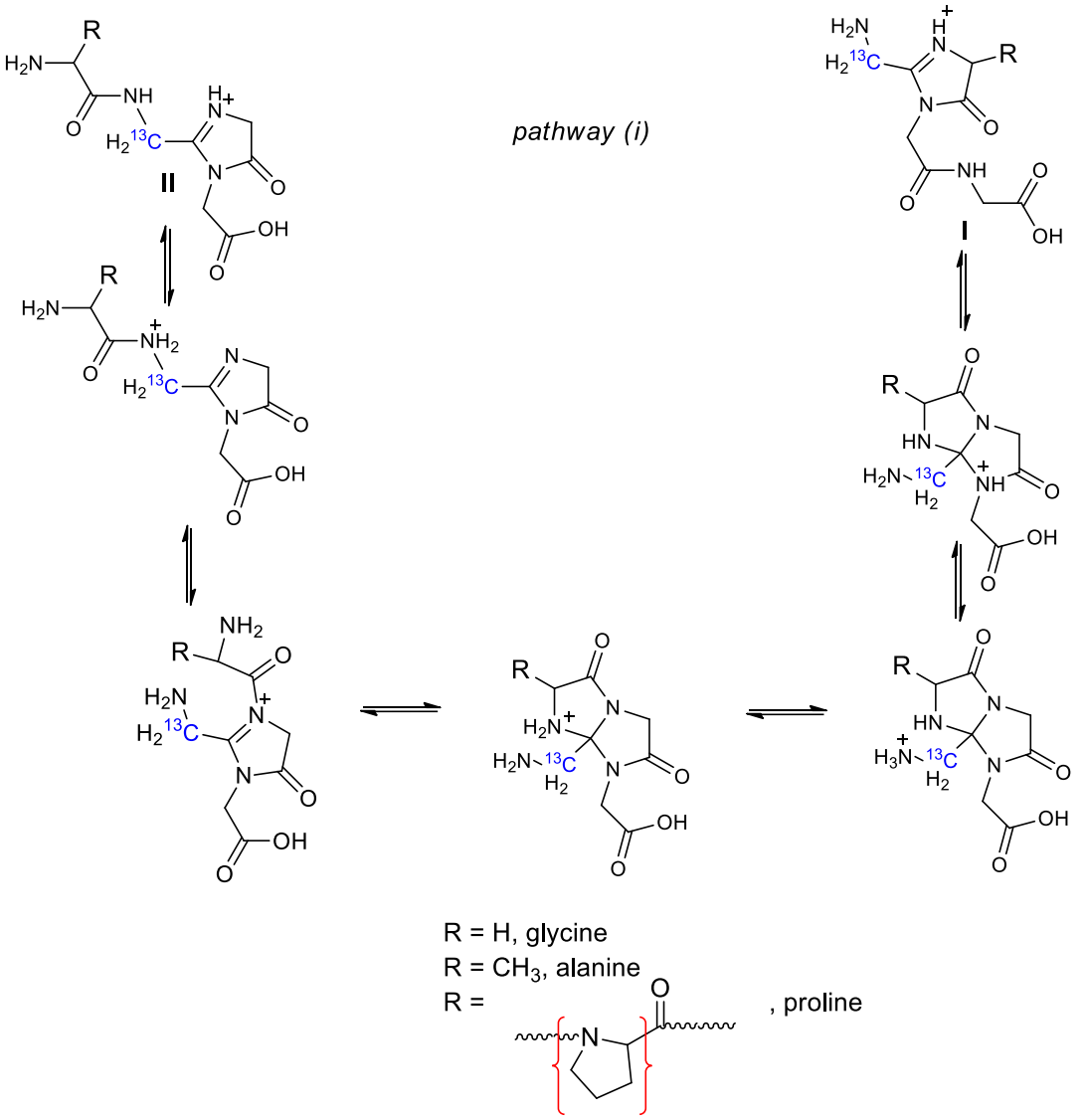


Figure 5.5 CID spectra of the nominal $[b_4]^+$ ions of AlaGlyGlyGly formed by loss of labelled oxygen from the second amide (a) with an ^{18}O label and (b) with an ^{18}O label and $^{13}C_\alpha$. CID spectra (a) and (b) were collected at a normalized collision energy of 17.



Scheme 5.1. Mechanism for the rearrangement of $[b_4]^+$ ions created by the loss of water from the 2nd amide. The tetrapeptides have three glycine residues and an alanine, a glycine, or a proline in the first position.

5.3 Loss of water from $[GlyAlaGlyGly + H]^+$

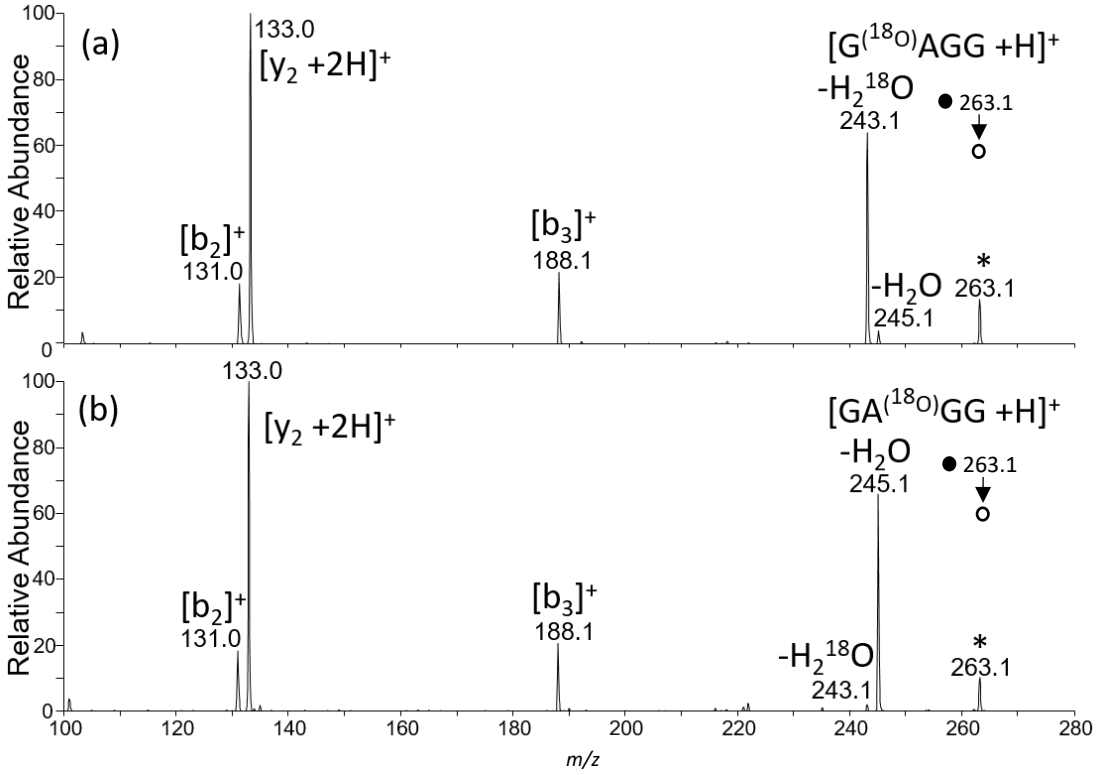


Figure 5.6 CID spectra of protonated GlyAlaGlyGly (a) with the ^{18}O label at the first amide oxygen (b) with the ^{18}O label at the second amide oxygen. CID spectra (a) and (b) were collected at a normalized collision energy of 20.

Isotopic labelling with ^{18}O showed that water loss is predominantly from the first amide and the abundances of the $[b_4]^+$ ions are much higher than those derived from $[AlaGlyGlyGly + H]^+$ and $[GlyGlyAlaGly + H]^+$. Water loss from the second amide is a minor pathway (Figure 5.6). The CID spectrum of $[Gly(^{18}O)AlaGlyGly + H - H_2{}^{18}O]^+$ showed only the loss of methanimine (Figure 5.7 (a)). ^{13}C labelling showed that methanimine is from the first residue, glycine (Figure 5.8 (a)). This fragmentation pattern is similar to that of $[GlyGlyGlyGly + H]^+$, where the major pathway is loss of methanimine.

The CID spectrum of $[GlyAla(^{18}O)GlyGly + H - H_2{}^{18}O]^+$ (Figure 5.7 (b)) showed losses of both methanimine and ethanimine. This fragmentation pattern is unlike those $[GlyGlyGlyGly + H - H_2O]^+$ and $[AlaGlyGlyGly + H - H_2O]^+$ ions, where water had been lost from the second amide. Both lost only

methanimine; consequently, [GlyAla(^{18}O)GlyGly + H - $H_2{}^{18}O$]$^+$ would have been expected to lose only ethanimine and the loss of methanimine indicates that another channel is competitive.

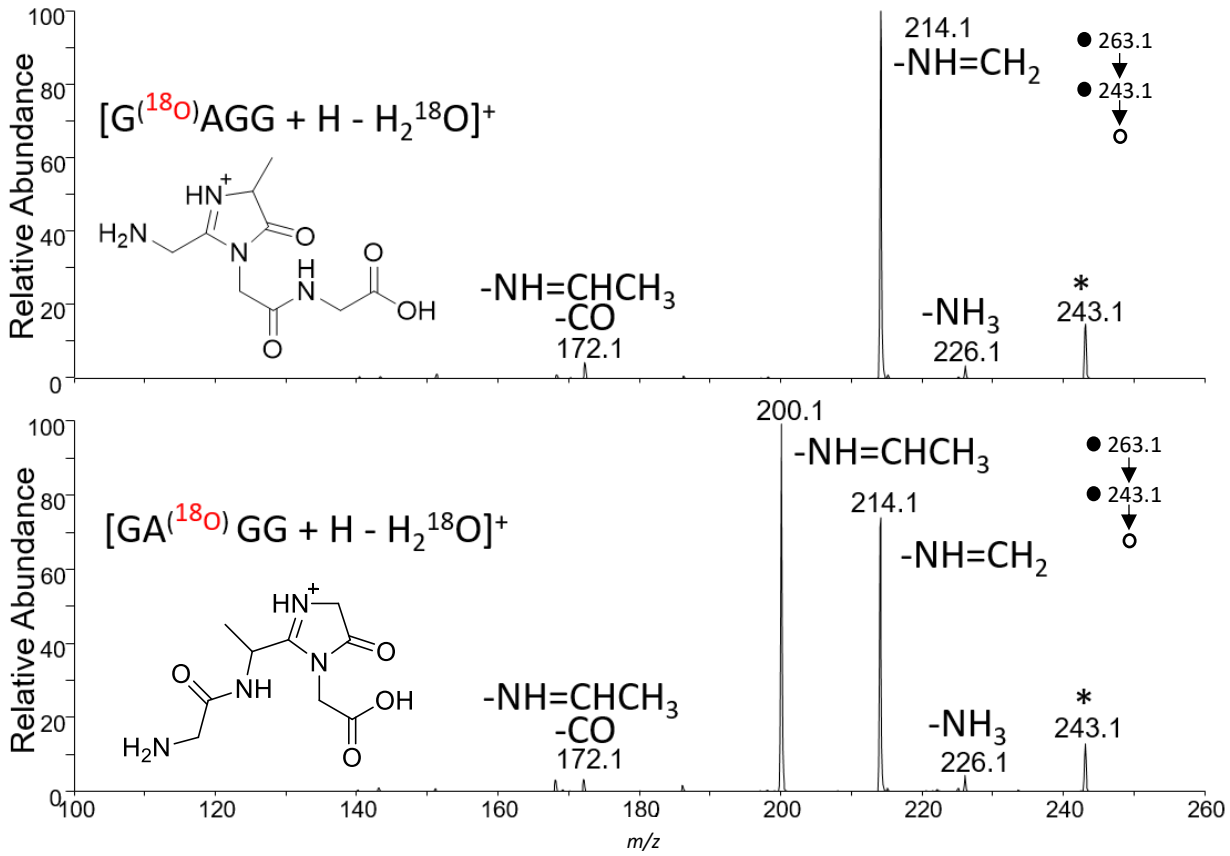


Figure 5.7 CID spectra of [b_4]$^+$ ions of GlyAlaGlyGly formed by loss of oxygen from the (a) first amide and (b) second amide. CID spectra (a) and (b) were collected at a normalized collision energy of 17.

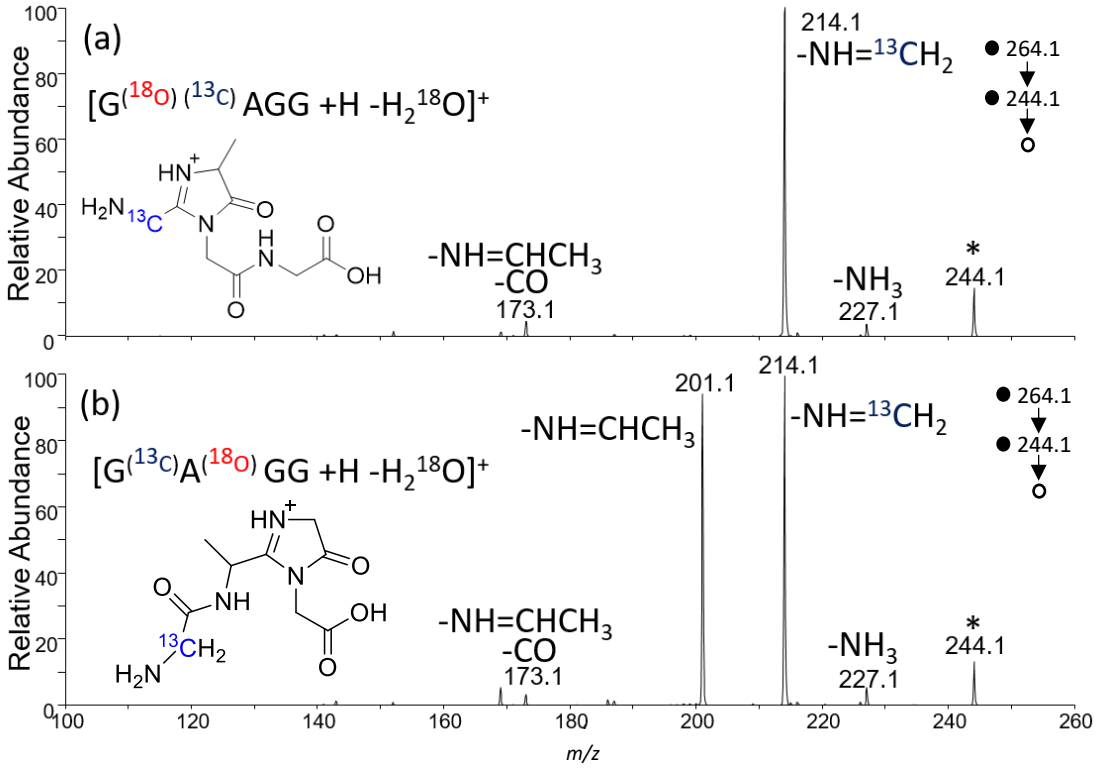
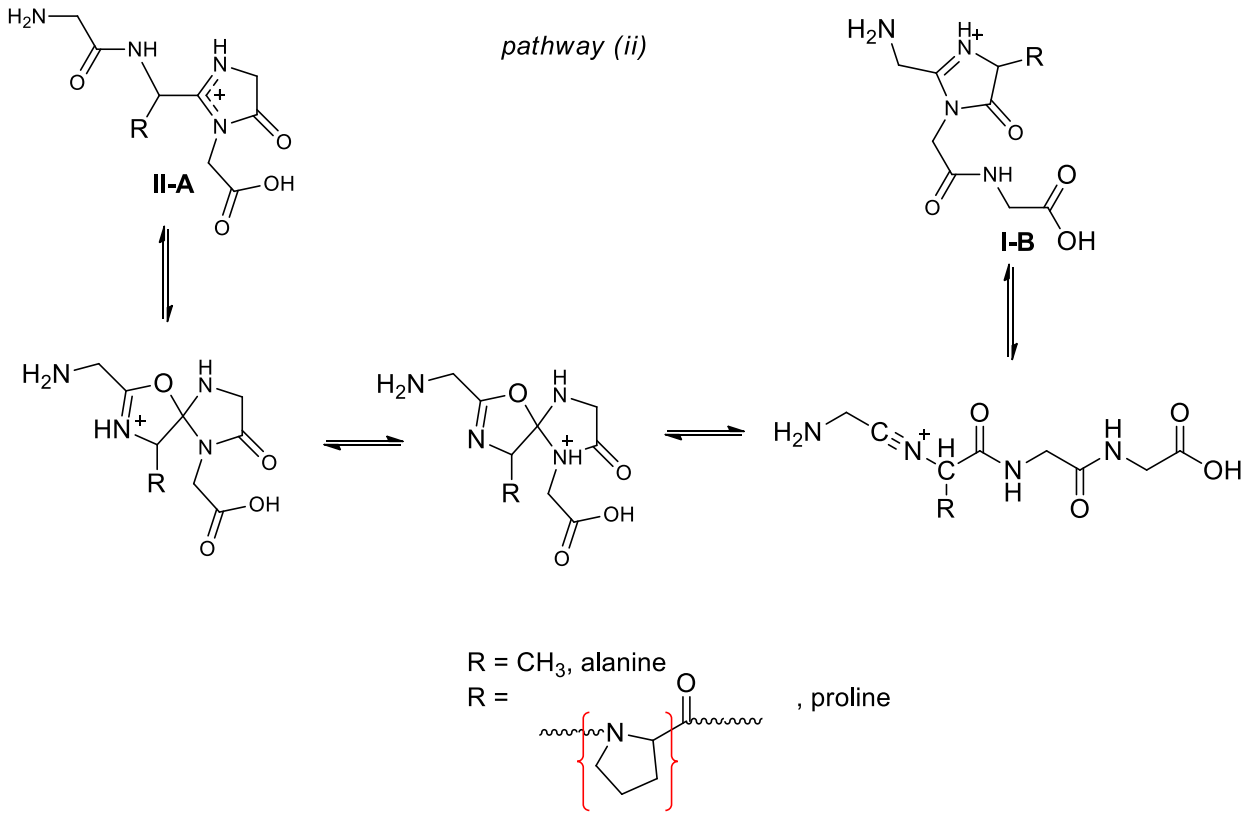


Figure 5.8 CID spectra of nominal $[b_4]^+$ ions of GlyAlaGlyGly with $^{13}C_\alpha$ formed by loss of oxygen from the (a) first amide and (b) second amide. CID spectra (a) and (b) were collected at a normalized collision energy of 17.

The fragmentation pattern of [GlyAla(^{18}O)GlyGly + H - $H_2{}^{18}O$]$^+$ is unlike any of the other fragmentation patterns from tetraglycines; ^{13}C labelling on the α-carbon was used to determine from which residue the methanimine came. ^{13}C labelling showed that loss of methanimine always came only from the first residue (Figure 5.8 (b)). Presumably, ethanimine comes from the alanine residue. The loss of ethanimine, after water loss from the second residue, suggests the formation of an imidazol-4-one created by the interconversion mechanism (pathway (i) Scheme 5.1).



Scheme 5.2 Mechanism for the rearrangement of the $[b_4]^+$ ion formed by loss of water from the second amide.

Ethanimine can be lost via the rearrangement given in Scheme 5.1 and pathway (ii) can explain the observed losses of methanimine (Scheme 5.2). Pathway (ii) is a higher energy pathway, as consistent with the energy-resolved curves of $[GlyAla(^{18}O)GlyGly + H - H_2{}^{18}O]^+$ (Figure 5.9). Formation of a 4-imidazolone is supported by the identical CID spectra of the ions formed after loss of methanimine (*m/z* 214) (Figure 5.11).

