

5-2009

# Investigations of the Fragmentation Spectra of Peptides Containing Lysine and its Non-Protein Amino Acid Homologs

Matthew Cooke Bernier  
*College of William and Mary*

Follow this and additional works at: <https://scholarworks.wm.edu/honorstheses>

---

## Recommended Citation

Bernier, Matthew Cooke, "Investigations of the Fragmentation Spectra of Peptides Containing Lysine and its Non-Protein Amino Acid Homologs" (2009). *Undergraduate Honors Theses*. Paper 316.  
<https://scholarworks.wm.edu/honorstheses/316>

This Honors Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of W&M ScholarWorks. For more information, please contact [scholarworks@wm.edu](mailto:scholarworks@wm.edu).

# Investigations of the Fragmentation Spectra of Peptides Containing Lysine and its Non-Protein Amino Acid Homologs

A thesis submitted in partial fulfillment of the requirement for the degree of  
Bachelor of Science in Chemistry from the  
College of William and Mary

By

Matthew Cooke Bernier

Accepted for \_\_\_\_\_  
(Honors, High Honors, Highest Honors)

\_\_\_\_\_  
Chair  
  
\_\_\_\_\_  
  
\_\_\_\_\_  
  
\_\_\_\_\_

Williamsburg, Virginia  
April 2009

## Table of Contents

<b>Table of Contents</b> .....	<b>i</b>
<b>Acknowledgements</b> .....	<b>ii</b>
<b>List of Figures</b> .....	<b>iii</b>
<b>Abstract</b> .....	<b>iv</b>
<b>Chapter 1 – Introduction</b> .....	<b>1</b>
1.1 – Protein Structure and Sequencing.....	1
1.2 – Lysine and its Homologs: Thermodynamic and Structural Effects on Fragmentation.....	4
<b>Chapter 2 – Experimental</b> .....	<b>8</b>
2.1 – Synthesis of Peptides.....	8
2.2 – ESI Ion Trap Mass Spectrometry.....	17
<b>Chapter 3 – Results</b> .....	<b>20</b>
3.1 – Lysine and Ornithine- Containing Dipeptide Fragmentations.....	20
3.2 – Lysine, Ornithine, Daba, and Dapa-Containing Pentapeptide Fragmentations.....	32
<b>Chapter 4 – Discussion</b> .....	<b>44</b>
4.1 – Important Features of Dipeptide Fragmentations.....	44
4.2 – Important Features of Pentapeptide Fragmentations.....	47
4.3 – Conclusions and Future Work.....	51
<b>References</b> .....	<b>53</b>

## **Acknowledgements**

I would like to start by thanking Dr. J. C. Poutsma for the opportunity to work on such an interesting project, his assistance in helping me finish this thesis, and rewarding me such an outstanding undergraduate research experience as a whole. I need to thank my research group members who gave me help when I needed it and made the Ionlab such a terrific place to do research. A special thanks is in order for Erica Tullo and Mary Disa Raulfs who helped me so much in learning the basics of conducting this research, and to Brandon Thompson and Chris Collins for helping me synthesize peptides after spraining my ankle. I would also like to thank all my friends, who kept me from being too stressed during the difficult grinds, and especially my parents who I could not have possibly done this without.

## List of Figures

Figure 1.1 – Amino Acid Structure.....	1
Figure 1.2 – Peptide Fragmentation Nomenclature Diagram.....	4
Figure 1.3 – Lysine, Ornithine, Daba, and Dapa.....	6
Figure 2.1 – Solid-Phase Peptide Synthesis Vessel Set-up.....	10
Figure 2.2 – Removal of Fmoc Protecting Group From N-Terminal.....	11
Figure 2.3 – Amino Acid Coupling-Step.....	13
Figure 2.4 – Simplified Diagram of ESI Ion Trap Mass Spectrometer.....	18
Figure 3.1 – CID Spectrum of Alanine-Lysine.....	20
Figure 3.2 – CID Spectrum of Alanine-Ornithine.....	21
Figure 3.3 – CID Spectrum of Lysine-Alanine.....	22
Figure 3.4 – CID Spectrum of Ornithine-Alanine.....	23
Figure 3.5 – CID Spectrum of Arginine-Lysine.....	24
Figure 3.6 – CID Spectrum of Arginine-Ornithine.....	25
Figure 3.7 – CID Spectrum of Lysine-Arginine.....	26
Figure 3.8 – CID Spectrum of Ornithine-Arginine.....	27
Figure 3.9 – CID Spectrum of Proline-Lysine.....	28
Figure 3.10 – CID Spectrum of Proline-Ornithine.....	29
Figure 3.11 – CID Spectrum of Serine-Lysine.....	30
Figure 3.12 – CID Spectrum of Serine-Ornithine.....	31
Figure 3.13 – CID Spectrum of Ala-Ala-Lys-Ala-Ala.....	32
Figure 3.14 – CID Spectrum of Ala-Ala-Orn-Ala-Ala.....	33
Figure 3.15 – CID Spectrum of Ala-Ala-Daba-Ala-Ala.....	34
Figure 3.16 – CID Spectrum of Ala-Ala-Dapa-Ala-Ala.....	35
Figure 3.17 – CID Spectrum of Ala-Lys-Ala-Ala-Ala.....	36
Figure 3.18 – CID Spectrum of Ala-Orn-Ala-Ala-Ala.....	37
Figure 3.19 – CID Spectrum of Ala-Daba-Ala-Ala-Ala.....	38
Figure 3.20 – CID Spectrum of Ala-Dapa-Ala-Ala-Ala.....	39
Figure 3.21 – CID Spectrum of Ala-Ala-Ala-Ala-Lys.....	40
Figure 3.22 – CID Spectrum of Ala-Ala-Ala-Ala-Orn.....	41
Figure 3.23 – CID Spectrum of Ala-Ala-Ala-Ala-Daba.....	42
Figure 3.24 – CID Spectrum of Ala-Ala-Ala-Ala-Dapa.....	43
Figure 4.1 – Serine-Ornithine Dipeptide.....	45
Figure 4.2 – Oxazolone b <sub>2</sub> ion.....	46

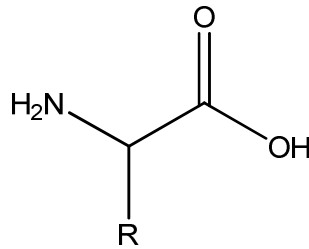
## Abstract

In this study, we investigated the fragmentation spectra of dipeptides containing lysine or ornithine and pentapeptides containing lysine, ornithine, Daba, or Dapa. Lysine is one of the 20 protein amino acids whereas ornithine, Daba, and Dapa are non-protein amino acids that resemble lysine. Each of these homologs has an amine side chain that is consecutively shorter than lysine's, making their differing effects on fragmentation interesting to observe. By looking at the fragments created by each peptide upon collision-induced dissociation (CID) in an ESI ion trap mass spectrometer, we could compare the fragments between the peptides and interpret those differences as a result of structural differences. Performing CID on lysine containing peptides, produced many more significant fragments than ornithine, and to similar extent, Daba, and Dapa containing peptides. Ornithine-containing pentapeptides consistently gave a dominant b fragment in which ornithine was the C-terminal of that fragment. Dapa and Daba-containing pentapeptides seemed to favor the loss of water from the parent in fragmentation, yet they also produced peaks that were in good agreements with the lysine containing pentapeptides. The amount of peaks for each amino acid did not follow a set pattern from lysine to Dapa, as the Daba-containing did at times contain more significant peaks than the ornithine. These comparisons are only a preliminary investigation into better understanding the mechanism with which amine side chains can affect fragmentation of the peptide backbone during CID.