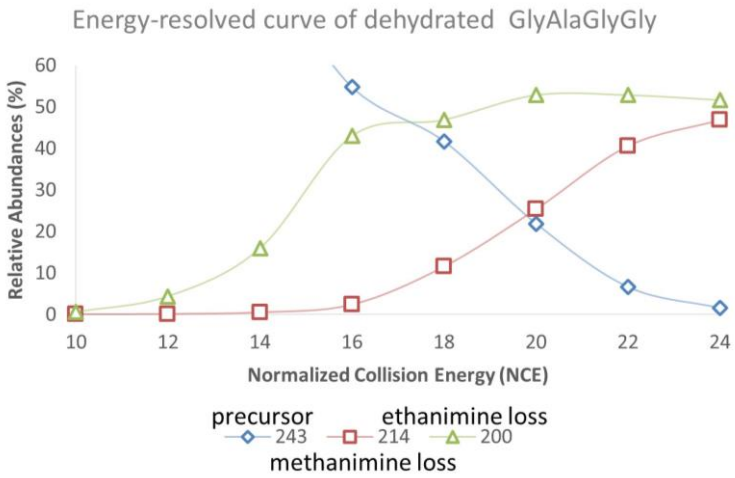
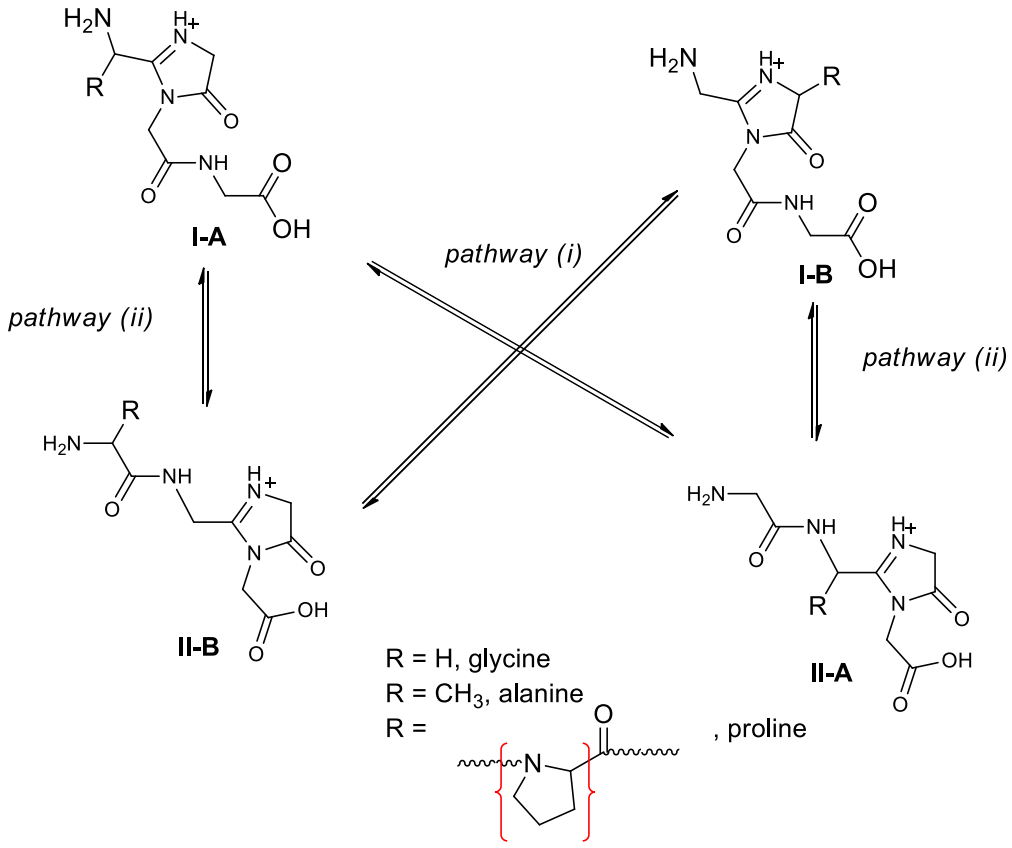


Figure 5.9 Energy-resolved curve of protonated GlyAlaGlyGly after water loss from the second residue, alanine.

The fragmentation pattern of the $[b_4]^+$ ions, formed by water loss, with three glycines and an alanine follow many fragmentation pathways of the $[b_4]^+$ ions formed by water loss from protonated tetraglycine. The pathways for loss of an imine, ammonia, amide bond cleavage in the loss of an amino acid residue and a glycine molecule from the C-terminus are described in Chapter 3 (Scheme 3.2).

To rationalize the imine losses from the N-terminus by C_α–C cleavage analogous to the formation of an $[a_1]^+$ ion from a $[b_2]^+$ ion,[62] the imines must relocate to the N-terminus in accordance with the isotopic labelling. [Ala(^{18}O)GlyGlyGly + H - $H_2{}^{18}O$]$^+$ and [Gly(^{18}O)AlaGlyGly + H - $H_2{}^{18}O$]$^+$ ions have the structures **I-A** and **I-B** and these can lose an imine directly from the N-terminus, from the residue in which water was initially lost. No interconversion mechanism is required to explain the losses observed in Figures 5.4 (a) and 5.8 (a). *Pathway (i)* describes an interconversion mechanism where the imine is initially located on the second residue and relocates to the N-terminus. [AlaGly(^{18}O)GlyGly + H - $H_2{}^{18}O$]$^+$ ions have structure **II-B** which rearranges into structure **I-B**, and the loss of methanimine is from the same residue from which water was lost. *Pathway (ii)* describes an interconversion mechanism where the loss of the imine is not from the same residue from which water was lost. [GlyAla(^{18}O)GlyGly + H - $H_2{}^{18}O$]$^+$ ions formed structure

II-A but were able to lose both methanimine and ethanimine in high abundances. Loss of ethanimine comes from the alanine residue suggesting the same kind of interconversion required previously in *pathway (i)* where structure **II-A** interconverts into structure **I-A** to allow for imine loss from the same residue in which water was initially lost. However, loss of methanimine was found to come from the first residue which requires an alternative interconversion, *pathway (ii),* to lose methanimine from the N-terminus where structure **II-A** interconverts into structure **I-B**. At high energies *pathway (i)* and *(ii)* are competitive for [GlyAla(^{18}O)GlyGly + H -$H_2{}^{18}O$]$^+$, but at low energies *pathway (i)* is dominant.



Scheme 5.3 The possible pathways for interconversion between $[b_4]^+$ ions formed from water loss.

5.4 Fragmentation of the ion at *m/z* 200 and *m/z* 214

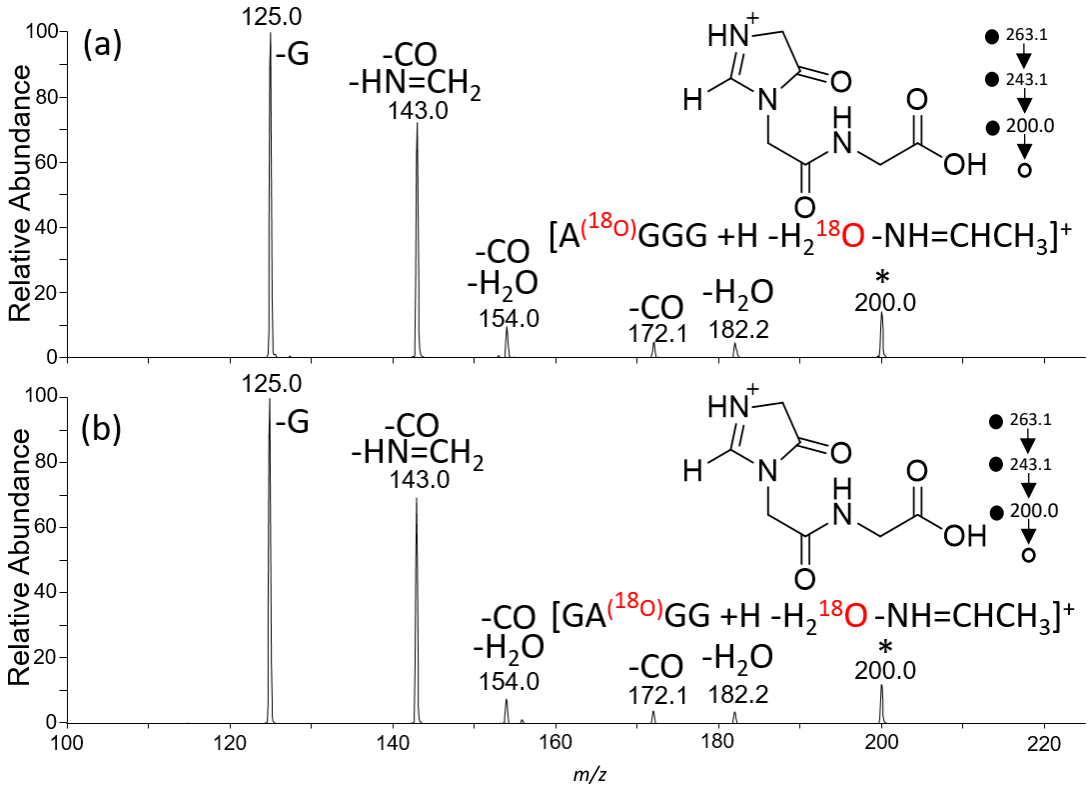


Figure 5.10 CID spectra for the dissociation of *m/z* 200 ions derived by loss of water followed by loss of ethanimine from protonated (a) AlaGlyGlyGly and (b) GlyAlaGlyGly. CID spectra (a) and (b) were collected at a normalized collision energy of 26.

The CID spectra of the *m/z* 200 ions of $[GlyAlaGlyGly + H - H_2O - HN{=}CHCH_3]^+$ and $[AlaGlyGlyGly + H - H_2O - HN{=}CHCH_3]^+$ were identical with that of $[GlyGlyGlyGly + H - H_2O - HN{=}CH_2]^+$ (*m/z* 200 ion), which suggests formation of the imidazole-4-one structure (Figure 5.10). For tetraglycine, the methanimine lost is from the residue that was C_α-C bonded to the imidazole-4-one ring (Scheme 3.2). Thus, for alanine, it is probable that the same C_α-C bond cleavage can occur for the same 4-imidazolone structure.

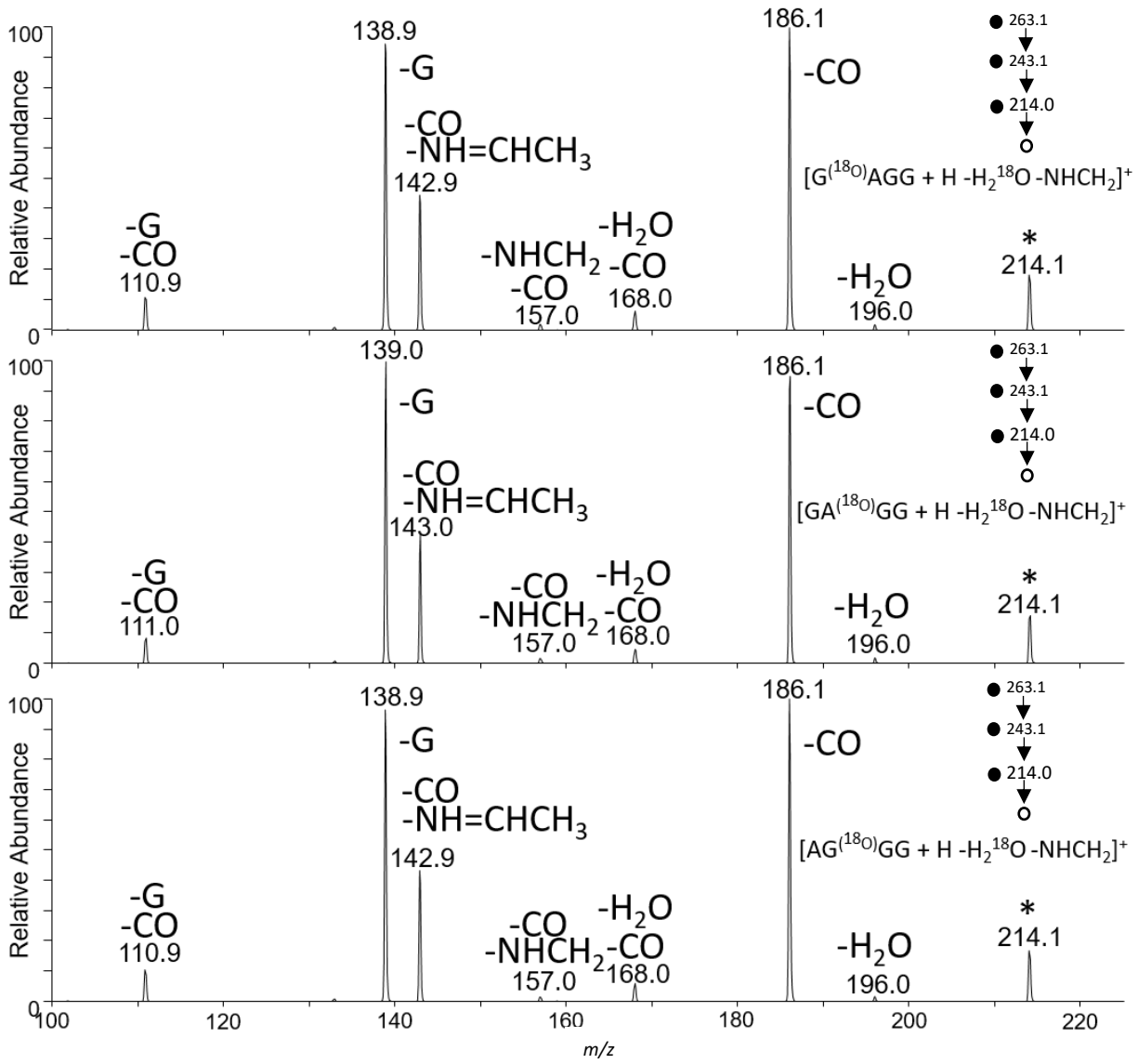
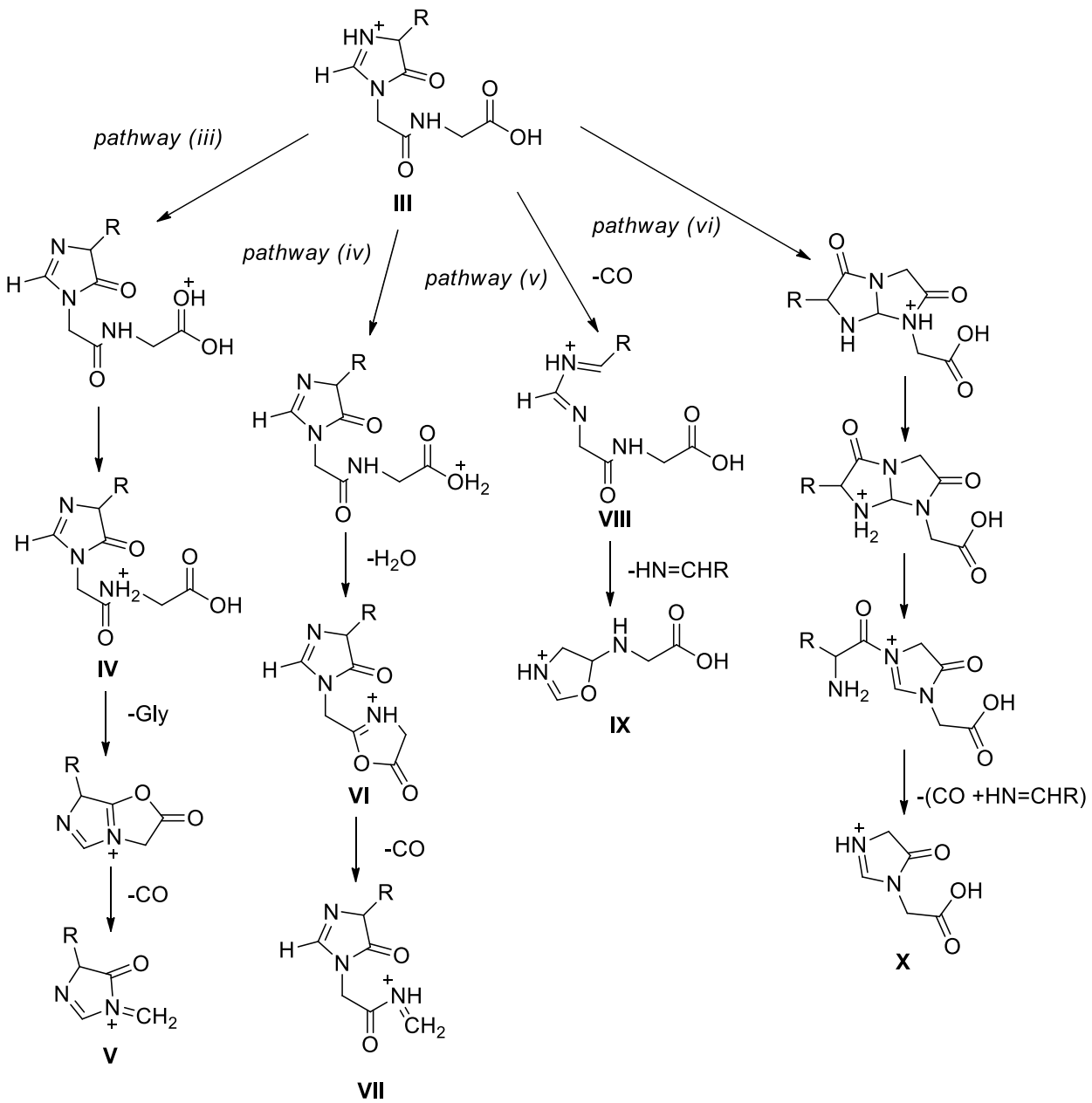


Figure 5.11 CID spectra for the dissociation of *m/z* 214 ions derived from the loss of water followed by loss of methanimine from protonated (a) GlyAlaGlyGly from the first residue, (b) GlyAlaGlyGly from the second residue and (c) AlaGlyGlyGly from the second residue. CID spectra (a), (b) and (c) were collected at a normalized collision energy of 25.

The CID spectra in Figure 5.11 are identical as well as are their energy-resolved curves (Appendix B. 22, 23 and 24). This suggests that their associated $[b_4]^+$ ions are able to interconvert into the same structure. The fragmentation pattern seen in Figure 5.11 can be rationalized using many of the same fragmentation pathways as seen with the 200 *m/z* ion, generated after loss of methanimine from [GlyGlyGlyGly + H - H_2O]$^+$. A proposed mechanism for the fragmentation of the *m/z* 214 ion is seen in

Scheme 5.4. The only notable difference between the CID spectra of [GlyGlyGlyGly + H - H_2O - $HN=CH_2$]$^+$ and of [AlaGlyGlyGly + H - H_2O - $HN=CH_2$]$^+$ and [GlyAlaGlyGly + H - H_2O - $HN=CH_2$]$^+$ is that CO loss has a higher relative abundance. The ion formed after loss of CO is an imine that in the case of an alanine residue has a methyl group that stabilizes the charge more than the hydrogen of a glycine residue. This may be why there is a larger abundance of structure **VIII** in the dissociation of ion *m/z* 214 as opposed to the corresponding ion from *m/z* 200. This is then followed by nucleophilic attack by the adjacent amide oxygen displacing an imine. For the *m/z* 200 ion loss of (CO + $HN=CH_2$) can occur in one step via pathway (vi) in Scheme 5.4. By adding an R group in the form of alanine, the structure with the double bond is stabilized. This may be why the *m/z* 214 ion would undergo ring opening via *pathway (vi)* while the *m/z* 200 ion does not.

Many of the fragmentation pathways are similar to that of the *m/z* 200 ion as seen in Scheme 5.4. Loss of water can be rationalized by formation of an oxazolone at the C-terminus (Structure **VI**). Subsequent CO loss is also observed and is typical of oxazolone fragmentation (Structure **VII**). Loss of a glycine molecule can also be explained by the formation of an oxazolone at the C-terminus in the cleavage of an amide bond (Structure I**V**). A subsequent loss of CO occurs giving Structure **V**. The fragmentation pattern of the *m/z* 214 ion, which is similar to that of the *m/z* 200 ion, suggests that this ion is a 4-imidazolone with structures only differing by the former having a methyl group.



Scheme 5.4 Mechanisms for the fragmentation of 4-imidazolones after the loss of an imine.

5.5 Loss of water from [ProGlyGlyGly + H]$^+$

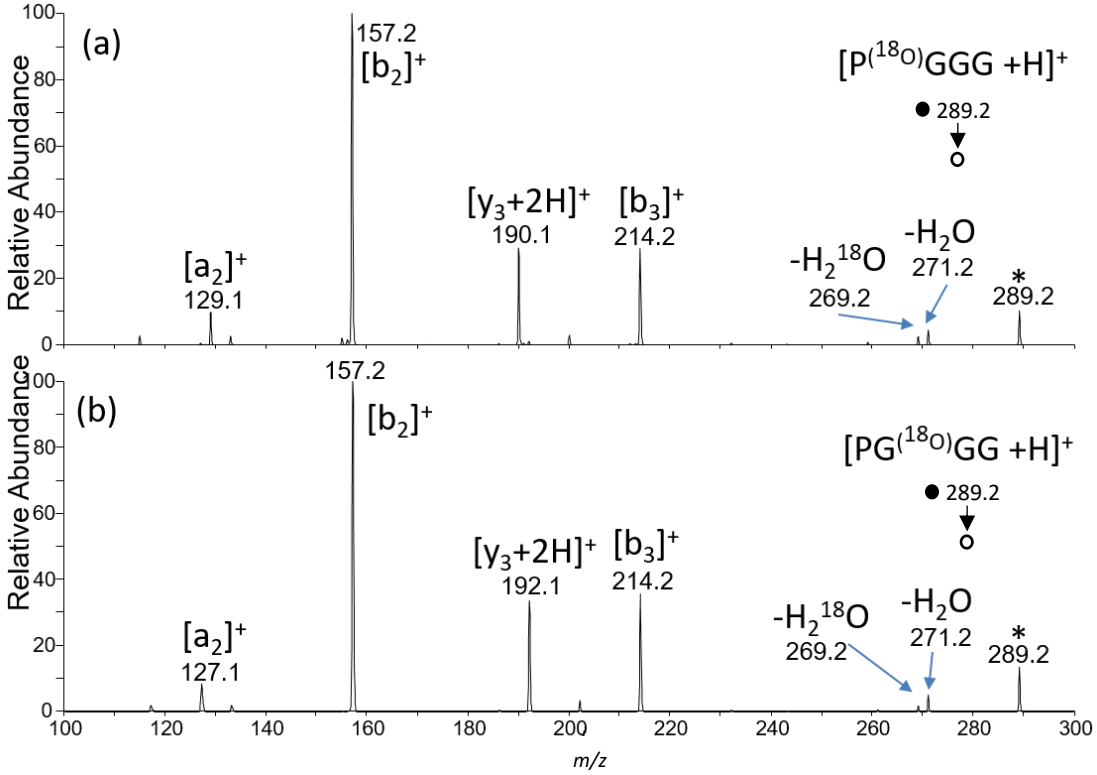


Figure 5.12 CID spectra of protonated ProGlyGlyGly (a) with the ^{18}O label at the first amide oxygen (b) with the ^{18}O label at the second amide oxygen. CID spectra (a) and (b) were collected at a normalized collision energy of 22.

Isotopic labelling with ^{18}O showed that water loss is in low abundance with considerably less than half coming from the first amide. Loss of water from the first amide is more dominant than the second but the two together do not account for all the water loss. Loss from the C-terminus is most likely the dominant pathway (Figure 5.12). The relative abundance of water loss is significantly less than what was observed from protonated tetraglycine from peptides containing an alanine.

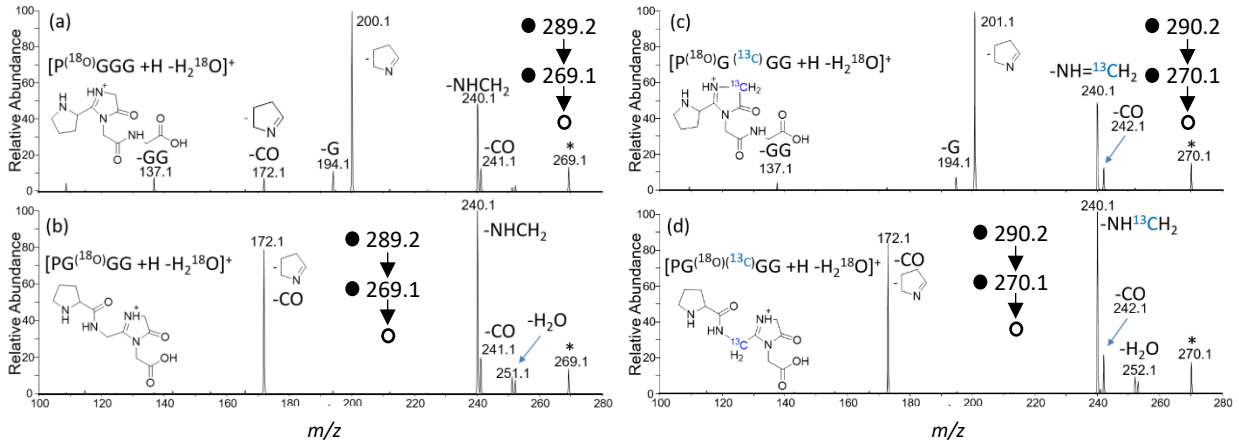


Figure 5.13 CID spectra of $[b_4]^+$ ions of ProGlyGlyGly formed by loss of water from (a) the first amide and (b) the second amide. CID spectra of $[b_4]^+$ ions of ProGlyGlyGly with $^{13}C_\alpha$ in the same residue in which water was lost from (c) the first amide and (d) the second amide. CID spectra (a), (b), (c) and (d) were collected at a normalized collision energy of 20.

The CID spectrum of $[Pro(^{18}O)GlyGlyGly + H - H_2{}^{18}O]^+$ ions showed the loss of 3,4-dihydro-2H-pyrrole (imine of proline) and methanimine. ^{13}C labelling revealed that the loss of methanimine is from the second residue, glycine (Figure 5.13). The loss of the imine of proline from structure **I-A** of the $[b_4]^+$ of protonated ProGlyGlyGly from the first amide can be rationalized by direct C_α-C cleavage at the N-terminus. Loss of methanimine from the same structure **I-A** requires interconversion as in *pathway (ii)* from **Scheme** 5.2. This would then be followed by C_α-C cleavage at the N-terminus.

The CID spectrum of $[ProGly(^{18}O)GlyGly + H - H_2{}^{18}O]^+$ showed the loss of methanimine and the loss of (imine of proline + CO). Loss of methanimine from structure **II-B** occurs by rearrangement described by *pathway (i)* followed by C_α-C cleavage at the N-terminus. Structure **II-B** can undergo amide bond cleavage of the first amide as described in Scheme 3.2 in the loss of (imine of proline + CO). Structure **II-B** can also undergo interconversion to structure **I-A** by *pathway (ii)* and cleave the C_α-C at the N-terminus followed by ring opening of the 4-imidazolone as described in **Scheme** 5.2. The energy resolved curve (Figure 5.14) also showed similar behavior to that of $[GlyAlaGlyGly - H_2O + H]^+$, as seen in (Figure 5.9). This may suggest

that structure **II-B** undergoes interconversion by *pathway (ii).* It may also be possible for the 4-imidazolone ring to open and lose (imine of proline + CO) directly from structure **II-B**.

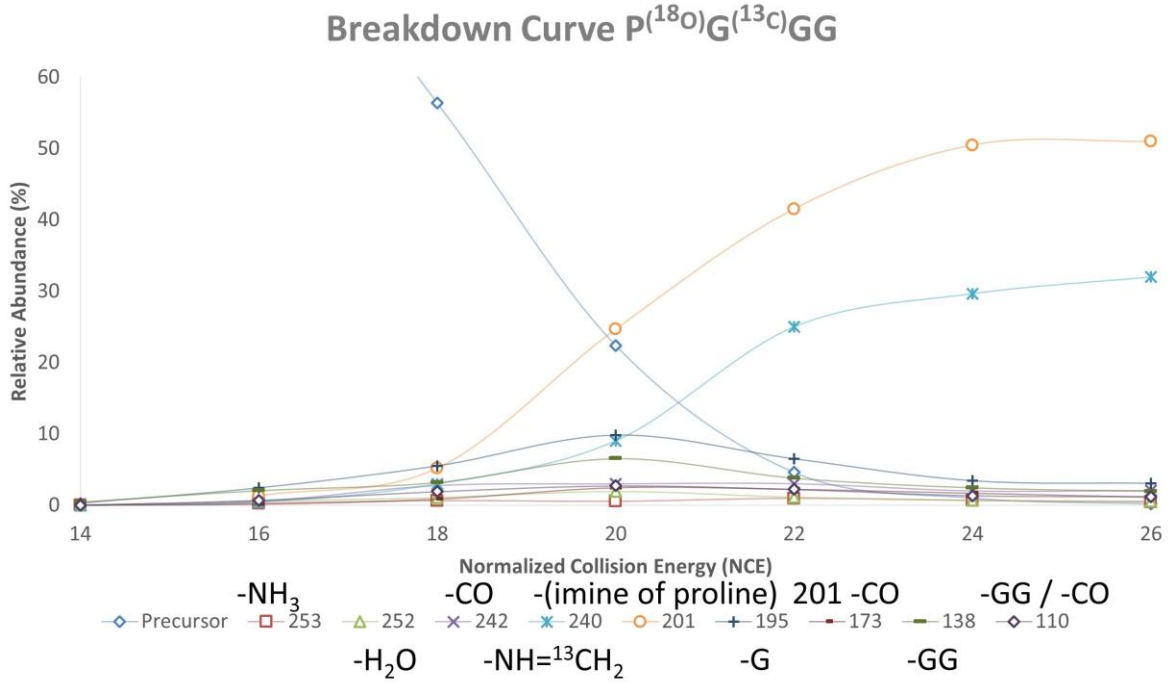


Figure 5.14 Energy-resolved curve of protonated ProGlyGlyGly after water loss from the first residue, proline.

5.6 Loss of water from [GlyProGlyGly + H]$^+$

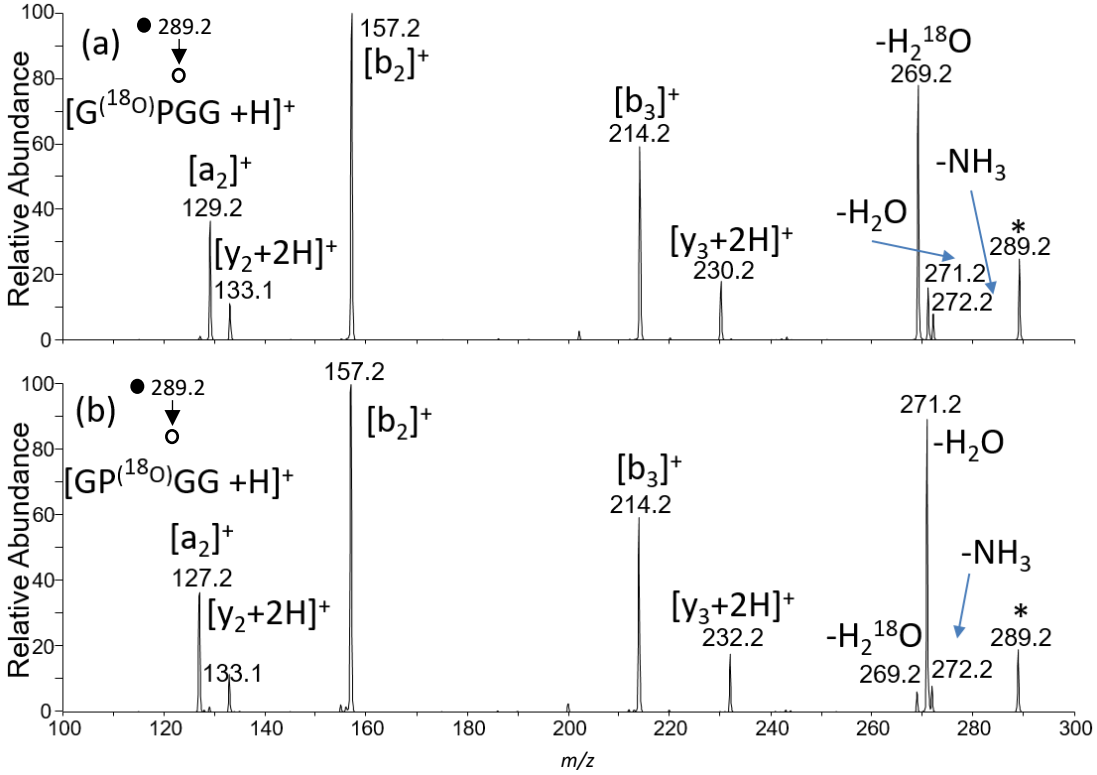


Figure 5.15 CID spectra of protonated GlyProGlyGly (a) with the ^{18}O label at the first amide oxygen (b) with the ^{18}O label at the second amide oxygen. CID spectra (a) and (b) were collected at a normalized collision energy of 22.

Isotopic labelling with ^{18}O showed that water loss is in high abundance and is predominantly from the first amide; loss of water from the second amide is a minor channel (Figure 5.15). This may be due to the low energy structure of the proline-containing imidazolone. The 5-membered ring structure of the proline stabilizes the positive charge within the imidazolone ring. The CID spectrum of [Gly(^{18}O)ProGlyGly + H - H_2O]$^+$ showed only the loss of methanimine and CO in low abundance. The CID spectrum of [GlyPro(^{18}O)GlyGly +H -H_2O]$^+$ showed the loss of 3,4-dihydro-2H-pyrrole (imine of proline), presumably from the proline residue, and loss of methanimine (Figure 5.16).

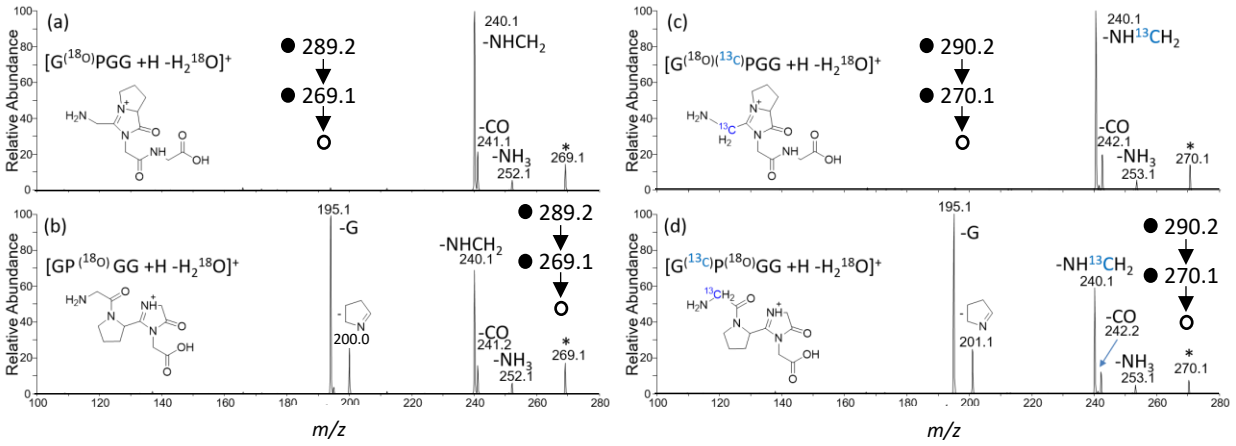


Figure 5.16 CID spectra of $[b_4]^+$ ions of GlyProGlyGly formed by loss of water from (a) the first amide and (b) the second amide. CID spectra of $[b_4]^+$ ions of GlyProGlyGly with $^{13}C_\alpha$ in the same residue from which water was lost in (c) the first amide and (d) the second amide. CID spectra (a), (b), (c) and (d) were collected at a normalized collision energy of 20.

^{13}C labelling showed that loss of methanimine always comes from the first residue, glycine. The loss of methanimine from [Gly(^{18}O)ProGlyGly + H - $H_2{}^{18}O$]$^+$, structure **I-B** can be rationalized by C_α-C cleavage at the N-terminus of the structure given the spectrum in Figure 5.16 (a) as described in Scheme 3.2. Loss of the imine of proline from [GlyPro(^{18}O) GlyGly + H - $H_2{}^{18}O$]$^+$, Structure **II-A**, is a low energy product and can be rationalized by interconversion by *pathway (i)* described in **Scheme** 5.1 followed by C_α-C cleavage to lose the imine from the same residue in which water was lost. Loss of methanimine from [GlyPro(^{18}O) GlyGly + H - $H_2{}^{18}O$]$^+$ can be rationalized by conversion of Structure **II-A** to **I-B** by *pathway (ii)* described in **Scheme** 5.2 followed by C_α-C cleavage to lose the imine from a different residue than that from which water was lost. The energy-resolved curve of [GlyPro(^{18}O) GlyGly + H - $H_2{}^{18}O$]$^+$, is also similar to that of [GlyAlaGlyGly - H_2O + H]$^+$, where loss of methanimine is a high energy pathway (Figure 5.17). This may suggest that both [GlyAla(^{18}O)GlyGly - $H_2{}^{18}O$ + H]$^+$, and [GlyPro(^{18}O)GlyGly + H - $H_2{}^{18}O$]$^+$ undergo similar interconversion mechanisms. The loss of glycine at the C-terminus is the dominant pathway from [GlyPro(^{18}O) GlyGly + H - $H_2{}^{18}O$]$^+$ but it is unclear as to why.

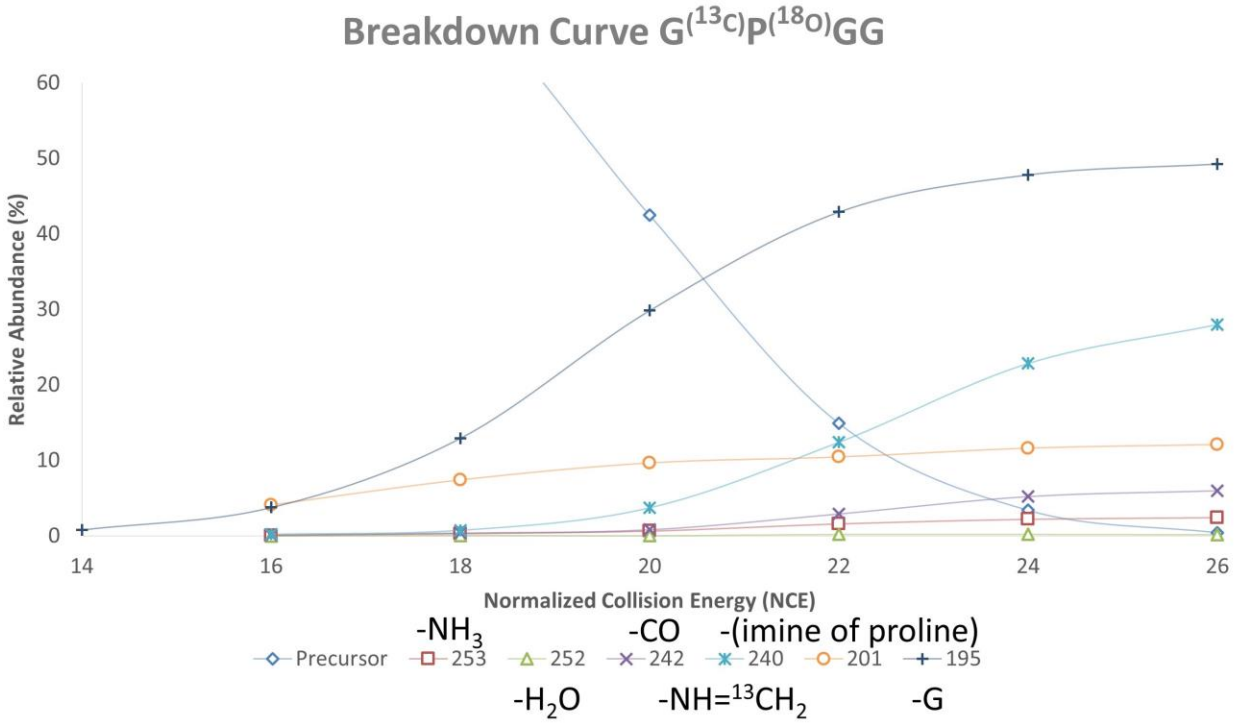


Figure 5.17 Energy-resolved curve of $[b_4]^+$ ion derived from protonated GlyProGlyGly after water loss from the second residue, proline.

5.7 Fragmentation of the ion at *m/z* 200 and ion at *m/z* 240

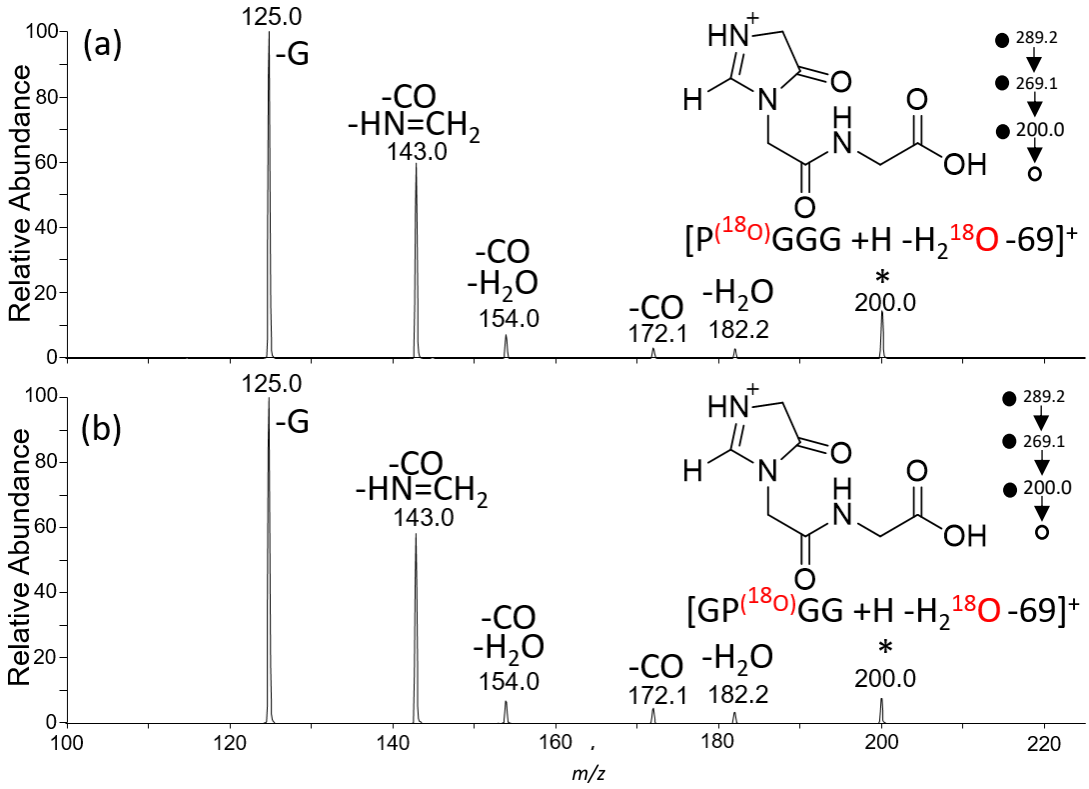


Figure 5.18 CID spectra are for the dissociation of *m/z* 200 ions derived from loss of water followed by loss of ethanimine from protonated (a) ProGlyGlyGly and (b) GlyProGlyGly. CID spectra (a) and (b) were collected at a normalized collision energy of 26.