## Chapter 1: Introduction Section

### 1.1 Protein Structure and Sequencing

There is no biological unit that is as essential to the everyday function of cells as the protein. The role of proteins in the body includes many different tasks such as transporting important molecules in and out of cells, acting as enzymes to catalyze biochemical reactions, regulating cell growth through hormone control, and acting as structural support for the cell [1]. The building blocks of all proteins used in the body are the 20 protein amino acids. These 20 amino acids are made up of a backbone with an amine group on one end, carboxylic acid group on the other, and a carbon with both a hydrogen molecule and a side chain specific to each amino acid attached. The basic structure of an amino acid is shown below on Figure 1.1:



**Figure 1.1:** Amino Acid Structure

To combine amino acid molecules into peptide chains, a condensation reaction occurs where the amine of one amino acid attaches to the carboxylic acid group of another. In this process, the hydroxyl group from the carboxylic acid and a hydrogen molecule from the amine are removed as water and a C-N peptide bond is formed. This unique bond, due to its slight double bond nature, gives amino acids a planar shape and allows for a great deal of flexibility, while retaining much needed stability [1].

Fully mature proteins are formed from long polypeptide chains of amino acids. First, the cell's DNA expresses a gene which encodes a linear series of amino acids to form a peptide chain. These long chains are synthesized by RNA and the repeated patterns of amino acids develop secondary structures like the alpha-helix and beta-pleated sheets due to the flexibility of the peptide bond and through hydrogen bonding interactions between the amino acid side chains [1]. With the shapes formed by these structures, peptides can wrap and fold around themselves, using combinations of these secondary structural patterns to develop a three dimensional or tertiary structure. With a distinguishing shape and structure, a protein unit can then interact with the molecules around it to carry out specific tasks, like form a channel through a cell membrane or create a hydrophobic cavity in which to perform an enzymatic digestion [1]. Finally, many polypeptide chains can combine to form proteins known as oligomers, which require the addition of multiple peptide subunits in order to carry out their function. Hemoglobin is a great example of an oligomer. This protein structure requires several interacting polypeptide subunits to carry out its task of transporting oxygen around the body. With the order of amino acids in a protein so crucial in developing its ultimate shape and function, it is no wonder that so much time is put in to sequencing them.

The study of proteomics, or protein sequencing, has become an important issue for biological science. The earliest way to sequence peptide chains was using Edman degradation. In this technique, the amino acid sequence is uncovered by identifying the amino acids one by one from the N-terminus. Each amino acid is chemically removed in step-wise fashion until the C-terminal amino acid is reached. The problems with this method are that it takes an incredibly long amount of time and that it adds the necessary task of isolating specific peptide fragments for sequencing. Furthermore, any modifications to the peptide can cause serious problems in

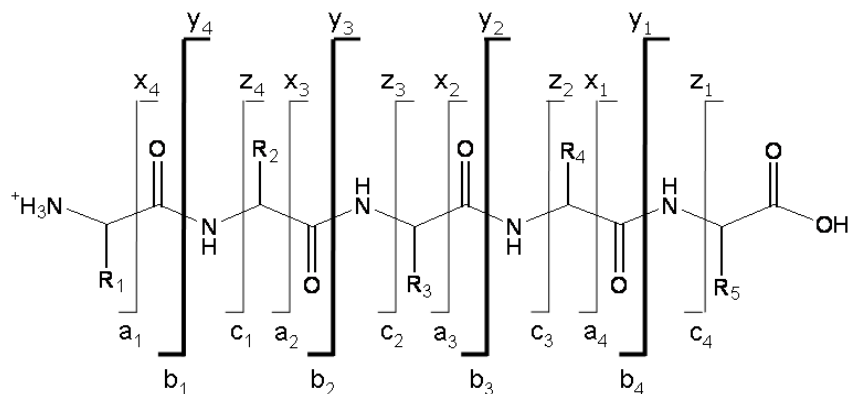


performing the degradation [2]. Due to these difficulties, there is absolutely no way that a long polypeptide could be sufficiently sequenced, much less the complex protein make-up of an entire cell.

A much more efficient way to sequence proteins was developed in the 1990's which utilizes mass spectrometry instrumentation [2]. The most common way to sequence proteins with mass spectrometry is known as the bottom-up or shotgun approach. In this method, a protein is digested using an enzyme, like trypsin, which breaks it up into several peptide fragments. Liquid chromatography is used to separate this peptide mixture before being ionized by a source, and then placed it into a mass spectrometer [3]. The mass spectrometer performs analysis on the peptides by separating them, using radio frequencies and direct currents, by their mass to charge ratios. In order to get the most coverage, the MS/MS mode is used. This technique, available to ion trap mass spectrometers, allows specific mass-to-charge ratios to be isolated and then fragmented using collision-induced dissociation (CID) [4]. A fragmentation pattern for the series of peptides from a protein is produced which gives a "fingerprint" that is unique to that protein. According to Brechi et al. this "fingerprint" analysis, using CID, allows for 97-100% of a protein sequence to be identified [5].

Since protein sequencing using fragmentation patterns has become such an important area of research, the ions produced from fragmentation have become heavily studied and their nomenclature has been standardized. The nomenclature for peptide fragmentation, developed by Roepstorff and Kohlman, classifies fragment ions as resulting from the breaking of one of the three backbone bonds and whether the charge stays with the N or C-terminus. The  $a_n$ ,  $b_n$ , and  $c_n$  ion fragments result from the charge staying with the N-terminal fragment, whereas the  $x_{m-n}$ ,  $y_{m-n}$ , and  $z_{m-n}$  ions result from the charge staying with the C-terminal fragment. The  $n$  subscript

denotes the number of R-groups present on the fragment ion while the m subscript represents the total amount of R-groups on the peptide chain [6,7]. A diagram of this nomenclature can be seen below on Figure 1.2:



**Figure 1.2:** Peptide Fragmentation Nomenclature Diagram [6]

From a CID fragmentation, peptide chains under study will commonly fragment into b, y, and at times, a ions [8]. It is from this mixture of fragment ions that the unique fragmentation patterns of various peptides fragment can be better understood. This is done by comparing the ratios of the fragment ions in peptides that are similar but containing small differences. The spectra differences can highlight the effects of these modifications, and ultimately unique traits of fragmentation can be elucidated.

## 1.2 Lysine and its Homologs: Thermodynamic and Structural Effects on Fragmentation

Wysocki et al. have proposed that the fragmentation of many peptides most commonly occur by what is known as the mobile proton model [9]. In this model, the collision energy produced by CID promotes a hydrogen on the peptide to move across the peptide backbone and directly effect the cleavage of specific peptide bonds. Furthermore, this mobile proton heavily

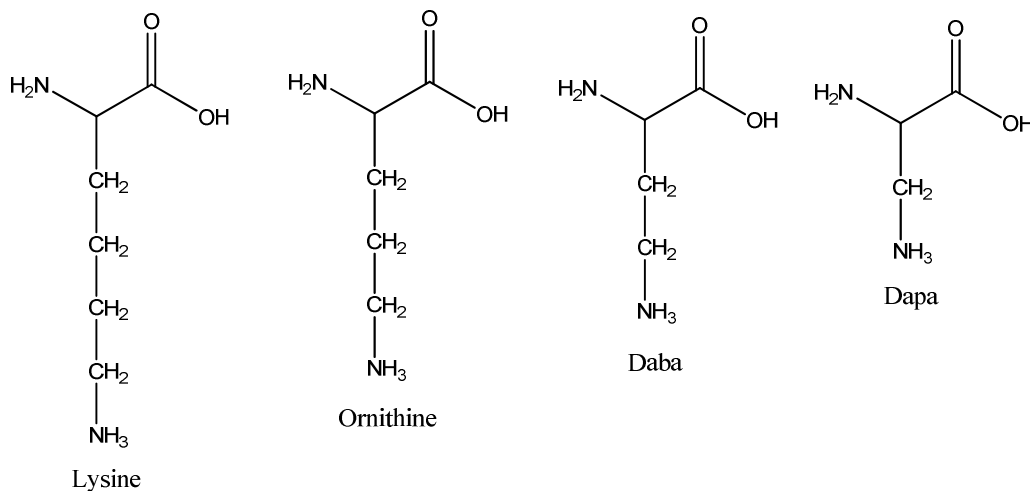
influences the fragmentation energy requirements and is strongly affected by the amino acids that make up the peptide chain [9,10]. The importance of amino acid side chains is most commonly attributed to its thermodynamic properties. When a certain side chain has an amine group and has a large proton affinity, this usually can sequester the proton, causing the presence of that amino acid to direct the fragmentation. Hunt et al. and Paisz et al. both have stated that the basicity of certain peptides cause the proton to move there while in the gas phase and if peptides are particularly basic, they may keep the proton before fragmentation on their amine side chains [9, 11, 12]. The most basic of the 20 protein amino acids are arginine, lysine, and histidine, which also have the most proton affinity in the gas phase [13, 14, 15, 16, 17].