The loss of the imine of proline as observed in Figure 5.18 for [GlyProGlyGly + H - H_2O - C_4H_7N]$^+$, [ProGlyGlyGly + H - H_2O - C_4H_7N]$^+$, result in the same CID spectra as seen from [GlyGlyGlyGly + H - H_2O - $HN{=}CH_2$]$^+$. The *m/z* 200 ions are formed by the same C_α-C bond cleavage resulting in the loss of the imine of proline instead of the imine of glycine, due to the proline substitution for glycine. Since the ions formed by the loss of the imine of proline from either peptide generated identical CID spectra, this also supports the idea of the presence of the imidazole-4-one structure and interconversions of residues.

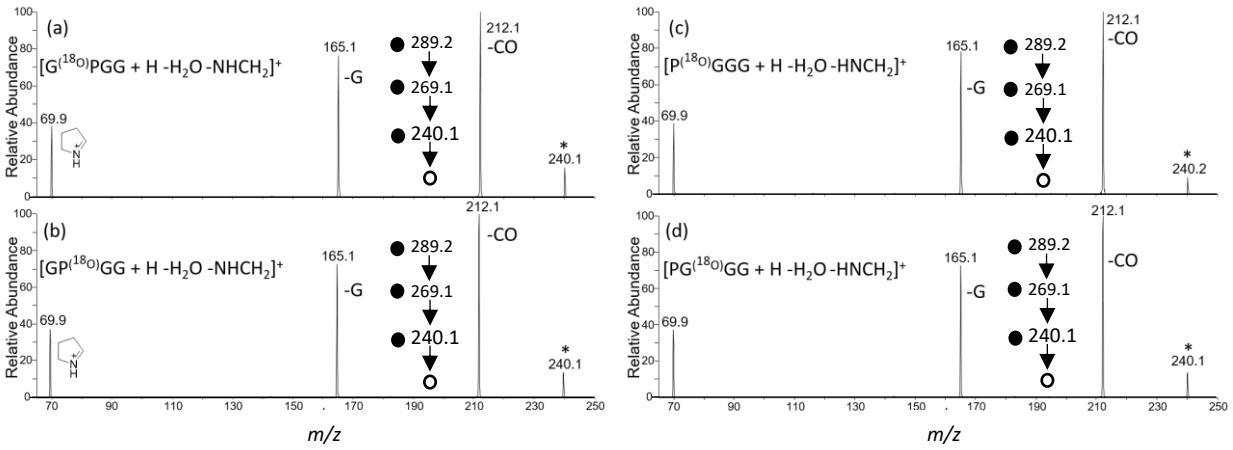


Figure 5.19 CID spectra of *m/z* 240 ion of GlyProGlyGly formed by loss of water from (a) the first amide and (b) the second amide, followed by loss of methanimine. CID spectra of *m/z* 240 ions of ProGlyGlyGly formed by loss of water from (c) the first amide and (d) the second amide, followed by loss of methanimine. CID spectra (a), (b), (c) and (d) were collected at a normalized collision energy of 20.

The ions produced by loss of methanimine as observed in Figure 5.19 for [GlyProGlyGly + H - H_2O - $HN=CH_2]^+$, [ProGlyGlyGly + H - H_2O - $HN=CH_2]^+$ gave identical CID spectra. The energy-resolved curves were also identical (Appendix. B. 25, 26, 27 and 28). This suggests that the ions are identical and that their respective $[b_4]^+$ ions interconvert. The fragmentation pathway of this ion is different from that of the other analogous ions as it lacks a mobile proton.

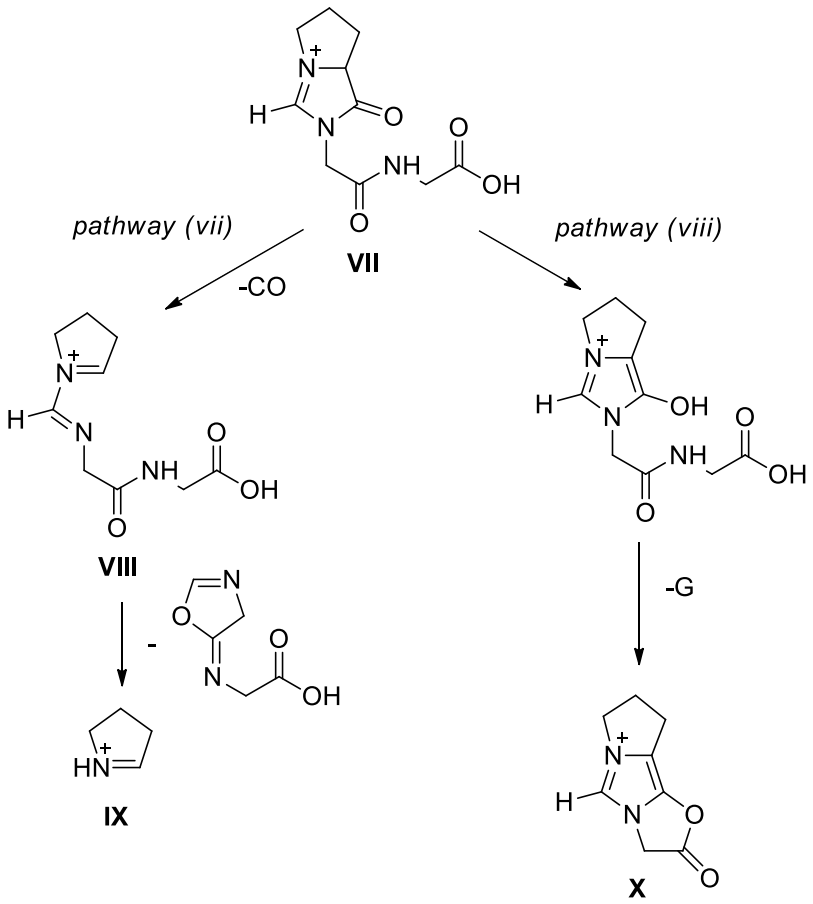
The pathways by which the *m/z* 240 ion dissociates or rearranges are given in Scheme 5.5.

(i) The loss of CO is a result of ring opening of the 4-imidazolone ring as seen in *pathway vii*. The resulting structure **VIII** has a positive charge localized on the nitrogen of the proline ring. The positive charge and the resulting double bond is stabilized by the proline ring.

(ii) Following the loss of CO, formation of the imine of proline *m/z* 70 ion is initiated by nucleophilic attack of the adjacent amide oxygen. The nitrogen of proline and the 5-membered ring stabilize both the double bond and the positive charge (structure **IX**).

(iii) Loss of glycine from the C-terminus requires a mobile proton, possibly by formation of the enol form the product of a 1,3 H shift. As the enol is part of an aromatic (6 π electron) ring that carries a positive charge, the OH will be slightly acidic and will

hydrogen bond to the C-terminal amide. Transferring the enol hydrogen to the amide nitrogen with concomitant cleavage of the C-N bond and formation of a C-O bond results in the loss of a glycine molecule and gives ion X.



Scheme 5.5 Fragmentation pathways of the *m/z* 240 ion

5.6 Conclusion

Water loss from the alanine or proline residues of XGlyGlyGly are in low abundance, possibly due to the steric hindrance of the side chain. The $[b_4]^+$ ions formed by loss of water from XXGlyGly; X=Gly, Ala, Pro (a peptide cannot contain Ala and Pro and contains only one non-glycine residue) show loss of imines to be the major pathway. When water is lost from a glycine residue then the resultant $[b_4]^+$ ion always results in only the loss of methanimine as the major pathway. Isotopic labelling confirmed that methanimine is lost from the same residue from which water was initially lost, from the glycine residue. This suggests the formation of a 4-imidazolone followed by interconversion, as previously described by the CID of protonated tetraglycine. However, when water is lost from an alanine or proline residue, both methanimine and the respective imines of the residue from which water was lost is present in high abundance at high collision energies. A secondary interconversion mechanism was discovered to rationalize the loss of methanimine from a different residue from which water was initially lost. Loss of the ethanimine or imine of proline can be explained by formation of a 4-imidazolone and a combination of the C-C_α cleavage and the various interconversion mechanisms. The *m/z* 200 ions **III**, always showed identical CID spectra with that from protonated GlyGlyGlyGly. The *m/z* 214 ions formed from losses of methanimine from alanine substituted tetraglycines showed common fragmentation pathways similar to the *m/z* 200 ion. The only notable exception being the increased loss of CO. Ring opening of the 4-imidazole ring is stabilized by the alanine side chain, to make this pathway more competitive. The m/z 240 ion produced by loss of methanimine from proline-containing tetrapeptides does not have a mobile proton and follows an uncommon fragmentation pathway.

Future Directions

Herein only peptides with aliphatic amino acids have been studied. The water loss from protonated peptides involving a greater variety of amino acids that have not been studied. Most likely the amino acids with heteroatoms in the side chain can act as both nucleophiles and receiving groups. The major pathway may not be the formation of the imidazole-4-one. Thus, IRMPD should be used to elucidate such structures. A similar research approach should be used where water loss from diglycines and triglycines involving a one amino acid substitution should be first be studied and then expand to longer peptides. Once the pathways for water loss is understood then a combination of amino acids should be studied. Deciphering the CID spectra for a long peptide sequence with multiple amino acid substitutions may prove challenging.

The role of steric or electronic effects in the water loss of protonated peptides should be further studied. Synthetic peptides involving the addition of electron-withdrawing groups such as halogens can be introduced to increase the susceptibility of the receiving groups to nucleophilic attack. Amino acids with electron donating groups such as tertiary amines can be introduced and compared with long aliphatic amino acid chain. Electronic and steric effects on water loss can then be determined by comparing these electron donating group, electron withdrawing group and aliphatic side chains.

References

(1) Yamashita, M.; Fenn, J. B. Electrospray Ion Source. Another Variation on the Free-Jet Theme. *J. Phys. Chem.* **1984**, *88* (20), 4451–4459.

(2) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and Polymer Analyses up Tom/z 100 000 by Laser Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1988**, *2* (8), 151–153.

(3) Wilkins, M. R.; Pasquali, C.; Appel, R. D.; Ou, K.; Golaz, O.; Sanchez, J. C.; Yan, J. X.; Gooley, A. A.; Hughes, G.; Humphery-Smith, I.; et al. From Proteins to Proteomes: Large Scale Protein Identification by Two-Dimensional Electrophoresis and Amino Acid Analysis. *Biotechnology. (N. Y).* **1996**, *14* (1), 61–65.

(4) James, P. Protein Identification in the Post-Genome Era: The Rapid Rise of Proteomics. *Q. Rev. Biophys.* **1997**, *30* (4), 279–331.

(5) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R.; III. Protein Analysis by Shotgun/bottom-up Proteomics. *Chem. Rev.* **2013**, *113* (4), 2343–2394.

(6) Yates, J. R.; Ruse, C. I.; Nakorchevsky, A. Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annu. Rev. Biomed. Eng.* **2009**, *11* (1), 49–79.

(7) Olsen, J. V; Ong, S.-E.; Mann, M. Trypsin Cleaves Exclusively C-Terminal to Arginine and Lysine Residues. *Mol. Cell. Proteomics* **2004**, *3* (6), 608–614.

(8) Switzar, L.; Giera, M.; Niessen, W. M. A. Protein Digestion: An Overview of the Available Techniques and Recent Developments. *J. Proteome Res.* **2013**, *12* (3), 1067–1077.

(9) Medzihradszky, K. F.; Chalkley, R. J. Lessons in de Novo Peptide Sequencing by Tandem Mass Spectrometry. *Mass Spectrom. Rev. 34* (1), 43–63.

(10) Steen, H.; Mann, M. The ABC's (and XYZ's) of Peptide Sequencing. *Nat. Rev. Mol. Cell Biol.* **2004**, *5* (9), 699–711.

(11) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. Why Are B Ions Stable Species in Peptide Spectra? *J. Am. Soc. Mass Spectrom.* **1995**, *6* (12), 1165–1174.

(12) Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. The Structure and Fragmentation of Bn, (n>3) Ions in Peptide Spectra. *J. Am. Soc. Mass Spectrom.* **1996**, *7* (3), 233–242.

(13) Harrison, A. G. To B or Not to B: The Ongoing Saga of Peptide B Ions. *Mass Spectrom. Rev.* **2009**, *28* (4), 640–654.

(14) Roepstorff, P.; Fohlman, J. Letter to the Editors. *Biol. Mass Spectrom.* **1984**, *11* (11), 601–601.

(15) Biemann, K. [25] Sequencing of Peptides by Tandem Mass Spectrometry and High-Energy Collision-Induced Dissociation. *Methods Enzymol.* **1990**, *193*, 455–479.

(16) Papayannopoulos, I. A. The Interpretation of Collision-Induced Dissociation Tandem Mass Spectra of Peptides. *Mass Spectrom. Rev.* **1995**, *14* (1), 49–73.

(17) Perkins, B. R.; Chamot-Rooke, J.; Yoon, S. H.; Gucinski, A. C.; Somogyi, A.; Wysocki, V. H. Evidence of Diketopiperazine and Oxazolone Structures for HA B_2^+ Ion. *J. Am. Chem. Soc.* **2009**, *131* (48), 17528–17529.

(18) Rodriquez, C. F.; Shoeib, T.; Chu, C. F.; Siu, K. W. M.; Hopkinson, A. C. Comparison between Protonation, Lithiation, and Argentination of 5-Oxazolones: A Study of a Key Intermediate in Gas-Phase Peptide Sequencing. *J. Phys. Chem. A* **2000**, *104* (22), 5335–5342.

(19) Nold, M. J.; Wesdemiotis, C.; Yalcin, T.; Harrison, A. G. Amide Bond Dissociation in Protonated Peptides. Structures of the N-Terminal Ionic and Neutral Fragments. *Int. J. Mass Spectrom. Ion Process.* **1997**, *164* (1–2), 137–153.

(20) Paizs, B.; Suhai, S. Fragmentation Pathways of Protonated Peptides. *Mass Spectrom. Rev.* **2005**, *24* (4), 508–548.

(21) Harrison, A. G.; Csizmadia, I. G.; Tang, T.-H. Structure and Fragmentation of b2 Ions in Peptide Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2000**, *11* (5), 427–436.

(22) Bernier, M. C.; Chamot-Rooke, J.; Wysocki, V. H. R vs. S Fluoroproline Ring Substitution: Trans/cis Effects on the Formation of B $_2$ Ions in Gas-Phase Peptide Fragmentation. *Phys. Chem. Chem. Phys.* **2016**, *18* (3), 2202–2209.

(23) Cooper, T.; Talaty, E.; Grove, J.; Van Stipdonk, M.; Suhai, S.; Paizs, B. Isotope Labeling and Theoretical Study of the Formation of a3* Ions from Protonated Tetraglycine. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (12), 1654–1664.

(24) Allen, J. M.; Racine, A. H.; Berman, A. M.; Johnson, J. S.; Bythell, B. J.; Paizs, B.; Glish, G. L. Why Are a3 Ions Rarely Observed? *J. Am. Soc. Mass Spectrom.* **2008**, *19* (12), 1764–1770.

(25) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. Sequence-Scrambling Fragmentation Pathways of Protonated Peptides. *J. Am. Chem. Soc.* **2008**, *130* (52), 17774–17789.

(26) Harrison, A. G.; Young, A. B.; Bleiholder, C. Scrambling of Sequence Information in Collision-Induced Dissociation of Peptides. **2006**, 10364–10365.

(27) Wu, B.-H. B.; Lau, J. K.-C.; Siu, K. W. M.; Hopkinson, A. C. Dissociation of [B 5] + Ions Containing an α-Methyltryptophan and Four Alanine Residues; Losses of CO 2 and the Oxazolone Ring. *Int. J. Mass Spectrom.* **2017**.

(28) Molesworth, S.; Osburn, S.; Stipdonk, M. Influence of Size on Apparent Scrambling of Sequence during CID of B-Type Ions. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (11), 2174–2181.

(29) Ballard, K. D.; Gaskell, S. J. Dehydration of Peptide [M + H] + Ions in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **1993**, *4* (6), 477–481.

(30) Reid, G. E.; Simpson, R. J.; O ’Hair, R. A. J. Probing the Fragmentation Reactions of Protonated Glycine Oligomers via Multistage Mass Spectrometry and Gas Phase Ion Molecule Hydrogen/deuterium Exchange. *1International J. Mass Spectrom.* **1999**, *190/191*, 209–230.

(31) Reid, G. E.; Simpson, R. J.; O ’Hair, R. A. J. A Mass Spectrometric and Ab Initio Study of the Pathways for Dehydration of Simple Glycine and Cysteine-Containing Peptide [M⁁H] ⁁ Ions. *Int. J. Mass Spectrom.* **1998**, *9* (9), 945–956.

(32) Armentrout, P. B.; Heaton, A. L. Thermodynamics and Mechanisms of Protonated Diglycine Decomposition: A Computational Study. *J. Am. Soc. Mass Spectrom* **2012**, *23*, 621–631.

(33) Mookherjee, A.; Van Stipdonk, M. J.; Armentrout, P. B. Thermodynamics and Reaction

Mechanisms of Decomposition of the Simplest Protonated Tripeptide, Triglycine: A Guided Ion Beam and Computational Study. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 739–757.

(34) Lam, K. H. B.; Lau, J. K.-C.; Lai, C.-K.; Hopkinson, A. C.; Siu, K. W. M. Interconversion between 4-Imidazolone Ions; Isomers of $[B_4]^+$ Derived from Protonated Tetraglycine. *J. Phys. Chem. B* **2017**, *121* (41), 9541–9547.

(35) Verkerk, U. H.; Zhao, J.; Van Stipdonk, M. J.; Bythell, B. J.; Oomens, J.; Hopkinson, A. C.; Siu, K. W. M. Structure of the $[M + H - H_2O]^+$ Ion from Tetraglycine: A Revisit by Means of Density Functional Theory and Isotope Labeling. *J. Phys. Chem. A* **2011**, *115* (24), 6683–6687.

(36) Lau, J. K. C.; Zhao, J.; Siu, K. W. M.; Hopkinson, A. C. Elimination of Water from the Backbone of Protonated Tetraglycine. *Int. J. Mass Spectrom.* **2012**, *316–318*, 268–272.

(37) Kisselev, A. F.; Akopian, T. N.; Woo, K. M.; Goldberg, A. L. The Sizes of Peptides Generated from Protein by Mammalian 26 and 20 S Proteasomes. Implications for Understanding the Degradative Mechanism and Antigen Presentation. *J. Biol. Chem.* **1999**, *274* (6), 3363–3371.

(38) Kisselev, A. F.; Akopian, T. N.; Goldberg, A. L. Range of Sizes of Peptide Products Generated during Degradation of Different Proteins by Archaeal Proteasomes. *J. Biol. Chem.* **1998**, *273* (4), 1982–1989.

(39) El-Aneed, A.; Cohen, A.; Banoub, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Appl. Spectrosc. Rev.* **2009**, *44* (3), 210–230.

(40) Banerjee, S.; Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Int. J. Anal. Chem.* **2012**, *2012*, 1–40.

(41) Zubarev, R. A.; Makarov, A. Orbitrap Mass Spectrometry. *Anal. Chem.* **2013**, *85* (11), 5288–5296.