It is the goal of this study to look at the fragmentation spectra of peptide chains that contain identical peptide sequences except for slight modifications differing in either thermochemistry or structure, and compare them to understand better how these changes affect fragmentation. Lysine, the second most basic of the protein amino acids (PAA), is an excellent candidate for this kind of analysis. Furthermore, one of the most commonly used enzymes for breaking up large proteins before using tandem mass spectrometry is trypsin. This is significant in that its digestion mechanism ends up cleaving peptides at the C-terminus of lysine. This means that every peptide chain being studied in a trypsin digested protein will contain the lysine amino acid on its C-terminus [18].

Besides being an essential amino acid in the cleaving of peptide bonds during enzymatic digestion for proteome studies, lysine also has three non-protein amino acid (NPAA) homologs that resemble it very closely in structure. While there are the 20 PAAs that are coded genetically by the body, there are also several hundred amino acids in nature that are not used by the body. Most of these are synthesized by plants and can sometimes be unsafe to ingest. This is because

many of these NPAAAs are very close in structure to certain PAAs and these can be misincorporated into peptides and ultimately mature proteins, which may result in the proteins not being able to perform their functions correctly. An example of this is arginine and its NPAA, canavanine. When leguminous plants which contain canavanine are ingested, there is evidence that this can lead to the development of lupus [19]. It is then reasonable to study the effects of slight changes in amino acids so that protein fragmentation studies can be used to identify when these misincorporations have occurred in the body.

The NPAAAs that resemble lysine include ornithine, 2,4-diaminobutanoic acid (Daba), and 2,3-diaminopropanoic acid (Dapa). They consecutively lose a methylene group from the amine side chain, where Dapa, the smallest of the three, has a single methylene followed by the amine. These molecules can be seen below on Figure 1.3:



**Figure 1.3:** Lysine, Ornithine, Daba, and Dapa

The strong basicity of these molecules can be attributed to the fact that they contain two amine groups. These amine groups can form intramolecular bonds, making very stable structures when the molecule has been protonated [20]. Poutsma et al. have calculated the proton affinities of

these four molecules using the extended kinetic method and corresponding high level theoretical calculations. A value of  $1004.2 \pm 8.0$  kJ/mol was suggested for lysine while values of  $1001.1 \pm 6.6$  kJ/mol,  $975.8 \pm 6.6$  kJ/mol,  $950.2 \pm 7.1$  kJ/mol were suggested for ornithine, Daba, and Dapa respectively [21]. While these values do differ somewhat, especially when comparing lysine's value to that of Daba and Dapa, they are still in a very reasonable range. The close proximity of these gas phase proton affinity measurements leads to the ability to compare how their side chain lengths may affect peptide fragmentations in mass spectrometer CID experiments. It is most reasonable to observe these effects between ornithine and lysine since their proton affinities are almost identical considering experimental error and that distinct differences in their fragmentation spectra would clearly be on account of the loss of methylene and the shortening of the side chain.

## Chapter 2: Experimental Section

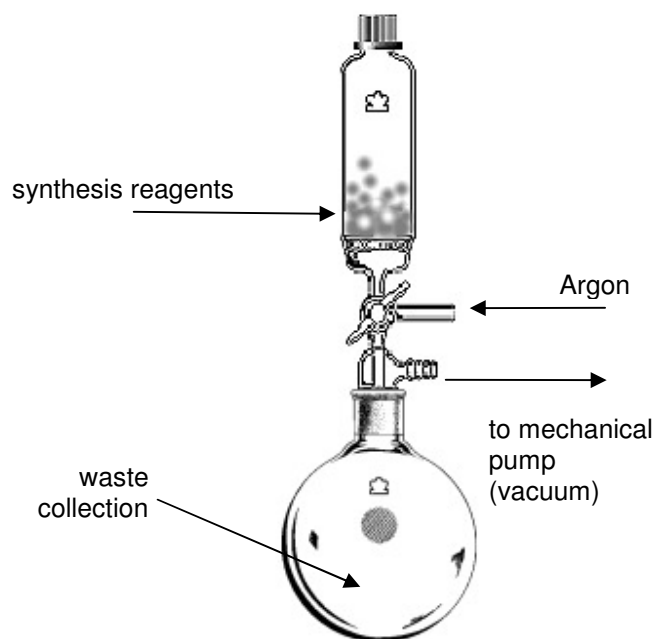
### 2.1 Synthesis of Peptides

All peptides were made using the Fmoc solid phase synthesis method. The exact procedures outlined in this experiment were suggested by Chan and White [22], Brechi [23], and an alternative procedure inspired by a method from Hood [24]. In this type of synthesis, the C-terminal amino acid is attached to an insoluble polyethylene bead, also known as a resin, on its C-terminus. In this case, a Wang resin (4-hydroxymethylphenoxy) was used as the support for synthesis. The amine side of all amino acids is protected with Fmoc (fluorenylmethyloxycarbonyl) to keep unwanted reactions from occurring before coupling to the following amino acid. First, the resin-attached amino acid is deprotected by removing the Fmoc from its N-terminus with piperidine. A second Fmoc-protected amino acid with a free C-terminus is subsequently coupled to the free amine group using the coupling reagents. This process repeats with deprotection followed by coupling until all the amino acids desired in the peptide chain are added. Following the addition of all necessary amino acids, a cleaving process occurs where the C-terminal amino acid is separated from the beaded resin. It is important to note that while some of the amino acids involved in the synthesis only required the Fmoc protecting group on their amine side, some amino acids required protecting groups on their side chains as well. Most notably this included the side chains of lysine and its homologs in order to make sure there were no unwanted reactions with the amine groups on these side chains. These side chains are protected by using protecting groups like tert-butyloxycarbonyl (Boc) and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf).

Although synthesis of ornithine Wang resin was attempted, ultimately all synthesis reagents, Fmoc-protected amino acids, and Fmoc Wang resins were purchased and used without any modifications. Fmoc-Ala-OH, Fmoc-Pro-OH, Fmoc-Orn(Boc)-OH, Fmoc-Dapa(Boc)-OH, Fmoc-Lys(Boc)-Wang Resin, Fmoc-Ala-Wang Resin, Fmoc-Val-Wang Resin, Fmoc-Phe-Wang Resin, Fmoc-Dapa(Boc)-Wang Resin, Fmoc-Daba(Boc)-Wang Resin, Fmoc-Gly-Wang Resin, Fmoc-Orn(Boc)-Wang Resin, Fmoc-Arg(Pbf)-Wang Resin, and Fmoc-Pro-Wang Resin were all purchased from ChemPep Inc. in Miami, Florida. Fmoc-Daba(Boc)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Gly-OH, Fmoc-Val-OH, and the coupling reagents benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBop) and 2-(6-Chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) were all purchased from Novabiochem in La Jolla, California. The coupling reagents N-hydroxybenzotriazole (HOBT) and N,N-diisopropylethylamine (DIPEA), the cleaving agents trifluoroacetic acid (TFA) and triisopropyl silane, and 2,6-dichlorobenzoyl chloride were purchased from Sigma-Aldrich, while methylene chloride (DCM), dimethylformamide (DMF), piperidine, and ethyl ether were purchased from Fischer Scientific. Acros Organics also provided some DMF.