(42) Hager, J. W. A New Linear Ion Trap Mass Spectrometer. *Rapid Commun. Mass Spectrom.* **2002**, *16* (6), 512–526.

(43) Kind, T.; Fiehn, O. Advances in Structure Elucidation of Small Molecules Using Mass Spectrometry. *Bioanal. Rev.* **2010**, *2*, 23–60.

(44) Daly, N. R. Scintillation Type Mass Spectrometer Ion Detector. *Rev. Sci. Instrum.* **1960**, *31* (3), 264–267.

(45) Bruins, A. P. Mechanistic Aspects of Electrospray Ionization. *J. Chromatogr. A* **1998**, *794*, 345–357.

(46) Smith, R. D.; Barinaga, C. J.; Udseth, H. R. Improved Electrospray Ionization Interface for Capillary Zone Electrophoresis-Mass Spectrometry. *Anal. Chem.* **1988**, *60* (18), 1948–1952.

(47) Burton, J. C.; Taborek, P. Simulations of Coulombic Fission of Charged Inviscid Drops. *Phys. Rev. Lett.* **2011**, *106* (14), 144501.

(48) Rayleigh, Lord. XX. On the Equilibrium of Liquid Conducting Masses Charged with Electricity. *Philos. Mag.* **1882**, *14* (87), 184–186.

(49) Hager, J. W. Recent Trends in Mass Spectrometer Development. *Anal. Bioanal. Chem.* **2004**, *378* (4), 845–850.

(50) Miller, P. E.; Denton, M. B. The Quadrupole Mass Filter: Basic Operating Concepts. *J. Chem. Educ.* **1986**, *63* (7), 617.

(51) Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K.; Horning, S. Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Anal. Chem.* **2006**, *78* (7), 2113–2120.

(52) Makarov, A.; Denisov, E.; Lange, O.; Horning, S. Dynamic Range of Mass Accuracy in LTQ Orbitrap Hybrid Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (7), 977–982.

(53) Olsen, J. V; de Godoy, L. M. F.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-Trap. *Mol. Cell. Proteomics* **2005**, *4* (12), 2010–2021.

(54) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Cooks, C. R. The Orbitrap: A New Mass Spectrometer. *J. Mass Spectrom.* **2005**, *40* (4), 430–443.

(55) Sleno, L.; Volmer, D. A. Ion Activation Methods for Tandem Mass Spectrometry. *J. Mass Spectrom.* **2004**, *39* (10), 1091–1112.

(56) Johnson, A. R.; Carlson, E. E. Collision-Induced Dissociation Mass Spectrometry: A Powerful Tool for Natural Product Structure Elucidation.

(57) Coin, I.; Beyermann, M.; Bienert, M. Solid-Phase Peptide Synthesis: From Standard Procedures to the Synthesis of Difficult Sequences. *Nat. Protoc.* **2007**, *2* (12), 3247–3256.

(58) Marecek, J.; Song, B.; Brewer, S.; Belyea, J.; Dyer, R. B.; Raleigh, D. P. A Simple and Economical Method for the Production of ^{13}C, ^{18}O-Labeled Fmoc-Amino Acids with High Levels of Enrichment: Applications to Isotope-Edited IR Studies of Proteins. *Org. Lett.* **2007**, *9* (24), 4935–4937.

(59) Kowalak, J. A.; Walsh, K. A. β-Methylthio-Aspartic Acid: Identification of a Novel Posttranslational Modification in Ribosomal Protein S12 from Escherichia Coli. *Protein Sci.* **1996**, *5* (8), 1625–1632.

(60) Balta, B. L.; Aviyente, V.; Lifshitz, C. Elimination of Water from the Carboxyl Group of GlyGlyH. *J. Am. Soc. Mass Spectrom.* **2003**, *14* (10), 1192–1203.

(61) Paizs, B.; Suhai, S. Theoretical Study of the Main Fragmentation Pathways for Protonated Glycylglycine. *Rapid Commun. Mass Spectrom.* **2001**, *15* (8), 651–663.

(62) Bythell, B. J.; Harrison, A. G. Formation of a1 Ions Directly from Oxazolone b2 Ions: An Energy-Resolved and Computational Study. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (5), 774–781.

Appendix A: Information about the Author

Journal Articles

Lam, K. H. B.; Lau, J. K.-C.; Lai, C.-K.; Hopkinson, A. C.; Siu, K. W. M. Interconversion between 4-Imidazolone Ions; Isomers of $[B_4]^+$ Derived from Protonated Tetraglycine. *J. Phys. Chem. B* **2017**, *121* (41), 9541–9547

Conferences

ASMS Conference on Mass Spectrometry and Allied Topics, San Diego, U.S.A 2018
Oral: Probing structures various cyclosporin analogues using differential mobility spectroscopy, collision induced dissociation and gas-phase hydrogen-deuterium exchange

International Mass Spectrometry Conference, Toronto, Canada 2016
Poster: The Water Loss from Protonated Polyglycines

Canadian Chemistry Conference and Exhibition, Toronto Canada 2017
Poster: Collision-induced Dissociation of $[b_4]^+$ Isomers Derived from Tetraglycine

ASMS Conference on Mass Spectrometry and Allied Topics, Indianapolis U.S.A 2017
Poster: Interconversion of protonated 4-imidazolone ions; isomers of $[b_4]^+$ derived from protonated tetraglycine

Research as a visiting student

Université Paris-Sud D'Orsay under Professor Philippe Maître working at the Center Laser Infrarouge D'Orsay (CLIO)

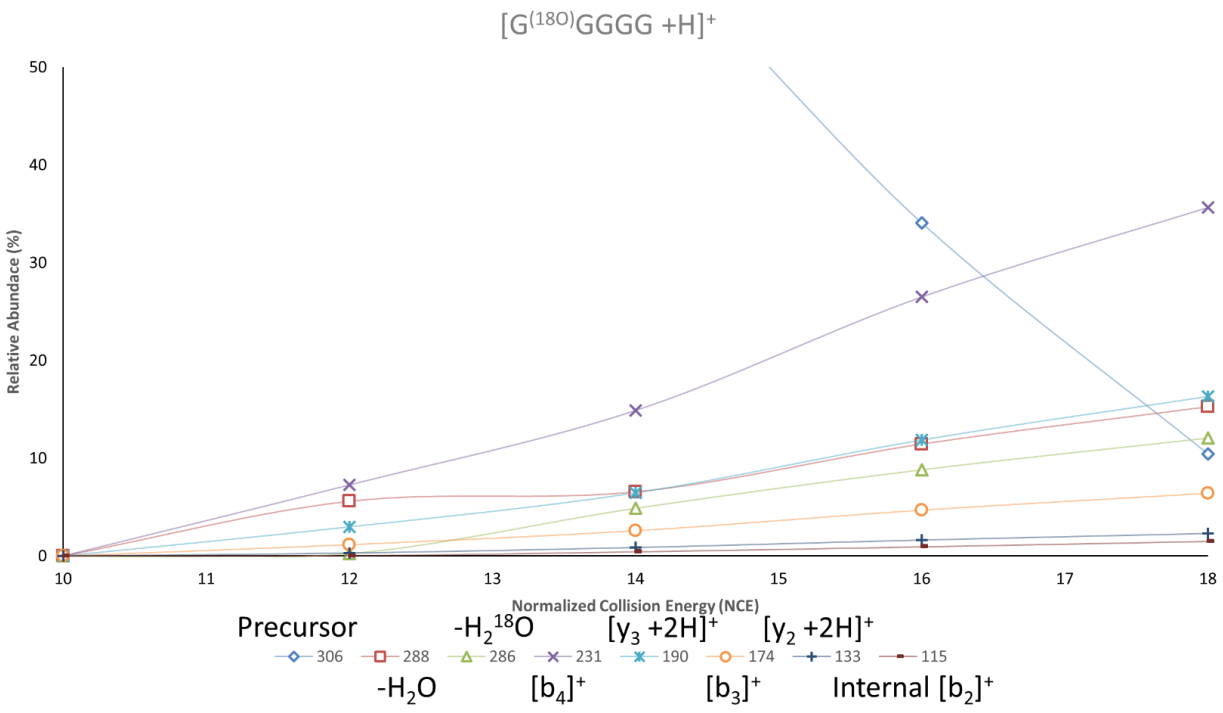
Internships

SCIEX, Research Department, a manufacturer of mass spectrometer technology

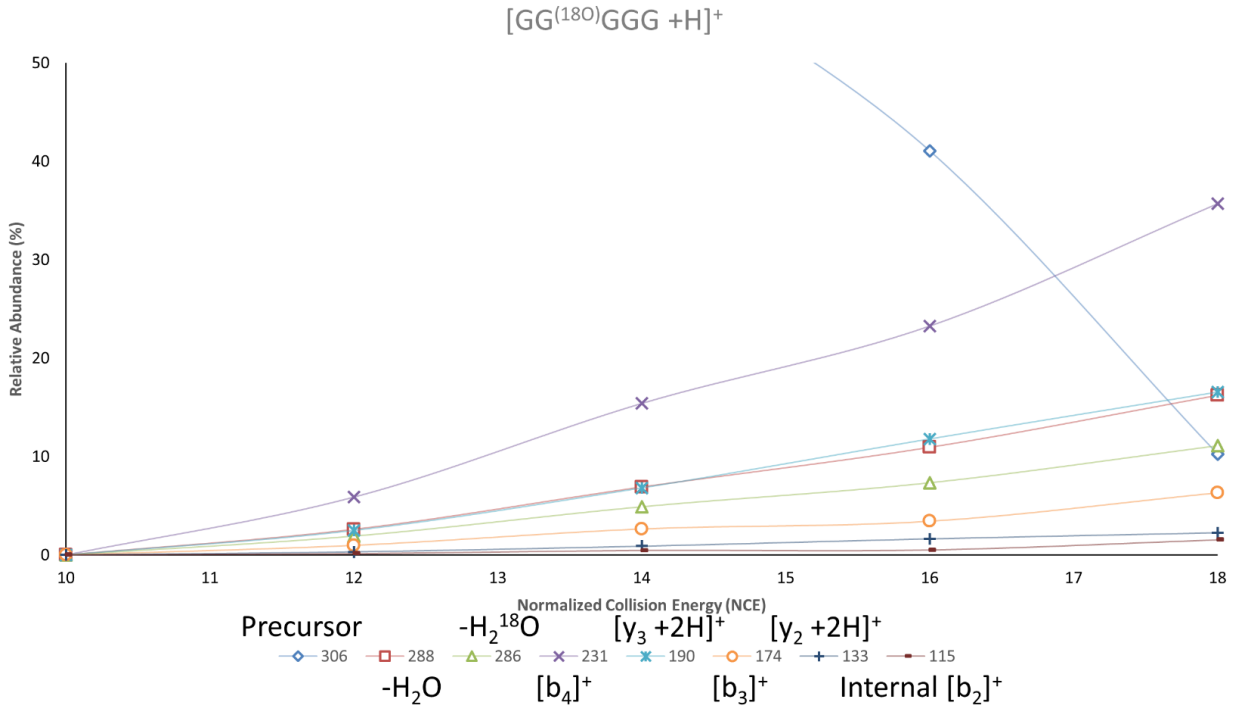
Disclosure

*Financial & competing interests disclosure K.H. Brian Lam is an intern at SCIEX, a manufacturer of mass spectrometer technology. The author has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed here.

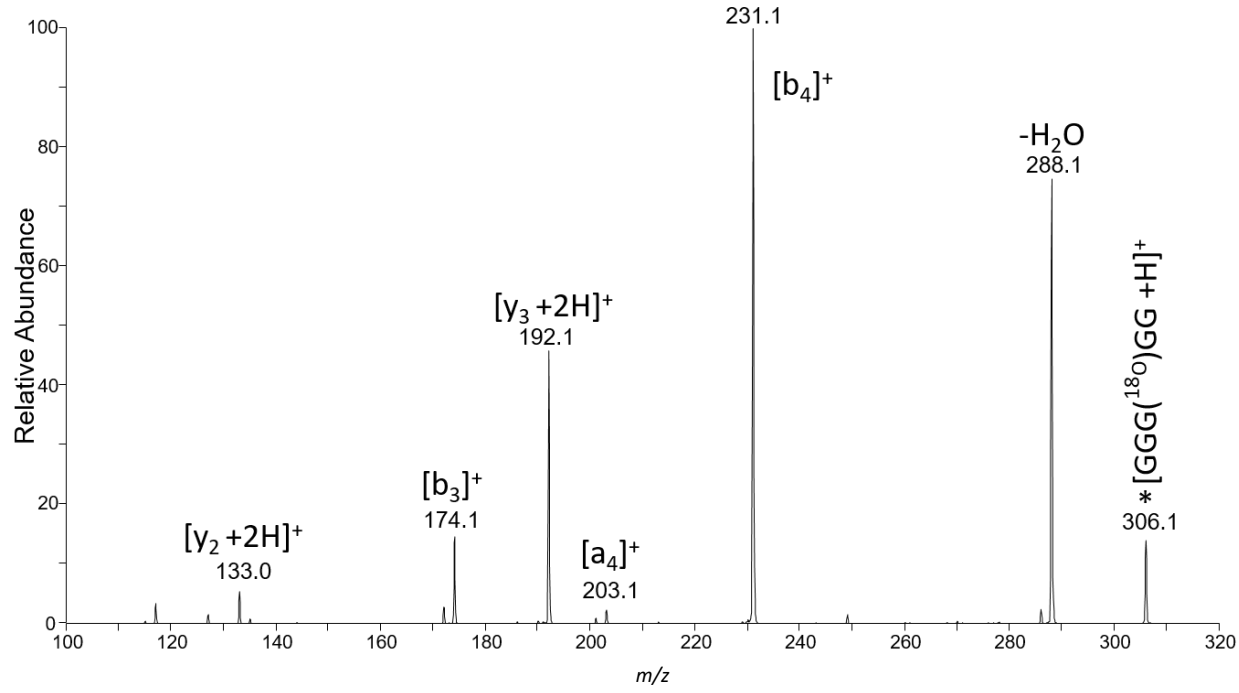
Appendix B: Supplemental Information



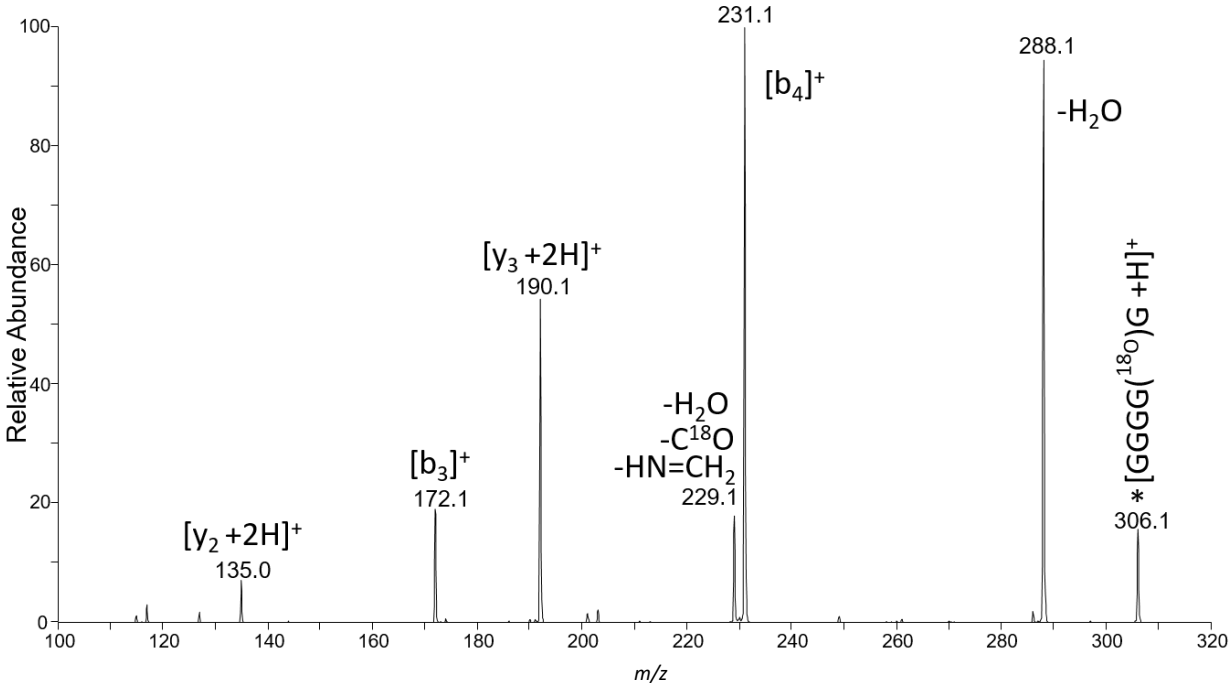
Appendix B. 1. Energy-resolved curves for ^{18}O labelled [GlyGlyGlyGlyGly +H]$^+$ in the first amide.



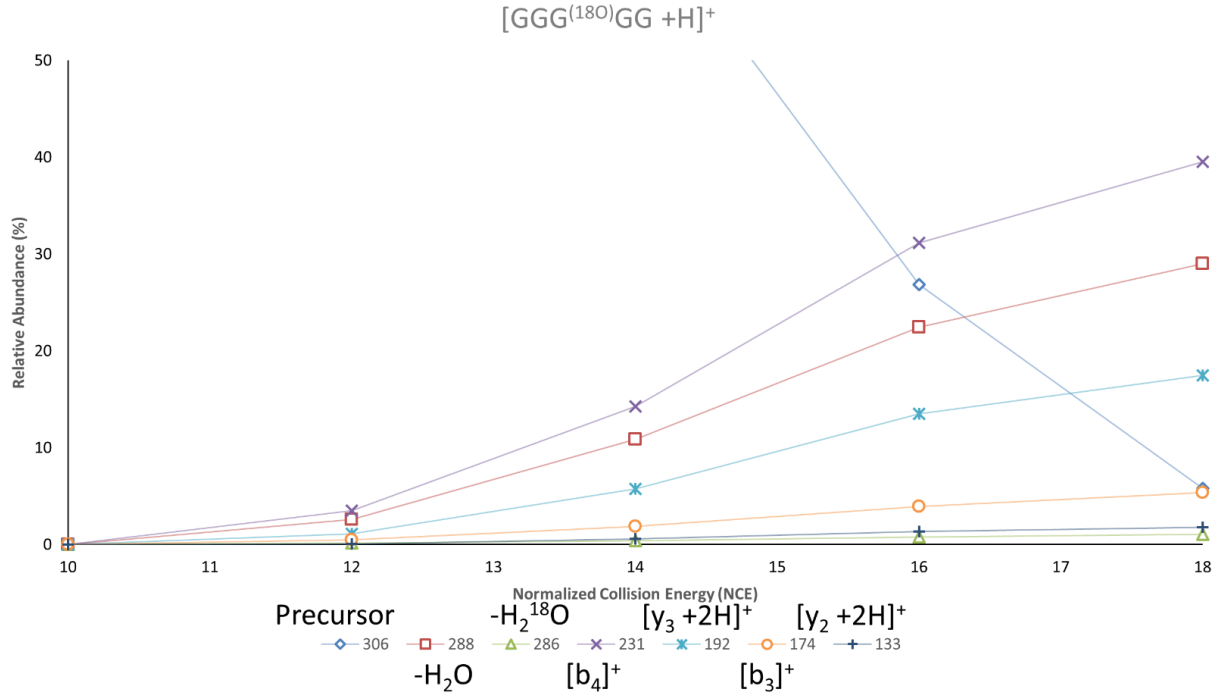
Appendix B. 2. Energy-resolved curves for ^{18}O labelled [GlyGlyGlyGlyGly +H]$^+$ in the second amide.