Synthesis was performed in a Kontes 50 mL cylindrical peptide synthesis vessel with a sealed in fritted disc. The disc allowed the insoluble resin beads to stay in the tube while solutions used to deprotect and couple to the beads could be pumped out of the column to allow the synthesis to continue. Two glass inlets were connected below the disc with a three way 2 mm PTFE stopcock connecting the three parts. The 8 mm side arm was connected to an argon tank used for bubbling the solution, while the bottom #1 vacuum hose connection was fitted with vacuum tubing and joined to either a pump or house vacuum set-up. This set-up allowed either

argon gas to flow up into the column or, by turning the valve, to pump out the liquid from the tube down a 20/40 drip joint into a 2000 milliliter round bottom flask, which is there to collect waste. This set-up can be seen below on Figure 2.1:



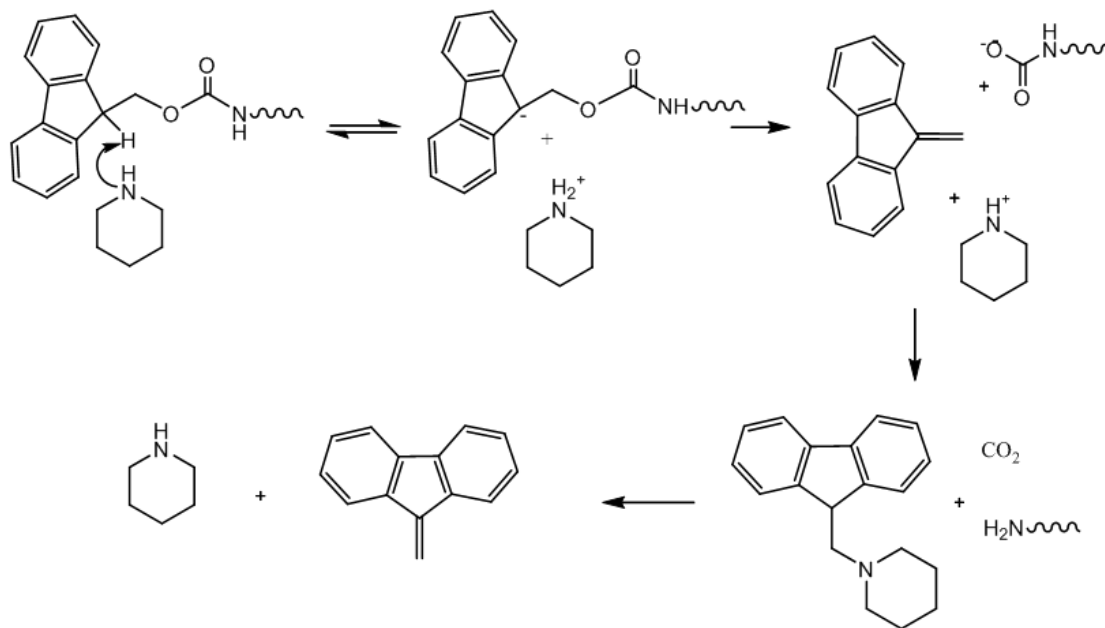
**Figure 2.1:** Solid-Phase Peptide Synthesis Vessel Set-Up

In order to explain the synthesis procedure in detail, a step-by-step account of the synthesis of AAKAA (Ala-Ala-Lys-Ala-Ala) will be described, as it was the first pentapeptide that was synthesized for this project. The first step is to put a roughly 2-10  $\mu$ -molar amount of Fmoc-Ala-Wang (0.57 mmol/g loading capacity) resin into the synthesis vessel and add to it a 50:50 mixture of DCM and DMF. The amount added was enough to fill the vessel roughly a centimeter so that the beads were sufficiently immersed in the mixture and the argon flow rate was set so that a small amount of bubbling could be seen. This process is done in order to swell the beaded resin so that it would be easier to access the amino acid to perform the deprotection and coupling steps. Immediately following the swelling step, the 50:50 mixture is vacuumed



into the waste chamber and two one-minute washes are performed with DMF in order to prepare the resin for the removal of the Fmoc protecting group.

The deprotection of the alanine on the Wang resin, or in other words, the removal of the Fmoc group from the amine, is facilitated by a 20:80 mixture of piperidine and DMF. The mixture is added in roughly the same amount as the swelling mixture so as to cover all the beads, and the solution is perturbed with the beads for a total of 25 minutes (5 min cycle followed by 20 min cycle). The reacting agent in this deprotection is the piperidine where the key step is the production of an aromatic cyclopentadiene intermediate from the deprotonation of the fluorene ring on the Fmoc protecting group. The cyclopentadiene intermediate transitions quickly into the form dibenzofulvene and is then attached to the piperidine, allowing the amine on the amino acid to drop the carboxyl group and prepare itself for the coupling of the next amino acid. (22) This reaction can be seen below in Figure 2.2:

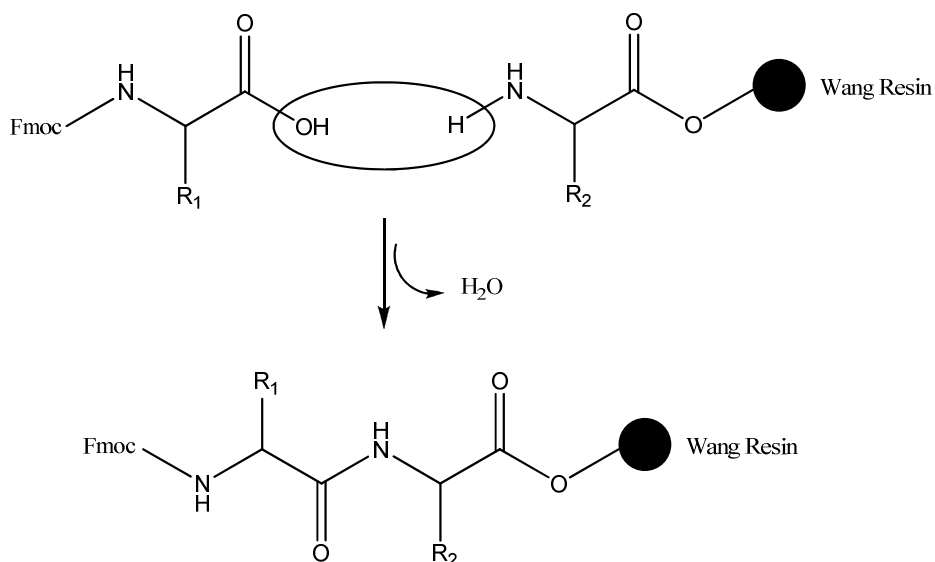


**Figure 2.2:** Removal of Fmoc Protecting Group from N-Terminal [22]

Following deprotection, two one-minute washes with DMF then four one-minute washes with DCM are done on the beads. The DMF serves to remove any residual piperidine, while the DCM dries out the beads for the next step. In order to identify whether the deprotection was successful and that there are a sufficient amount of free amines to couple with, the ninhydrin test is performed. The test, developed by Kaiser et al., is based on heating a small sample of the beaded resin in a solution of 1g of ninhydrin (triketohydrindene hydrate) per 10 mL of n-butanol to about 120 °C and observing the color of the solution after some time has passed [25]. This is usually done by poking the film of beads on the fritted disc with a small glass pipette, which picks up a few of the beads. The pipette is placed in a test tube with a small amount of the ninhydrin mixture, and the whole test tube is placed in a beaker of heated water for 4-6 minutes. If the solution stays yellow, then no free amines are present while an indigo blue color shows that there are free amines. After the deprotection step, the indigo blue color is expected, since it should have removed the Fmoc-protecting group and created free amines. Once this color is observed after 5 or so minutes have passed, then the synthesis procedure continues with the coupling step. If the color is yellow at this time, then more of the piperidine: DMF mixture is added to the vessel for 5-10 minutes. As Chan and White have described, this test is not able to detect deprotection of proline due to its secondary amine nature, and that larger chains of amino acids can often show a “false negative” [22].

After successful deprotection, the alanine attached to the resin is coupled to an Fmoc-Alanine-OH, which is the next amino acid from the C-terminus, in the pentapeptide AAKAA. First, two one-minute washes with DMF are done to saturate the beads which were dried by DCM. The coupling reagents are then added into the vessel and are gently bubbled with the beads for 90 minutes. To make this coupling solution, a 2 molar equivalent amount of the Fmoc-

Alanine-OH, HOBT, and PyBop, and a 4 molar equivalent amount of DIPEA, were placed in the synthesis vessel with some DMF. Chan and White suggested that 5 eq. amino acid, HOBT, and PyBop be used and 10 eq. DIPEA at a total time of one hour. However, it was suggested by Brechi that only 2 eq. amino acid, HOBT, and PyBop, while 4 eq. DIPEA was needed for the reaction to work at an extended time of 1.5 hours. The most efficient way to deliver the reagents into the vessel is to put the solid amino acid, PyBop, and HOBT into a vial and dissolve them with a portion of DMF. Then the solution is poured into the vessel and the DIPEA is injected directed into this mixture with a syringe while agitation with argon is occurring. In this reaction, the Fmoc-Alanine-OH is coupled with the beaded alanine in a condensation reaction, where H<sub>2</sub>O is produced from the OH and an H from the amine and a C-N bond is formed between the free amine and carbonyl. This reaction can be seen below in Figure 2.3:



**Figure 2.3:** Amino Acid Coupling-Step

It should be noted that an alternate coupling procedure was performed on the last three pentapeptides that were synthesized. After reading the publication by Hood et al. new reagents

were used on this step and the time in which the mixture was agitated was greatly reduced. (24) Instead of using PyBop, HOBt, and DIPEA with the amino acid, a mixture of HCTU and DIPEA were added to the synthesis vessel with the amino acid. The amino acid and HCTU were added in excess of 6 eq. and DIPEA of 12 eq. While this is a much larger amount of coupling reagents than the other procedure, the time of agitation is reduced to only 10 minutes and gives just as much visible product as the former method. It would be wise to consider the time efficiency of this second method, but it is important to note that if time is not an issue, and especially if only dipeptides are needed, that the other, more economical procedure might be a better choice for coupling.

Once the coupling step has finished, another wash with DMF and DCM then ninhydrin test is performed. This time, since the newly added Fmoc-protected amino acid should have reacted with all the free amines from the beaded resin, the solution should remain yellow when heated. If a blue color does form, then a wash with DMF followed by a repeated addition of coupling solution is necessary [22].

This procedure of deprotection followed by coupling is repeated for each amino acid. The only difference is that if the synthesis has not been stopped, there is no need to swell the beads for 30 minutes before each new deprotection. However, if the synthesis procedure has been delayed, the beads are swelled again before synthesis resumes. After the second alanine is attached, the procedure is repeated to couple Fmoc-Lysine-OH, Fmoc-Alanine-OH, and a final Fmoc-Alanine-OH. This gives a peptide chain attached to the bead of Fmoc-Ala-Ala-Lys-Ala-Ala-Wang resin.

After all wanted amino acids have been added to the chain, the next step is to cleave the peptide chain from the beaded resin. The first step in this process is to swell the resin again for

another 30 minutes. The beads are then washed with DMF and the last Fmoc is deprotected in the same manner as all of the other deprotection steps. A last wash and ninhydrin test are performed and the beads are dried out again with DCM. This is important since there should be no traces of residual DMF during the cleaving procedure.

To cleave the peptide from the beads, a solution of 95% trifluoroacetic acid, 2.5% deionized water, and 2.5% triisopropyl silane is added to the vessel. In this mixture, the TFA reacts with both the Wang resin linker and the side chain protecting groups on lysine and the other protecting groups for other di- and pentapeptides (ornithine, Daba, Dapa, arginine, and serine). The purpose for triisopropyl silane and water is to act as scavengers, whose role is to trap the very reactive cations produced from the removal of the resin and protecting groups. If left unquenched, these species could easily react with the product peptides. Therefore, these scavengers are added to reduce the chance of modifications on the desired peptide chain.

Due to the volatile nature of the cleaving solution, after about 10 mL of the solution is put into the vessel, parafilm is placed over the top of the vessel with holes poked on the top in order for ventilation. The argon flow is adjusted to a low rate so as not to boil of the solution. Another important reason to keep the flow low if there was a lot of bubbling, the beads could stick to the glass above the cleavage solution level. This would keep any uncleaved beads from undergoing the cleaving process and result in a lower yield. The beads are treated with this solution for 90 minutes and the solution is then pumped down the fritted disc into a clean 100 mL round bottom flask, which replaces the waste container below the drip joint. Immediately after collecting the cleaving solution into the flask, 30 mL of cold ethyl ether are added to the flask and the mixture is put in the freezer. Usually, the peptide is visible in the mixture as soon as the ether is added, but the solution is still kept in the freezer overnight to let as much product crash out as possible.

Once the pentapeptide in the ether has had enough time to crash out, the mixture is poured into test tubes and is centrifuged for about one minute. The ether is decanted out and the film of peptide left in the test tube is dissolved in a 25 mL mixture of 49.5 % water, 49.5% methanol, and 1% acetic acid. This 25 mL solution is placed in a capped vial and stored in a freezer to be analyzed under tandem mass spectrometry.

This same process was used for all pentapeptides and dipeptides. Other pentapeptides synthesized included AAOAA, AADabaAA, AADapaAA, ALAAA, AOAAA, ADabaAAA, ADapaAAA, AAAAL, AAAAO, AAAADaba, and AAAADapa. Dipeptides synthesized include AlaOrn, AlaLys, SerOrn, SerLys, ArgOrn, ArgLys, ProLys, ProOrn, OrnArg, LysArg, OrnAla, LysAla, and OrnLys. It should be noted that AADapaAA, ADabaAAA, and ADapaAAA were all synthesized with the HCTU/DIPEA coupling reagents as opposed to the PyBop, HOBT, DIPEA method of all the other peptides.

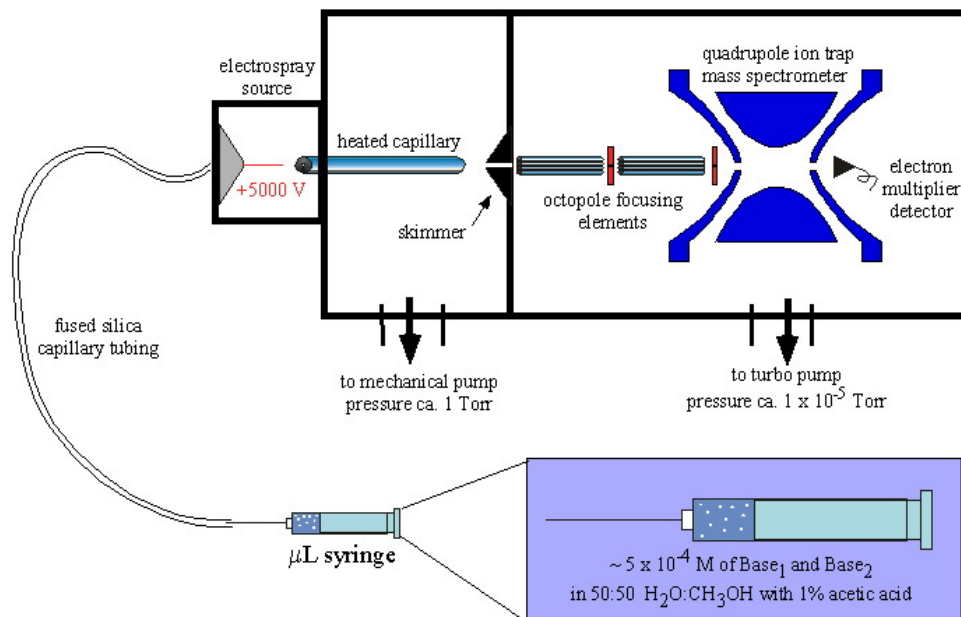
When first synthesizing ornithine containing peptides, an attempt was made to make Fmoc-Orn-Wang Resin by combining Fmoc-Orn-OH with plain hydroxyl based Wang resin. This technique, known as the DCB method, was based on the procedure outlined by Chan and White [22]. In this reaction, 2,6-dichlorobenzoyl chloride (DCB) is placed in the synthesis vessel with the plain OH-Wang resin and the protected amino acid along with DMF and pyridine. The reaction results in the formation of an anhydride between the DCB and Fmoc-amino acid. This anhydride goes on to promote esterification of the hydroxyl on the Wang resin with the C terminus of the amino acid, attaching the amino acid to the bead. In this procedure, the resin is added to the vessel and washed with DMF and excess DMF is added to form a solution of the resin and DMF. Secondly, 5 eq. of the amino acid (Fmoc-Orn-OH) and 8.25 eq. pyridine are added to the DMF mixture in the vessel and this is agitated until the solid amino

acid dissolves. DCB is then added and the mixture is bubbled for about 18 hours. After the 18 hours, the beads are washed and in the attempts made, the synthesis was continued with a coupling of the bead to alanine so as to form the dipeptide AlaOrn continuing the synthesis with the same steps as described above. This procedure was attempted about three times, and each time no AlaOrn was found after being analyzed with mass spectrometry. It was therefore necessary to order specially made Fmoc Ornithine Wang resin from ChemPep. Given the length of the synthesis, it is probably for the better that the product was ordered from an outside source rather than synthesized in lab anyway.