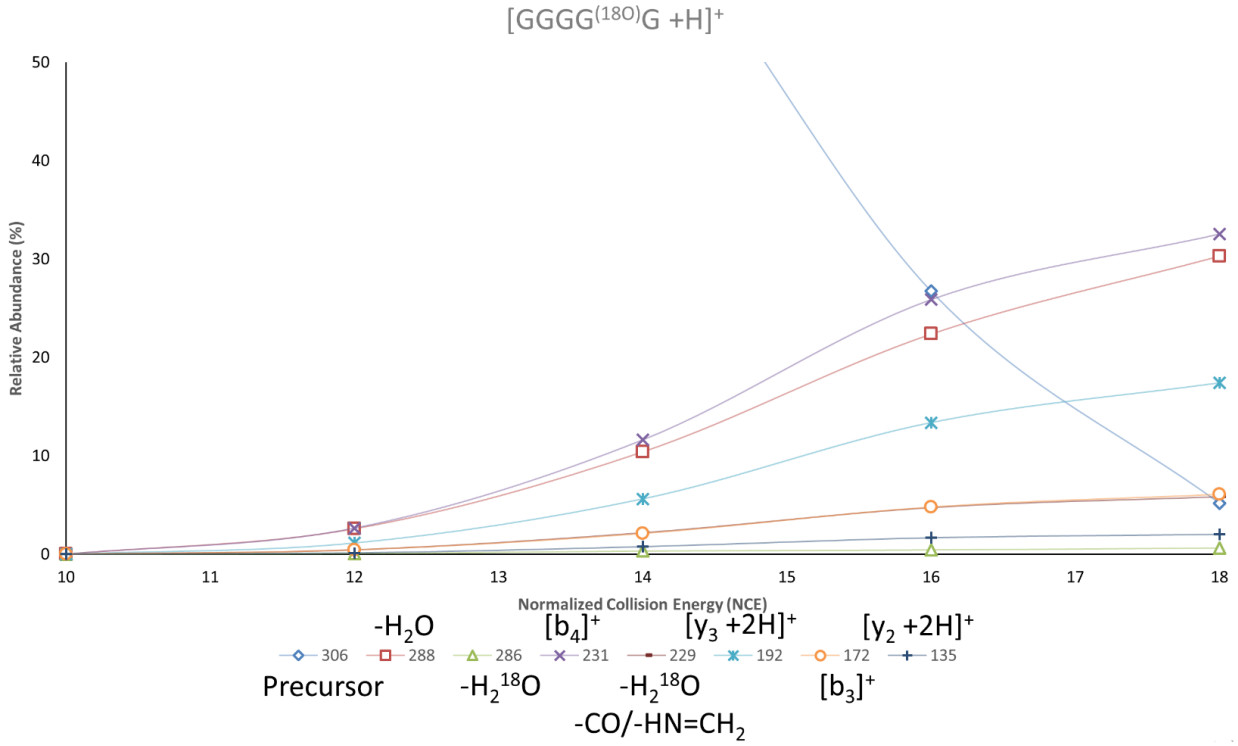
Appendix B. 3. CID spectra of pentaglycine with the third amide oxygen ^{18}O labelled. CID spectra were collected at a normalized collision energy of 20.



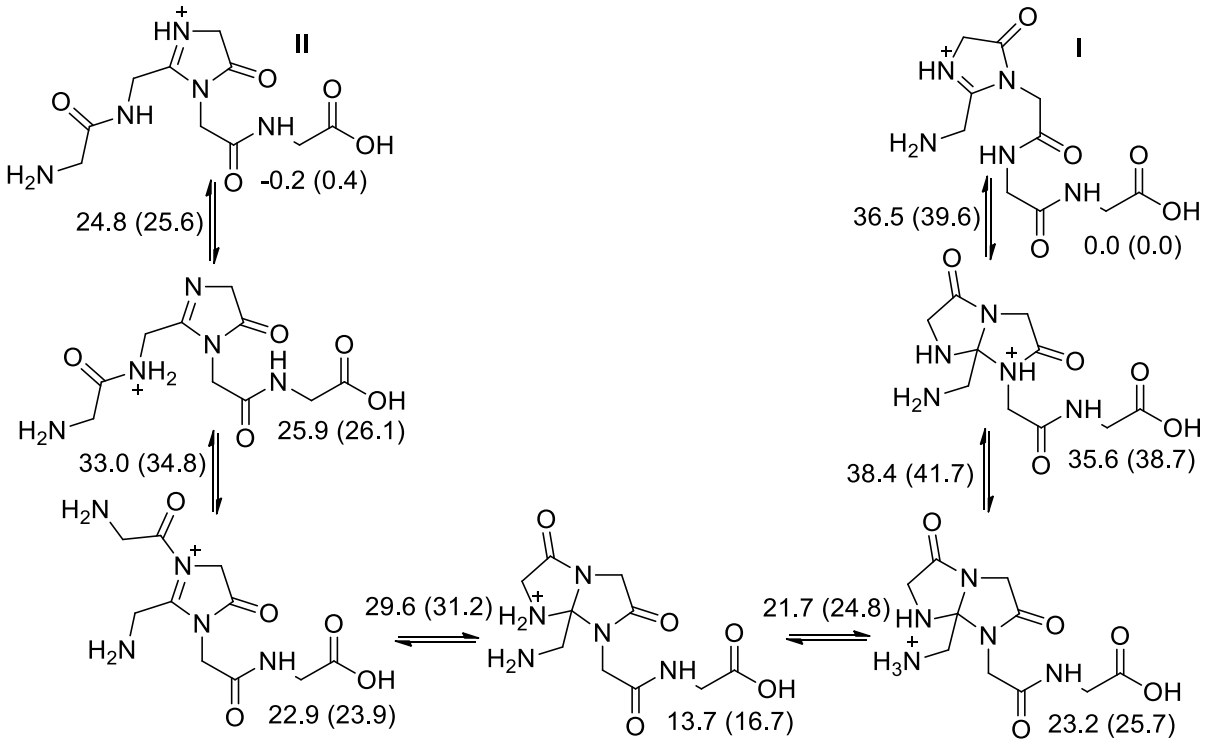
Appendix B. 4. CID spectra of pentaglycine with the fourth amide oxygen ^{18}O labelled. CID spectra were collected at a normalized collision energy of 20.



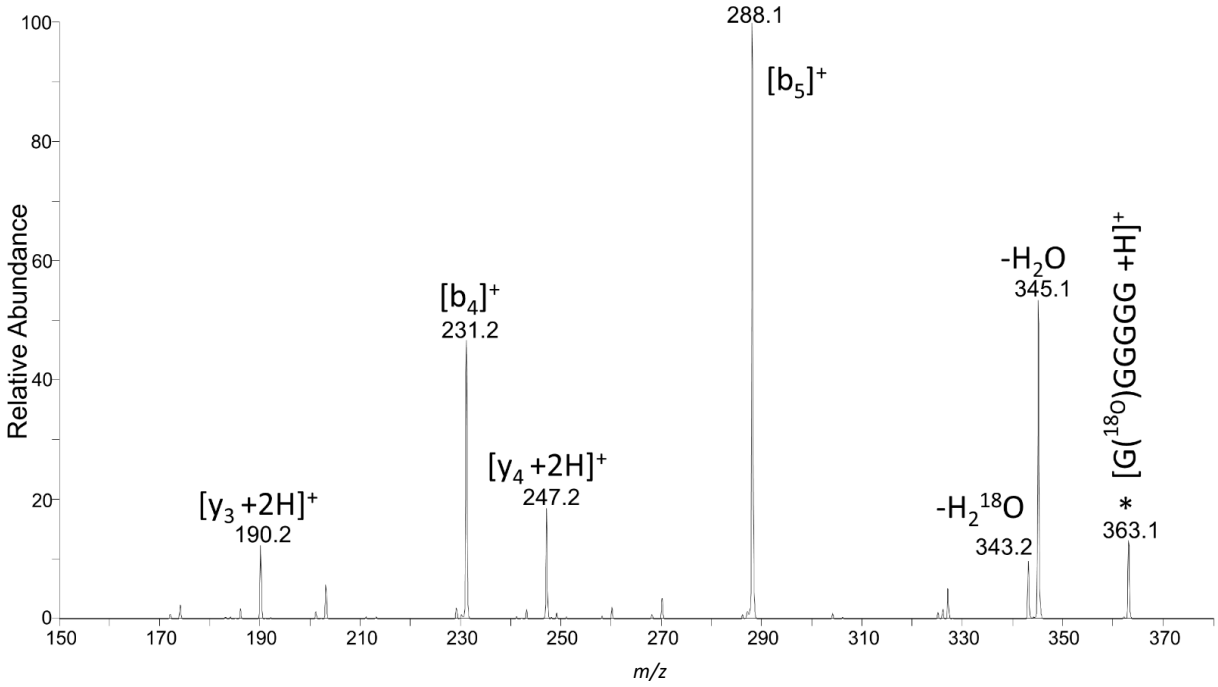
Appendix B. 5. Energy-resolved curves for ^{18}O labelled [GlyGlyGlyGlyGly +H]$^+$ in the third amide.



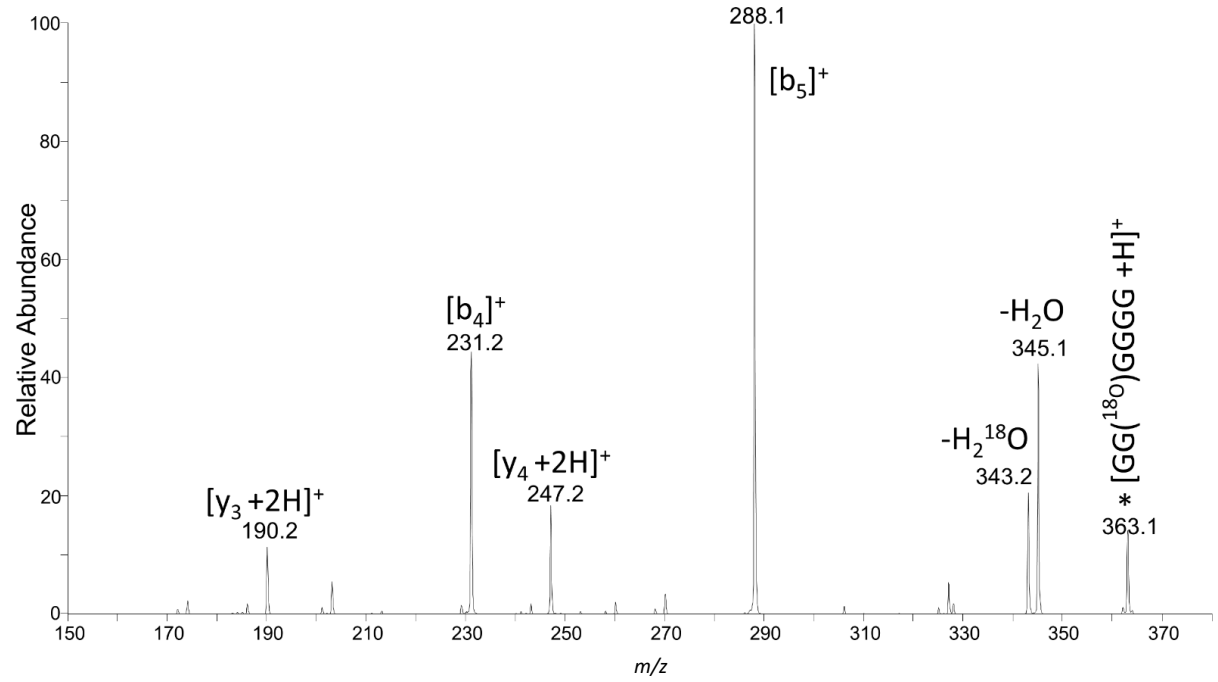
Appendix B. 6. Energy-resolved curves for ^{18}O labelled [GlyGlyGlyGlyGly +H]$^{+}$ in the fourth amide.



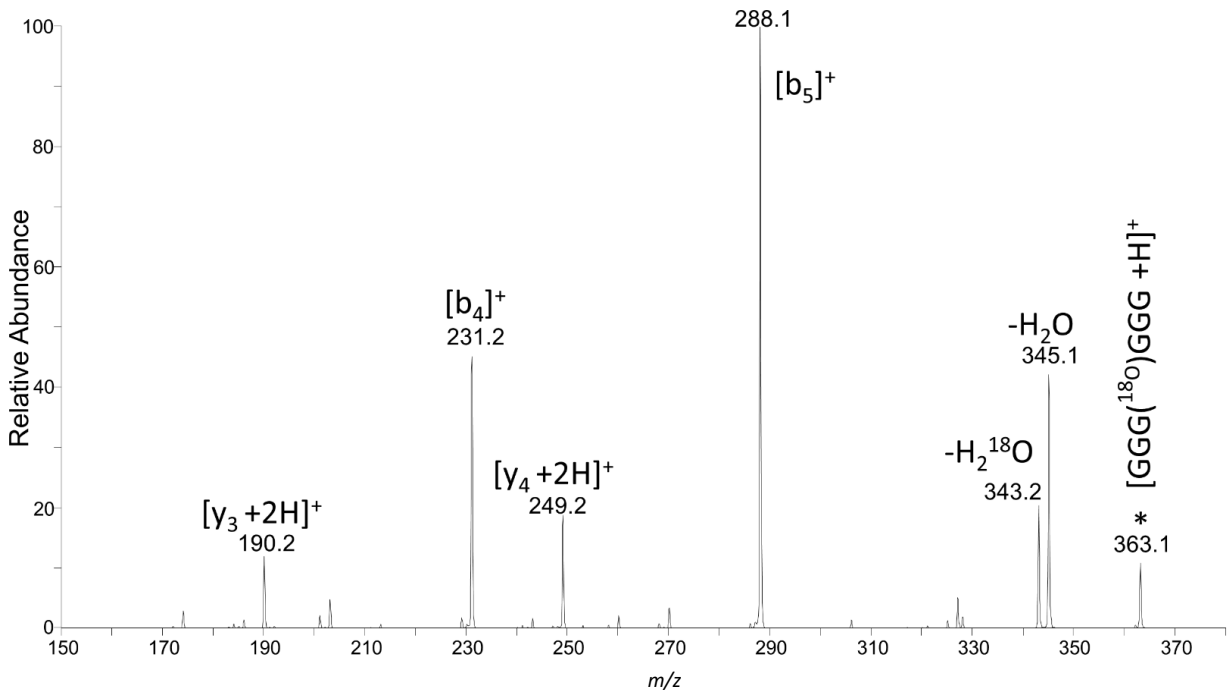
Appendix B. 7. Mechanism for the rearrangement of 2nd amide dehydrated pentaglycines. Relative enthalpies and (free energies) both in kcal mol^{-1} as calculated at the and **B3LYP/6-311++G(d,p) level**.



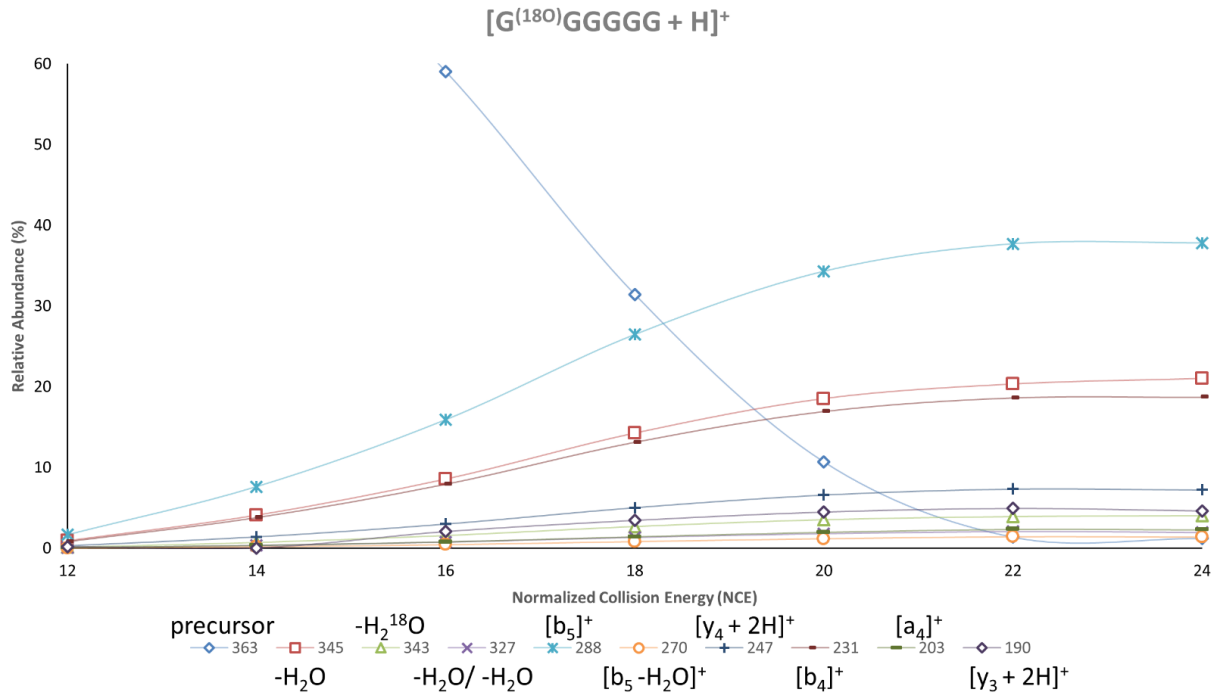
Appendix B. 8. CID spectrum of protonated hexaglycine with an ^{18}O label in the first amide. The CID spectrum was collected at a normalized collision energy of 21.



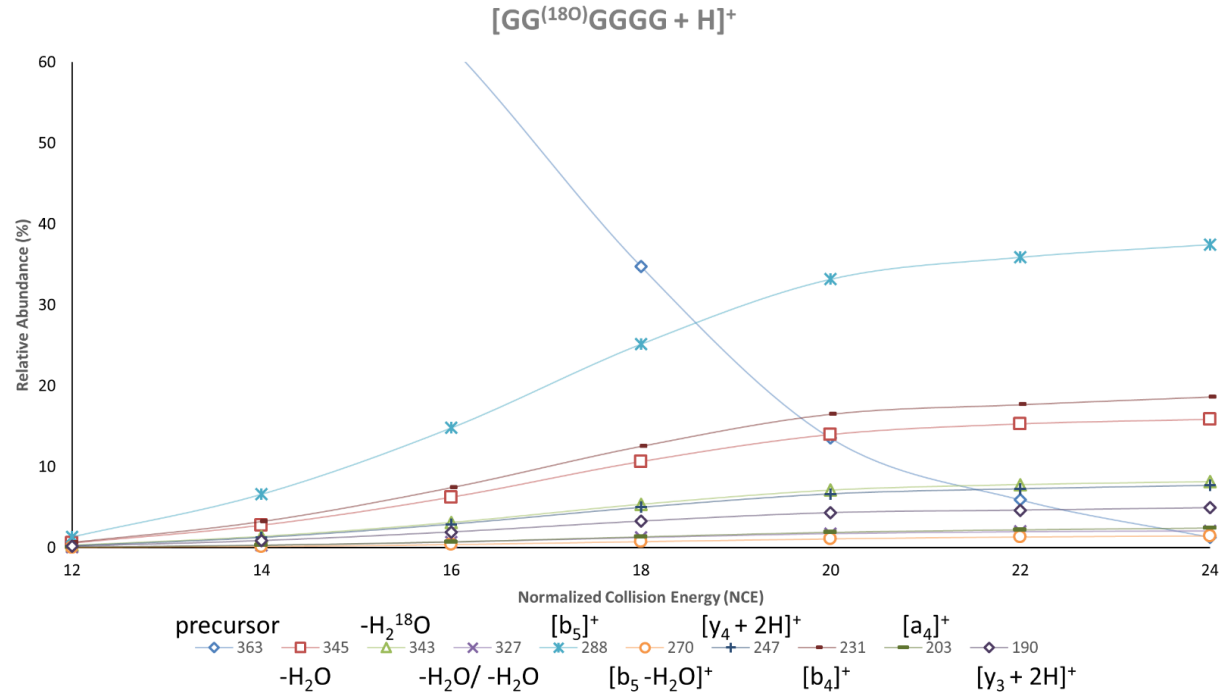
Appendix B. 9. CID spectrum of protonated hexaglycine with an ^{18}O label in the second amide. The CID spectrum was collected at a normalized collision energy of 21.