## **2.2 ESI Ion Trap Mass Spectrometry**

After synthesizing, cleaving, and storing a desired peptide chain in a vial of the H<sub>2</sub>O:MeOH solvent solution, the solution was eventually analyzed in a Finnigan LCQ-DECA ion trap mass spectrometer. This type mass spectrometer contains an ESI source that uses a voltage to create charged droplets of solvent with the analyte in the middle. The droplets go through a heated capillary which serves to boil off all the solvent surrounding the molecules so that protonated molecules of interest in the gas phase travel through the rest of the instrument. Immediately after the heated capillary is a set of focusing octopoles that can optimize the flow of ions going through them and allow the intensities of specific ions to be optimized. Following the focusing elements is an ion trap mass spectrometer with the ability of isolating specific mass to charge ratios and performing collision-induced dissociation (CID). In CID, the ions flowing through the instrument are accelerated into a neutral gas, in this case helium, and this produces collisional energy that breaks the ions into several fragments [4]. These fragments are scanned

and the fragmentation spectra of molecules under study can then be seen. A simplified diagram of this instrument can be seen below in Figure 2.4:



**Figure 2.4:** Simplified Diagram of ESI Ion Trap Mass Spectrometry

To introduce the peptides into the instrument a clean 500  $\mu\text{L}$  Gastight glass syringe is filled with the MeOH:H<sub>2</sub>O containing the peptide to the 500  $\mu\text{L}$  mark. It is then placed on kdScientific flow rate controller that is set at 900  $\mu\text{L}/\text{hour}$  and set to run until 500  $\mu\text{L}$  of solution have been dispensed. The peptide solution reaches the capillary needle which is operated at +4500 V and is electrostatically directed into the heated capillary, which is typically set at 125 °C with a capillary voltage of 45 V. The sheath gas (nitrogen gas), which forces the now ionized peptides through the rest of the instrument into the trap, is set to a flow rate of 20 arb. The typically starting settings for the two octopoles used as focusing elements are -3.5 V for multipole 1 offset, -28.5 V for multipole 2 offset, -92.00 V for lens voltage, 400 V for multipole RF amplitude, and -74.00 V for entrance lens. These settings can be changed when tuning to a



particular  $m/z$  is necessary. If this is the case, the settings for each will be tuned to an optimal position to find a specific peak and in this project, this is usually always done for the  $m/z$  of the protonated peptide chain of interest.

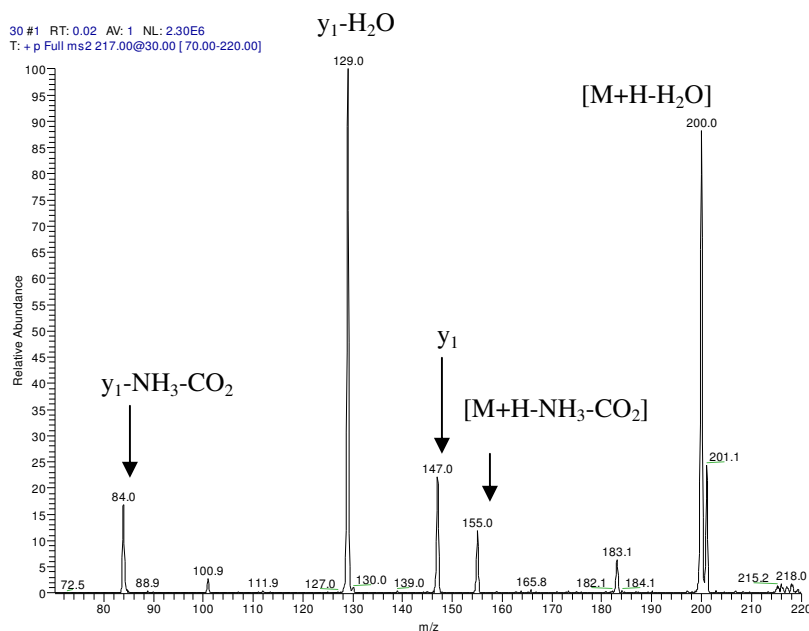
In the case of  $[\text{AlaAlaLysAlaAla}+\text{H}]^+$ , the peak expected is 431  $m/z$ . If this peak is seen it is isolated in the ion trap beginning with an isolation width of 10 au. After tuning, the signal is adjusted by reducing the isolation width until there are no other peaks observed in the spectrum. It is important to maintain an ion count of at least  $1.00 \text{ E } +5$ . However, the ion count for the parent ion is usually between  $1.00 \text{ E } +6$  and  $1.00 \text{ E } +7$ . Once the peak of interesting has been completely isolated and is at a reasonable intensity, the ion is activated at 10, 20, 30, 40, and 50 % collision energy using CID. A picture of each spectra at each energy level is recorded to be studied. The activation Q for each peptide fragmentation is set at 0.250 and the activation time is set to 30 msec.

## Chapter 3: Results Section

### 3.1 Lysine and Ornithine-containing Dipeptide Fragmentations

In this and the following section, the spectra of fragmentation will be presented and certain points of interest will be briefly mentioned. A more detailed discussion of these issues will be presented in the following chapter.

$MS^2$  fragmentation of Ala-Lys, with a mass-to-charge ratio of 218 m/z, was fragmented at 10, 20, 30, 40, and 50 % collision energy. The following spectrum shows the spectrum of fragments at 30%.

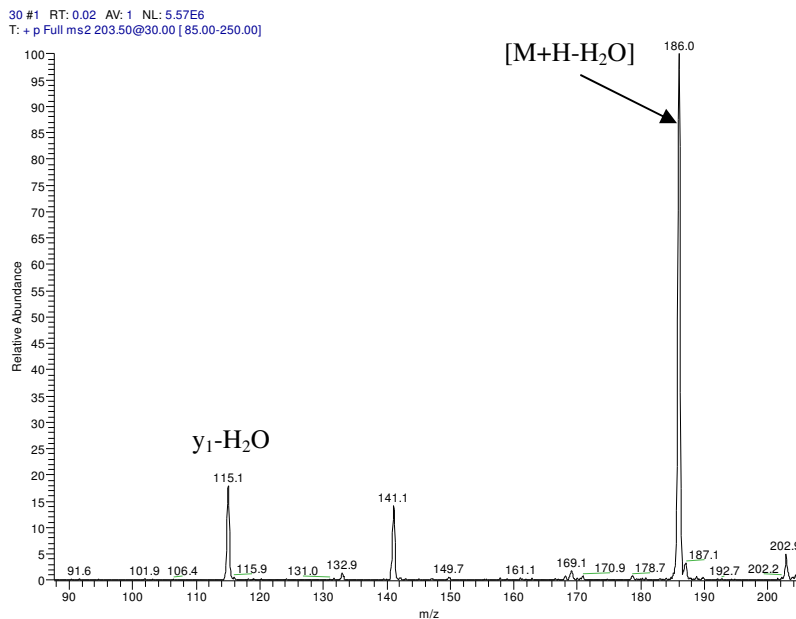


**Figure 3.1:** CID Spectrum of Alanine-Lysine

The main peak in the spectrum was  $y_1$ -water, while the other main peaks include loss of water from the parent ion. Being only a dipeptide, the amount of fragmentation is small and mostly includes these losses of water and the COOH group of mass 45 m/z. Other peaks present,

besides the ones identified, include a very small signal at 72.5, 100, 183, and 201 m/z. From this abundance of  $y_1$  based fragments, it appears that the more basic lysine favors keeping the charge.

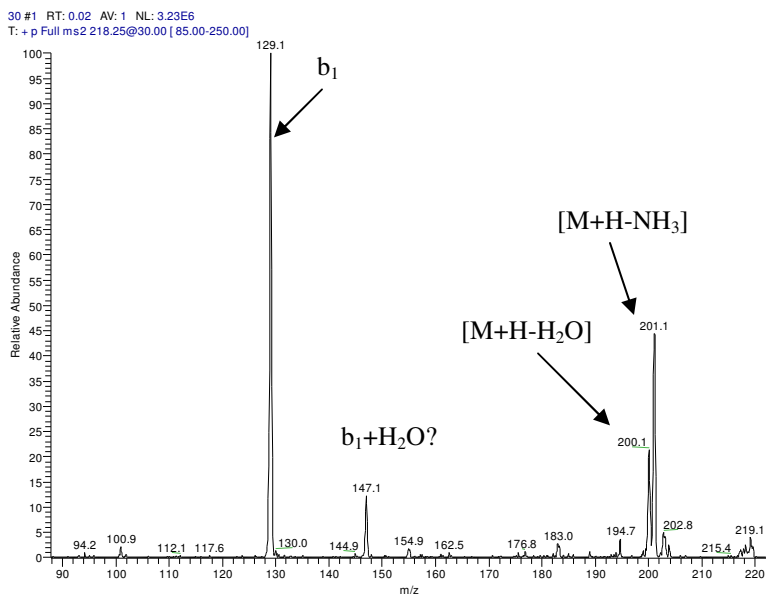
Ala-Orn was synthesized and isolated in MS<sup>2</sup> at 204 m/z, and its fragmentation spectrum at 30 % collision energy is shown below in Figure 3.2:



**Figure 3.2:** CID Spectrum of Alanine-Ornithine

Only three significant peaks are present in this spectrum. This includes the parent ion with the loss of water, which is the most abundant, as well as the  $y_1$  fragment minus water, and a peak at 141 m/z, which could possibly be the loss of ammonia and a CO<sub>2</sub> fragment. The abundance of these few fragments is a significant difference from the amount of peaks present in the Ala-Lys.

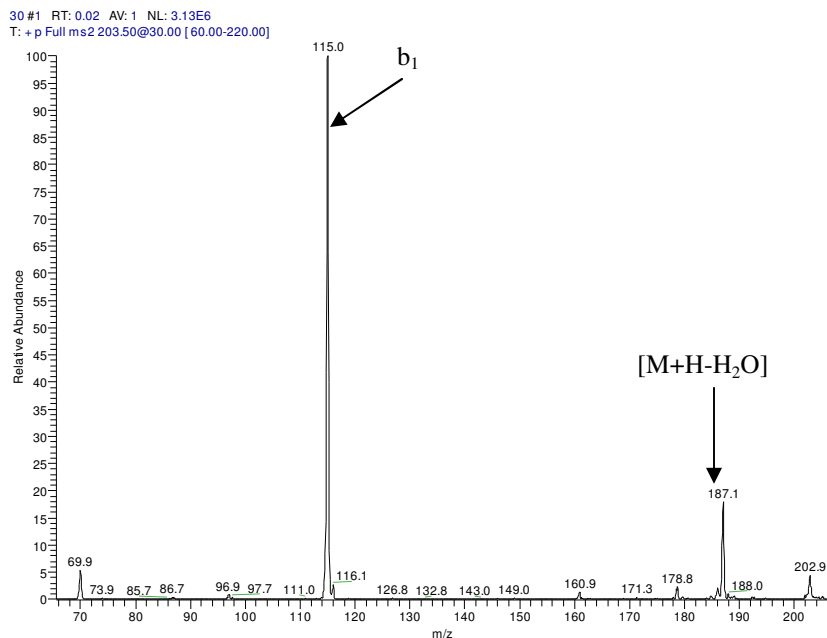
Having examined the fragments of the Ala-Lys and Ala-Orn dipeptides, it seemed reasonable to switch the position of the amino acid groups. Below, on Figure 3.3, is the fragmentation at 30 % collision energy of Lys-Ala, which has a mass of 218 m/z:



**Figure 3.3:** CID Spectrum of Lysine-Alanine

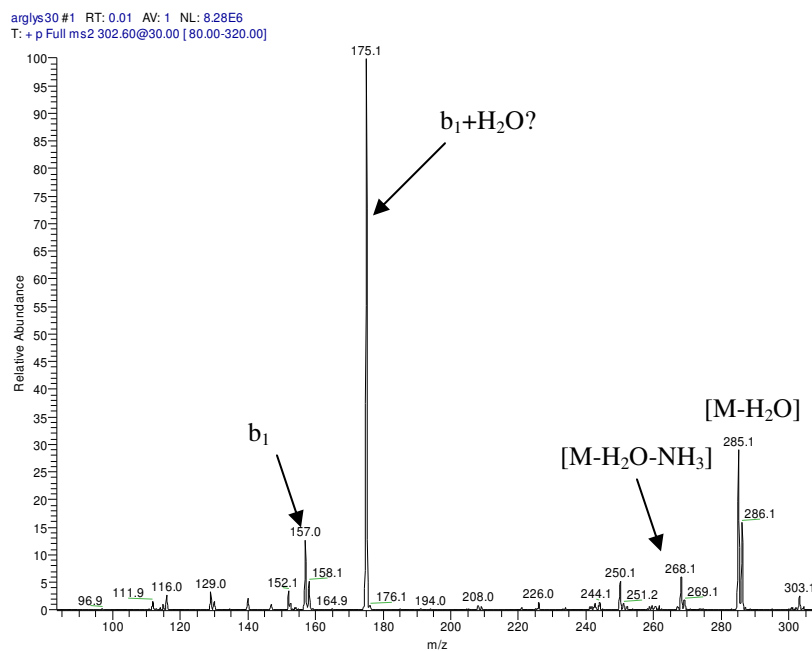
The most abundant fragment here is the b<sub>1</sub> ion, while there is also the parent with loss of water and also with loss of the amine group. A fairly large peak at mass-to-charge 147, which is the mass of protonated lysine on its own, is also present. This is a strange peak since b fragments never give the amino acid in its original mass.

The 30 % CE spectrum of Orn-Ala, mass-to-charge 204, presents a much simpler fragmentation pattern. The only fragments of interest include the loss of water and the  $b_1$ . However, no fragment at 133 m/z, which is the mass of plain ornithine as was found in Lys-Ala with lysine, was found. Like the Ala-Xxx comparison of the two, the Orn dipeptide showed a much simpler fragmentation. This spectrum is shown below on Figure 3.4



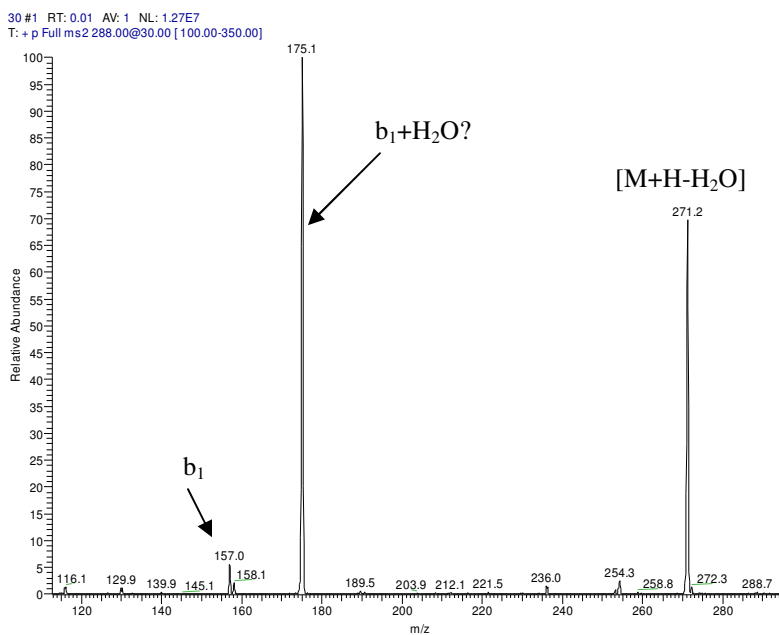
**Figure 3.4:** Spectrum of Ornithine-Alanine