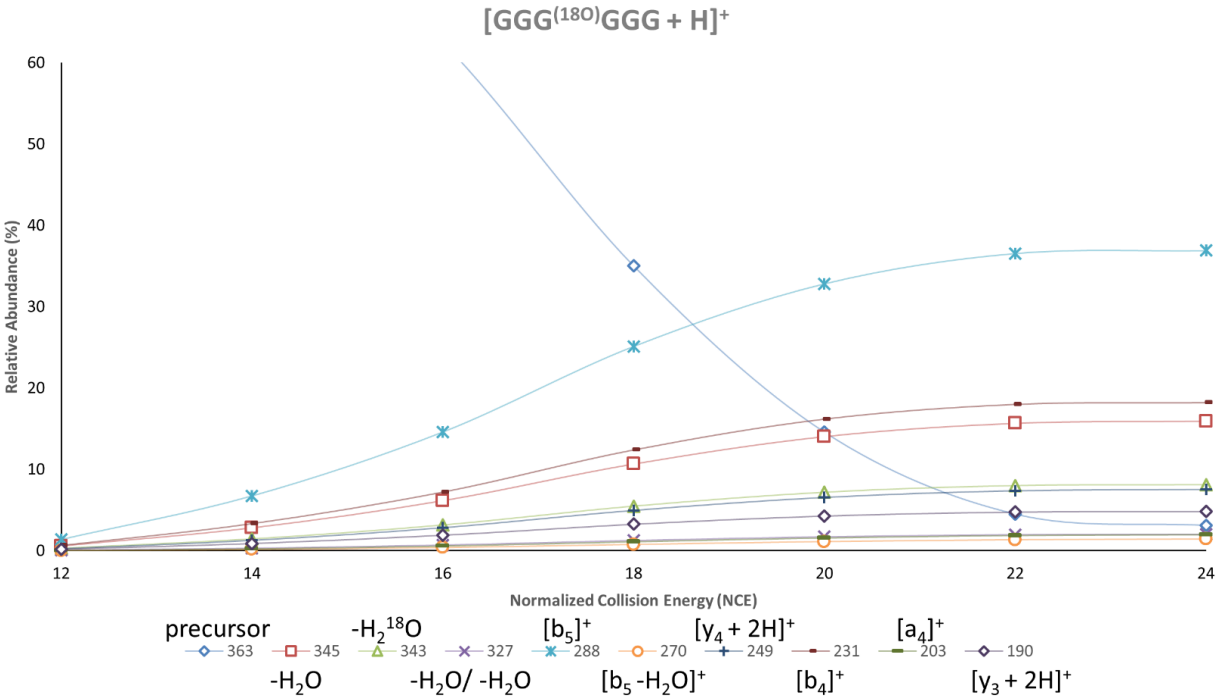
Appendix B. 10. CID spectrum of protonated hexaglycine with an ^{18}O label in the third amide. The CID spectrum was collected at a normalized collision energy of 21.



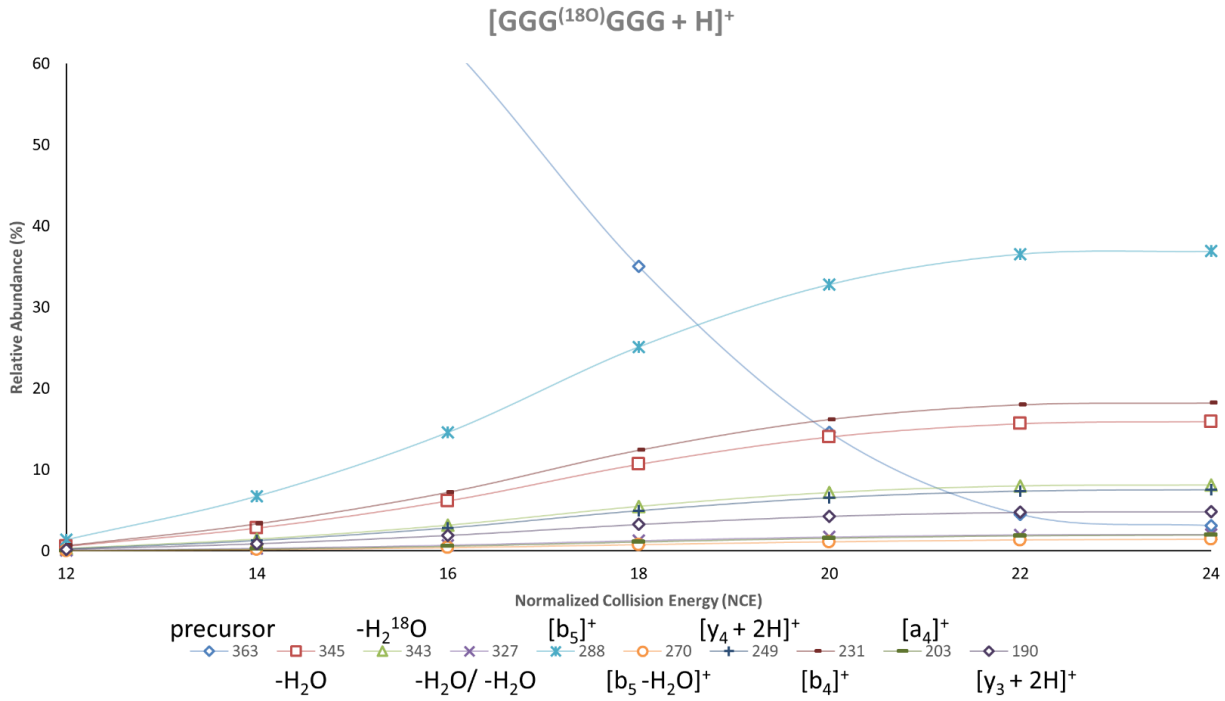
Appendix B. 11. Energy-resolved curves for ^{18}O labelled [GlyGlyGlyGlyGlyGly +H]$^+$ in the first amide.



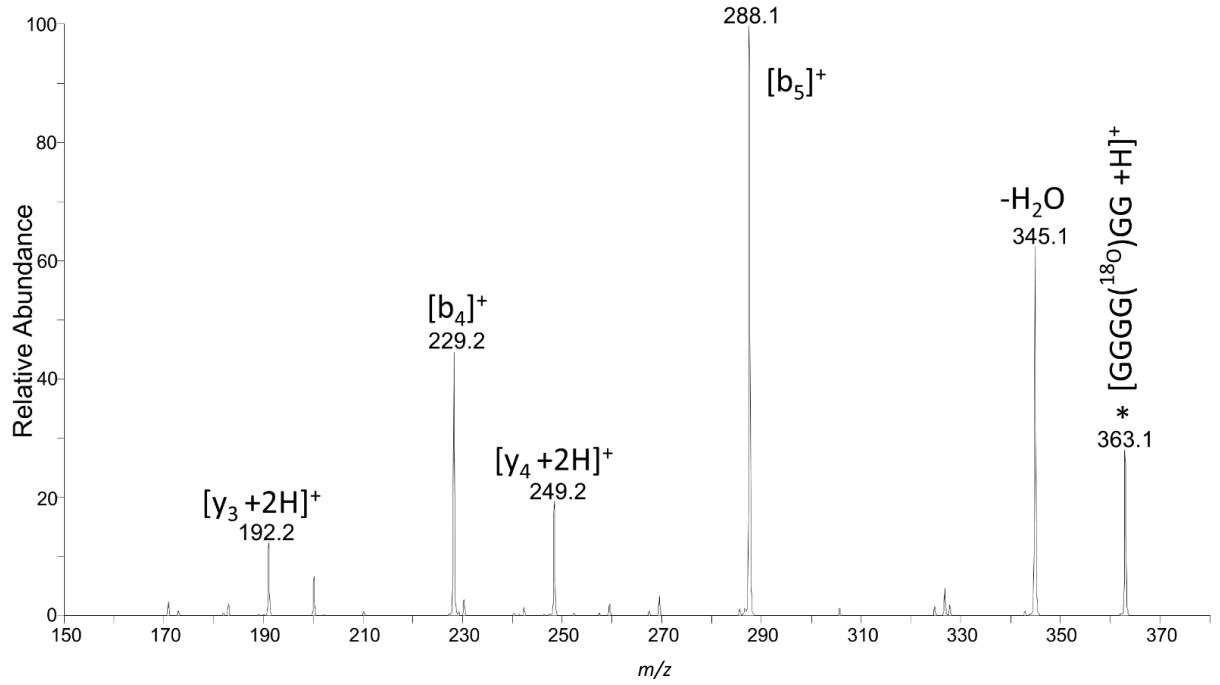
Appendix B. 12. Energy-resolved curves for ^{18}O labelled [GlyGlyGlyGlyGlyGly +H]$^+$ in the second amide.



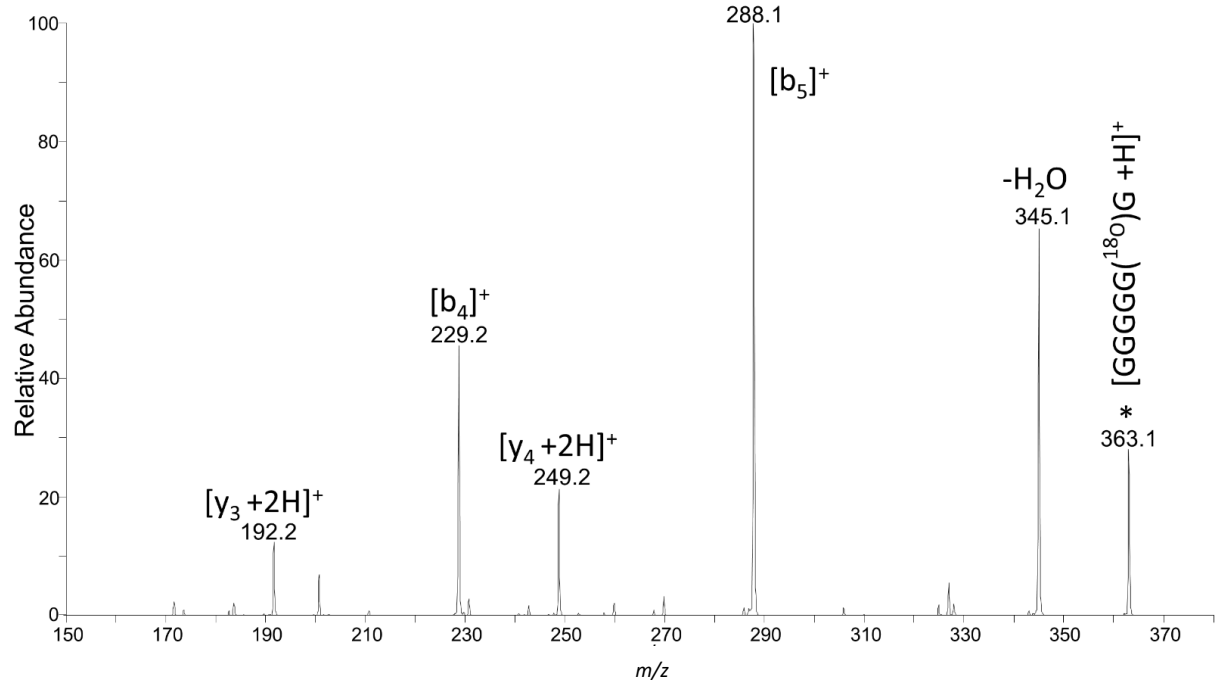
Appendix B. 13. Energy-resolved curves for ^{18}O labelled [GlyGlyGlyGlyGlyGly +H]$^+$ in the third amide.



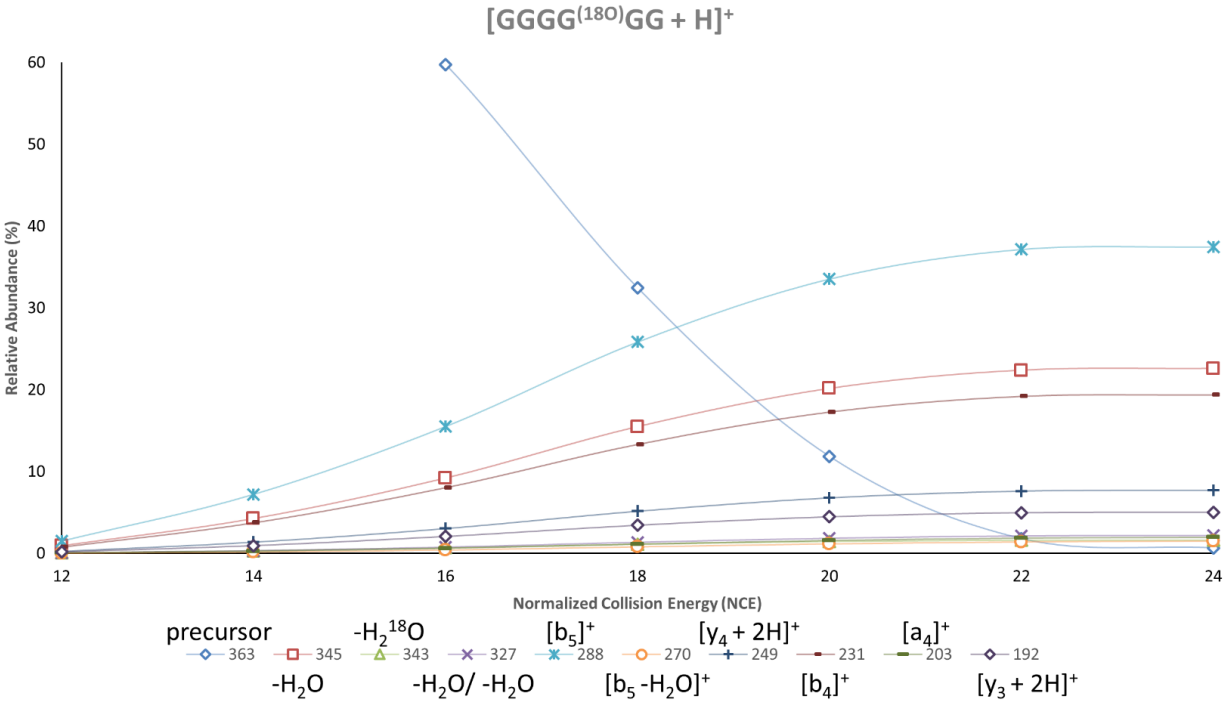
Appendix B. 14. Energy-resolved curves for ^{18}O labelled [GlyGlyGlyGlyGlyGly +H]$^+$ in the fourth amide.



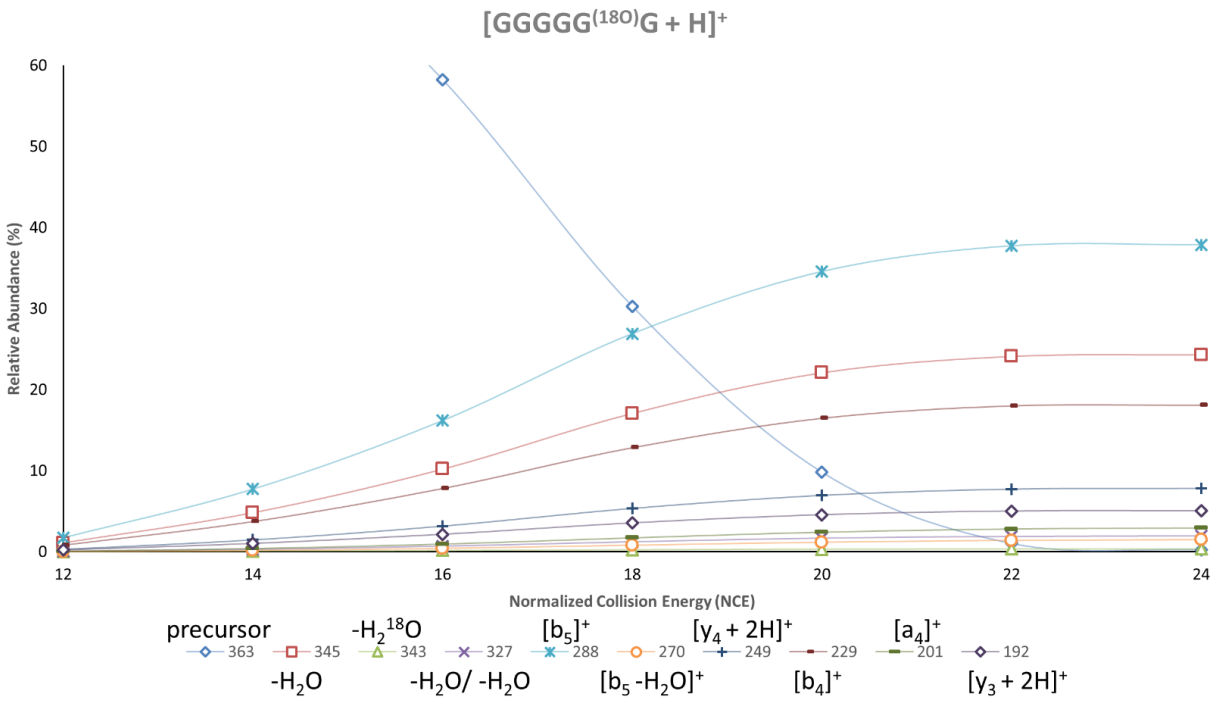
Appendix B. 15. CID spectrum of protonated hexaglycine with an ^{18}O label in the fourth amide. The CID spectrum was collected at a normalized collision energy of 21.



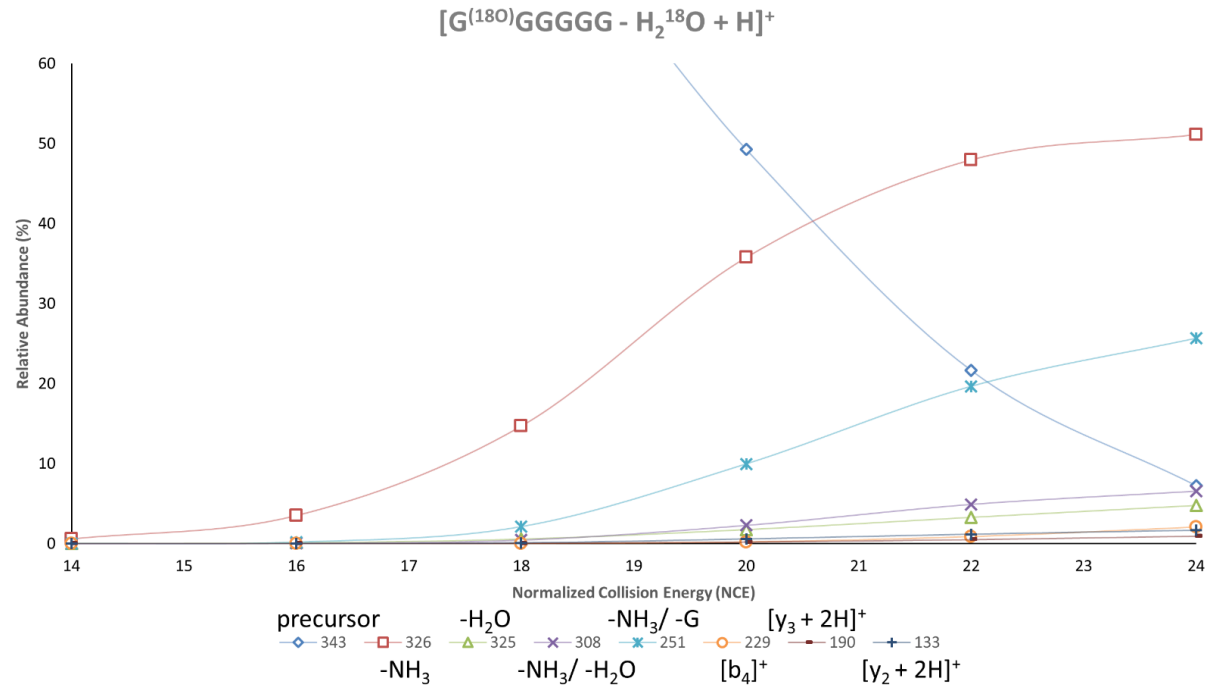
Appendix B. 16. CID spectrum of protonated hexaglycine with an ^{18}O label in the fifth amide. The CID spectrum was collected at a normalized collision energy of 21.



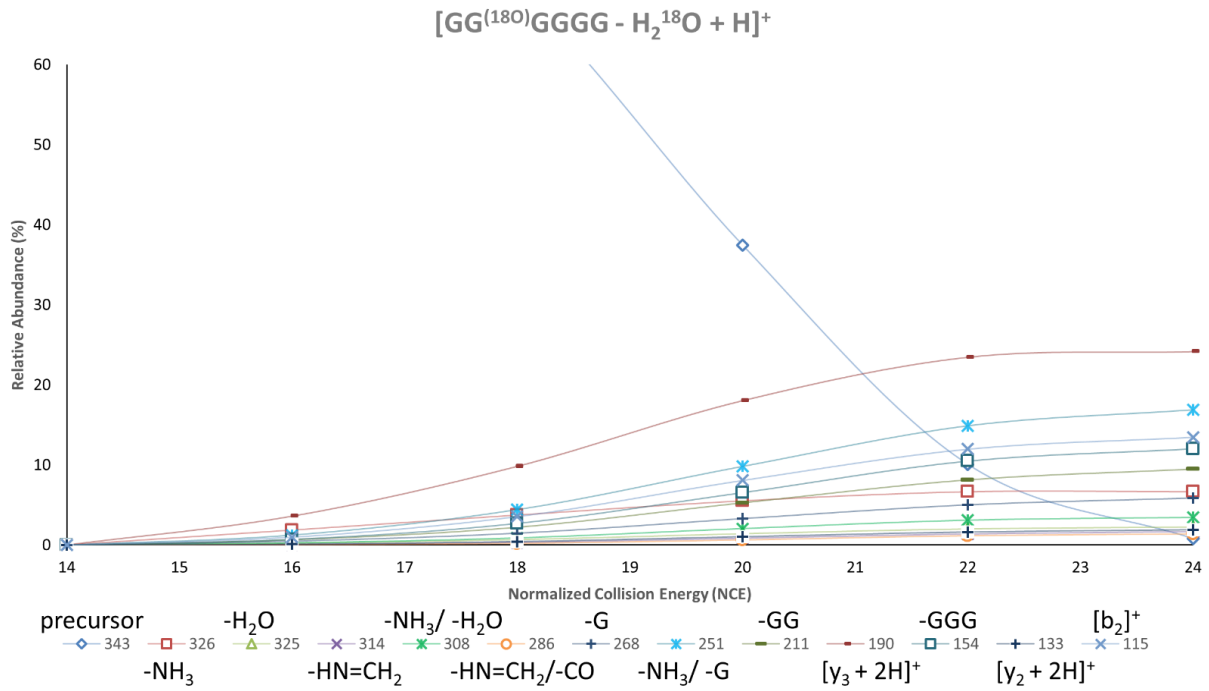
Appendix B. 17. Energy-resolved curves for ^{18}O labelled [GlyGlyGlyGlyGlyGly +H]$^+$ in the fourth amide.



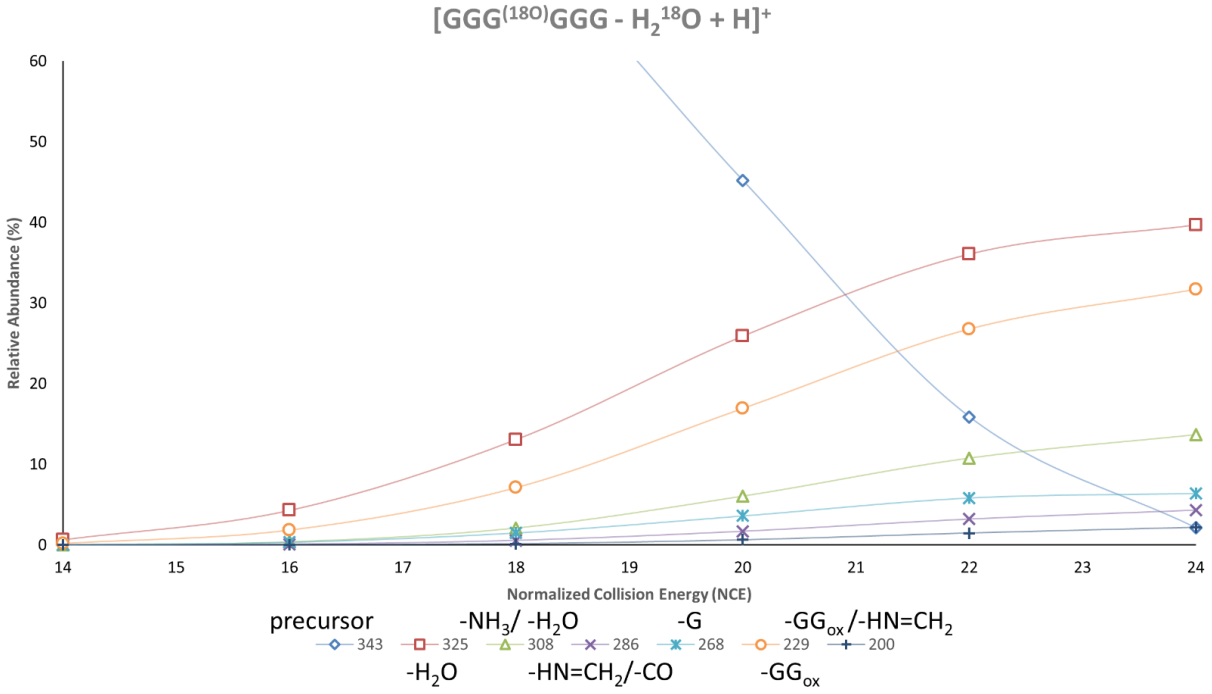
Appendix B. 18. Energy-resolved curves for [GlyGlyGlyGlyGlyGly + H]$^{+}$ ^{18}O-labelled in the fifth amide.



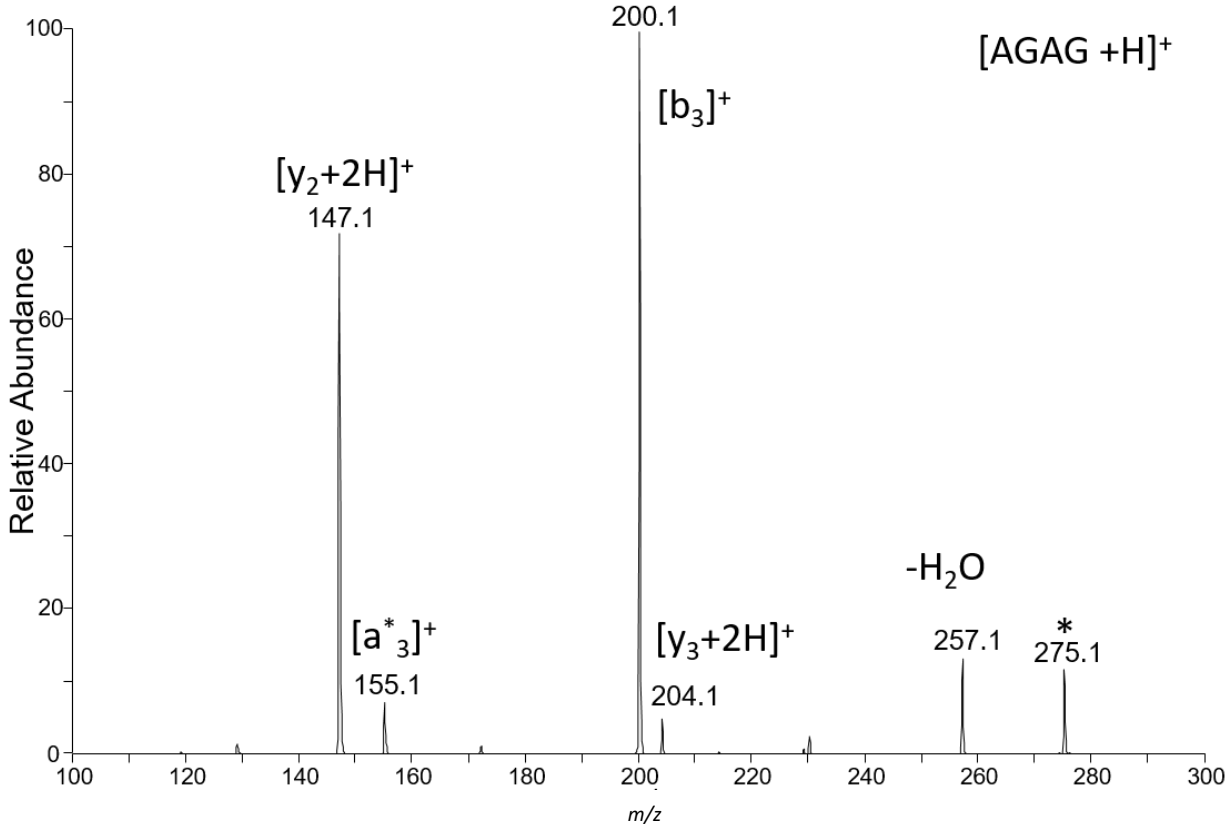
Appendix B. 19. Energy-resolved curves for ^{18}O labelled [Gly$^{(18O)}$GlyGlyGlyGlyGly -H$_2$^{18}O+H]$^{+}$ in the first amide.



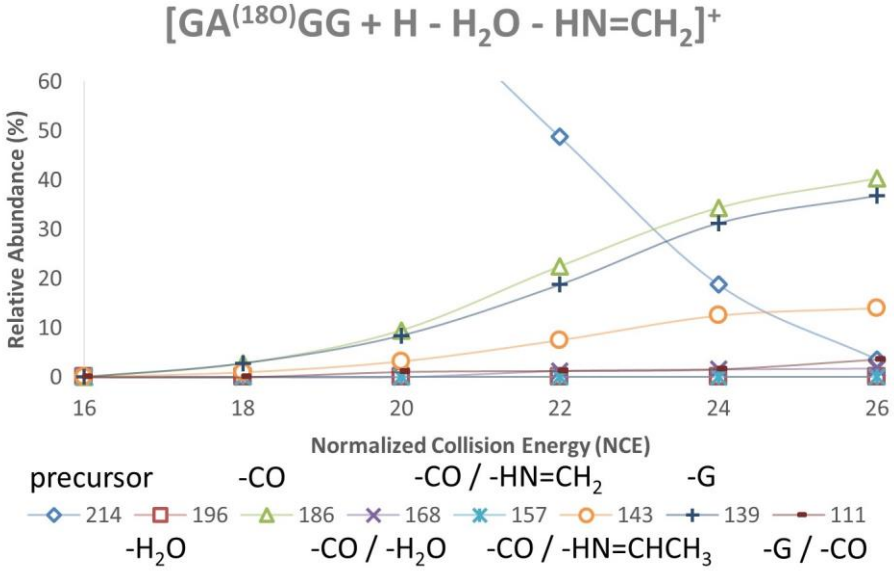
Appendix B. 20. Energy-resolved curves for ^{18}O labelled [GlyGly$^{(18O)}$GlyGlyGlyGly -$H_2{}^{18}O$+H]$^+$ in the first amide.



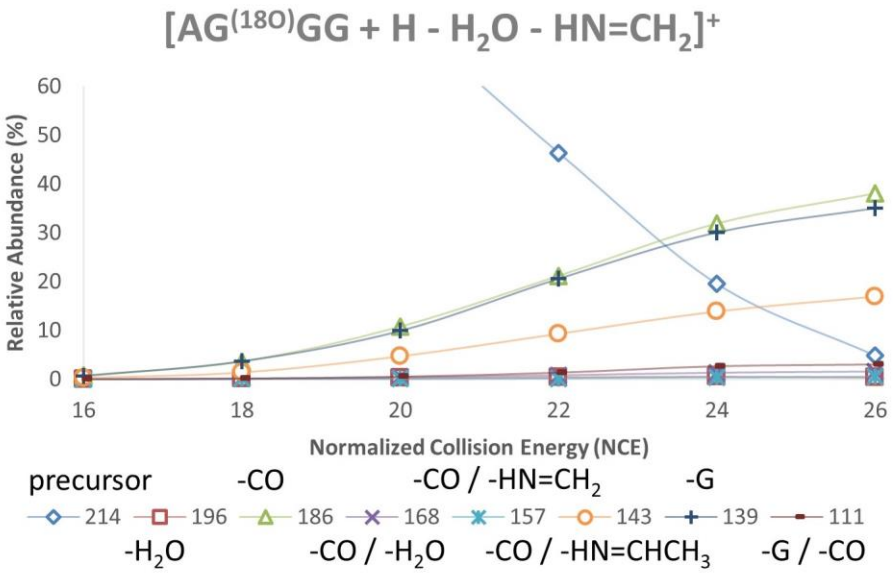
Appendix B. 21. Energy-resolved curves for ^{18}O labelled [GlyGlyGly$^{(18O)}$GlyGlyGly -$H_2{}^{18}O$+H]$^+$ in the first amide.



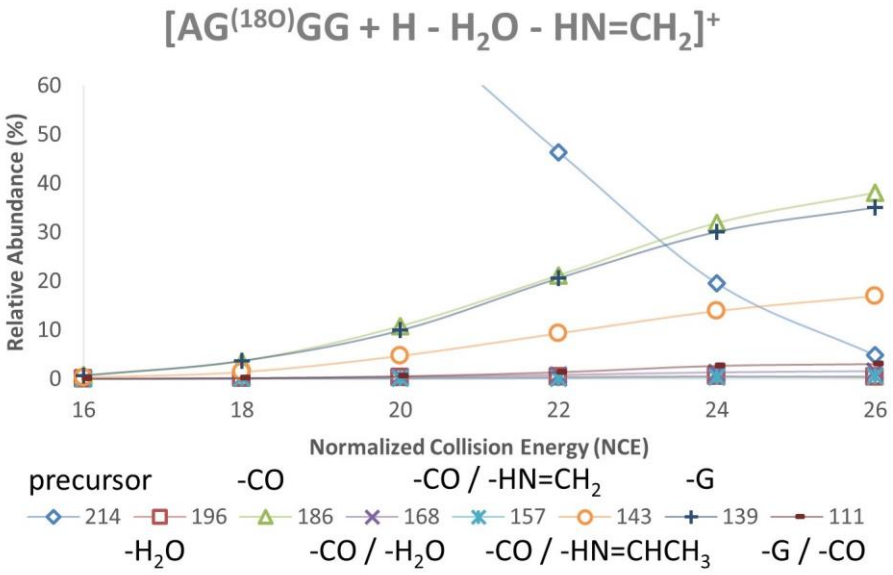
Appendix B. 22. CID spectra of protonated AlaGlyAlaGly. CID spectra were collected at a normalized collision energy of 20.



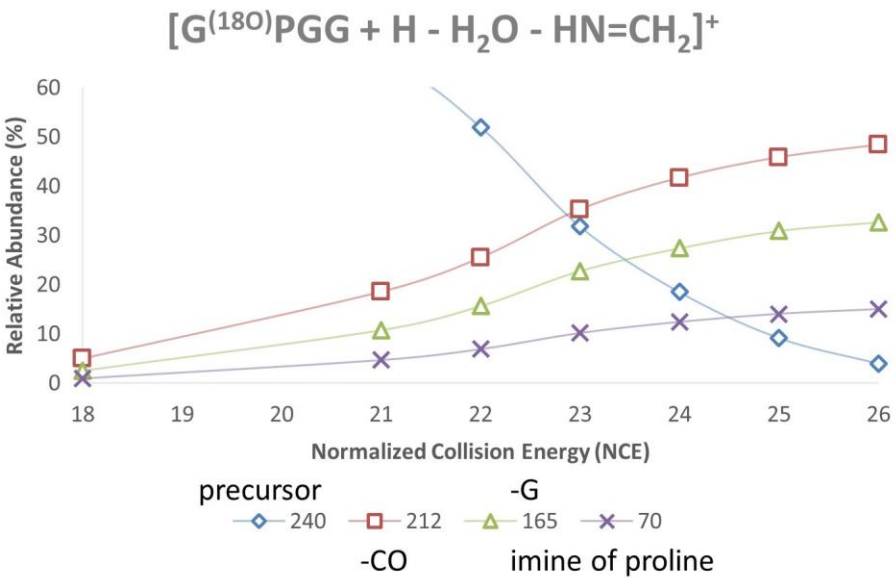
Appendix B. 23. Energy-resolved curves for ^{18}O labelled [GlyAla$^{(18O)}$GlyGly -$H_2{}^{18}O$ -$HN=CH_2$ +H]$^+$ in the first amide.



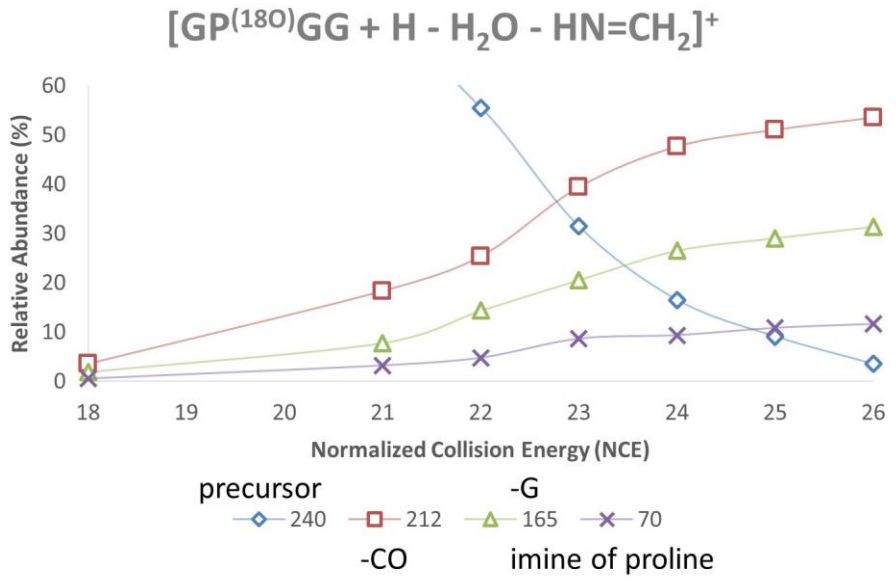
Appendix B. 24. Energy-resolved curves for ^{18}O labelled [Ala$^{(18O)}$GlyGlyGly -$H_2{}^{18}O$ -$HN=CH_2$ +H]$^+$ in the first amide.



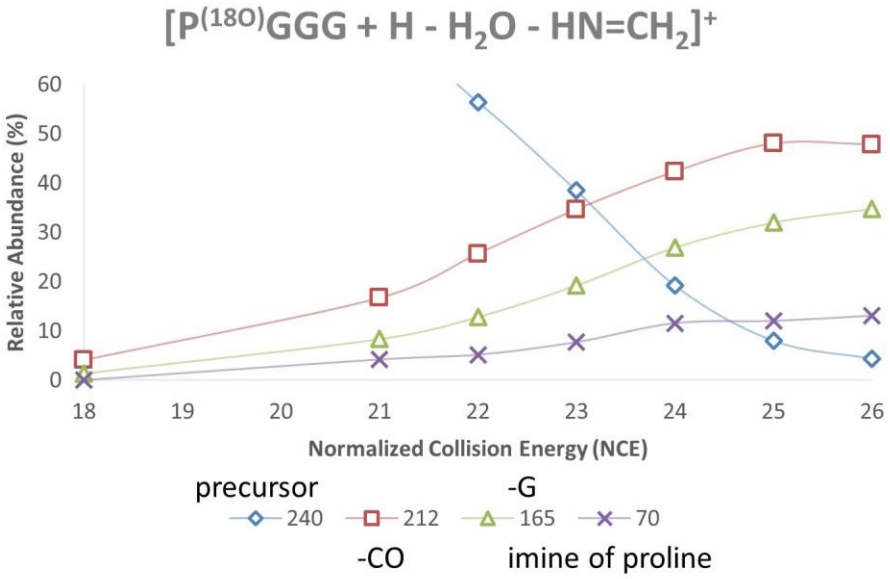
Appendix B. 25. Energy-resolved curves for ^{18}O labelled [AlaGly$^{(18O)}$GlyGly -$H_2{}^{18}O$ -$HN=CH_2$ +H]$^+$ in the first amide.



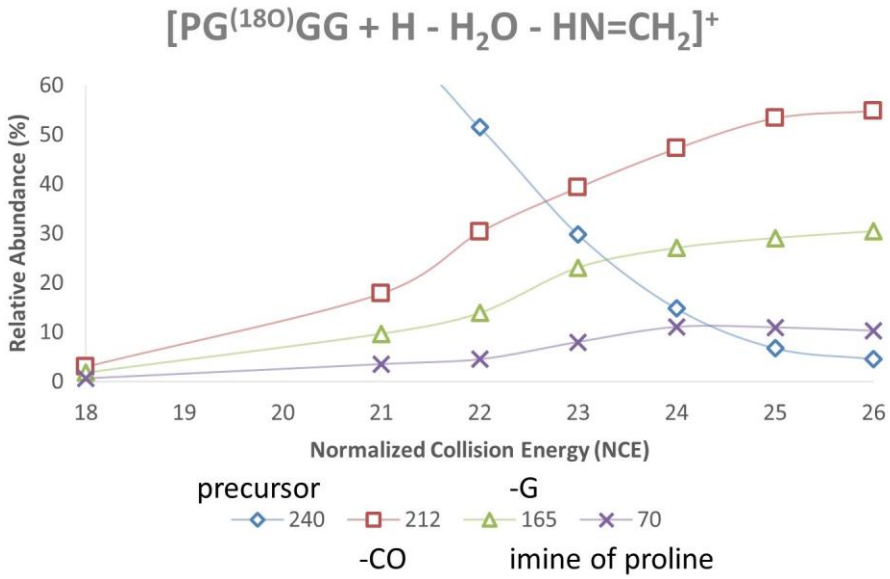
Appendix B. 26. Energy-resolved curves for ^{18}O labelled [Gly$^{(18O)}$ProGlyGly -$H_2{}^{18}O$ -$HN=CH_2$ +H]$^+$ in the first amide.



Appendix B. 27 Energy-resolved curves for ^{18}O labelled [GlyPro$^{(18O)}$GlyGly -$H_2{}^{18}O$ -$HN=CH_2$ +H]$^+$ in the first amide.



Appendix B. 28. Energy-resolved curves for ^{18}O labelled [Pro$^{(18O)}$GlyGlyGly -$H_2{}^{18}O$ -$HN{=}CH_2$ +H]$^+$ in the first amide.



Appendix B. 29. Energy-resolved curves for ^{18}O labelled [ProGly$^{(18O)}$GlyGly -$H_2{}^{18}O$ -$HN{=}CH_2$ +H]$^+$ in the first amide.