The dipeptide Arg-Lys was synthesized and its mass at 303 m/z was isolated and fragmented with the same 30 % energy. The fragmentation spectrum gave both the parent ion with loss of water, ammonia, and with both. Another important fragment is the 175 m/z peak which is very similar to the 147 m/z fragment in the Lys-Ala fragmentation. It appears to be a  $b_1$  fragment with water added, but this is a strange fragmentation, since b fragments are usually cyclic and don't give the full amino acid. Another interesting thing to note is that there is an absence of a y fragment, showing that the arginine side chain is far more basic and forces the proton to stay with the arginine fragment. This fragmentation spectrum is shown below on Figure 3.5:



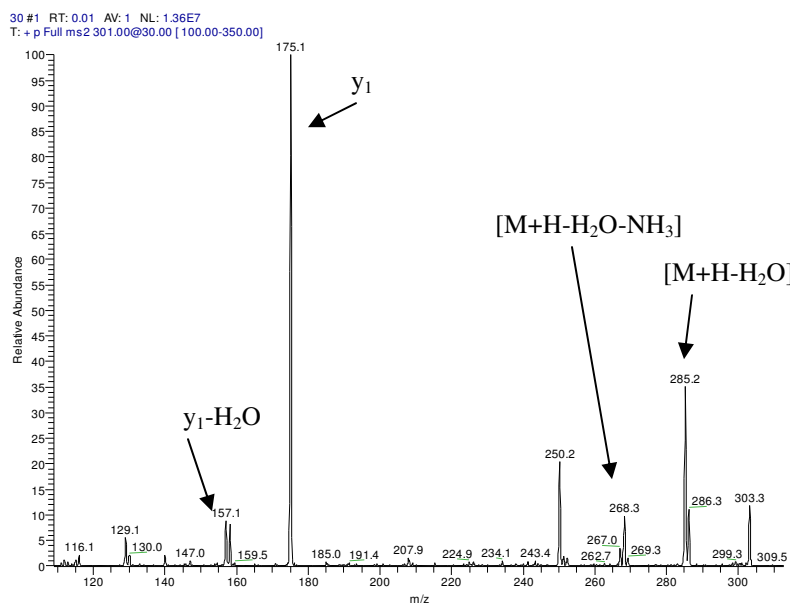
**Figure 3.5:** CID Spectrum of Arginine-Lysine

The Arg-Orn, with mass of 289 m/z, as a 30 % CE spectrum which again shows a characteristic lack of fragments. There are only three significant fragments, the most abundant of which is the  $b_1 + H_2O$ , which has been present in the Lys-Ala and Arg-Lys fragmentation spectra as well. However, there are not a significantly larger amount of peaks on Arg-Lys, most likely on account of the dominance of the 175 m/z peak since arginine is so basic an amino acid. This spectrum is shown below on Figure 3.6:



**Figure 3.6:** CID Spectrum of Arginine-Ornithine

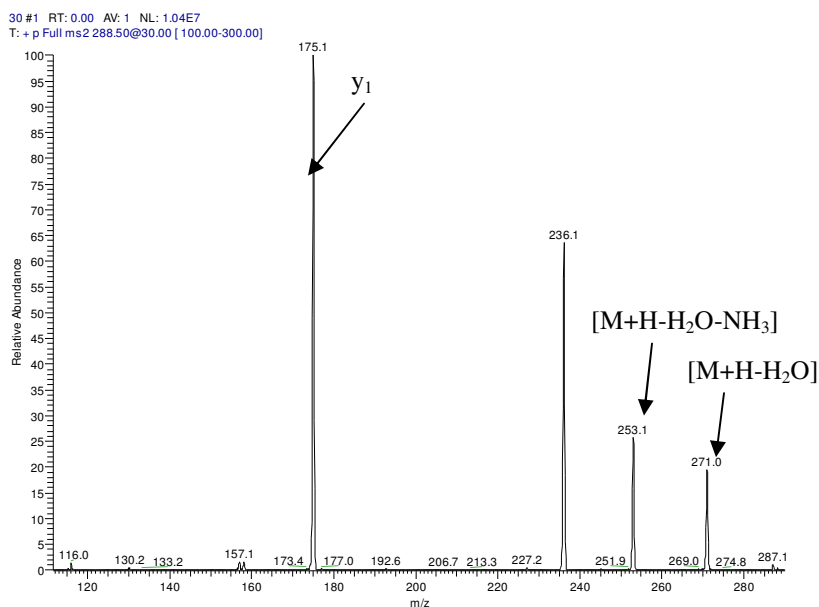
After synthesizing and fragmenting Lys-Arg, with mass of 303 m/z, the spectrum was recorded and the fragment ions at 30 % collision energy are presented. It is interesting to note that a very similar spectrum to Arg-Lys was produced. The fragment ion  $y_1$ , which is actually the same mass as  $b_1-H_2O$ , is the most dominant peak yet again. Again, the basicity of arginine dominates the fragmentation mechanism. Moreover, the parent with loss of both water and ammonia are obvious products, as well as the peak at 250 m/z. The only real difference in these two spectra is the abundances of the fragment ions, with 250 m/z being larger here and the fragment at 129 m/z also more significant. This spectrum is shown below on Figure 3.7:



**Figure 3.7:** CID Spectrum of Lysine-Arginine

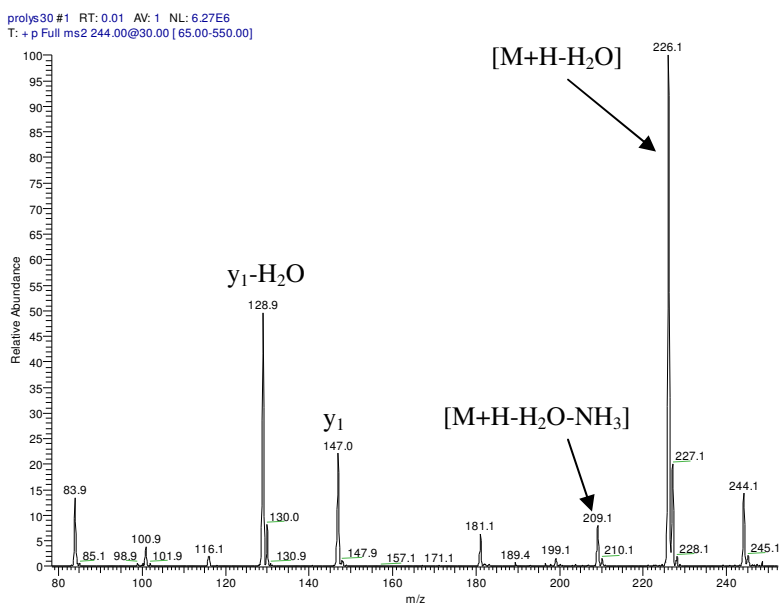


The Orn-Arg fragmentation spectra, with a parent mass of 289 m/z, was taken at 30 % CE as with previous dipeptides. It contains a simple spectrum with the abundant  $y_1$  peak as with Lys-Arg. The most interesting difference between this spectrum and that of Lys-Arg is the higher abundance of the 236 m/z as compared to the 250 m/z, which are the same fragment when taking the difference between lysine and ornithine into consideration. Also to note is that the  $y_1$ -H<sub>2</sub>O is barely visible in this spectrum. Furthermore, it shows spectrum with much fewer peaks than the previous lysine-containing dipeptide.



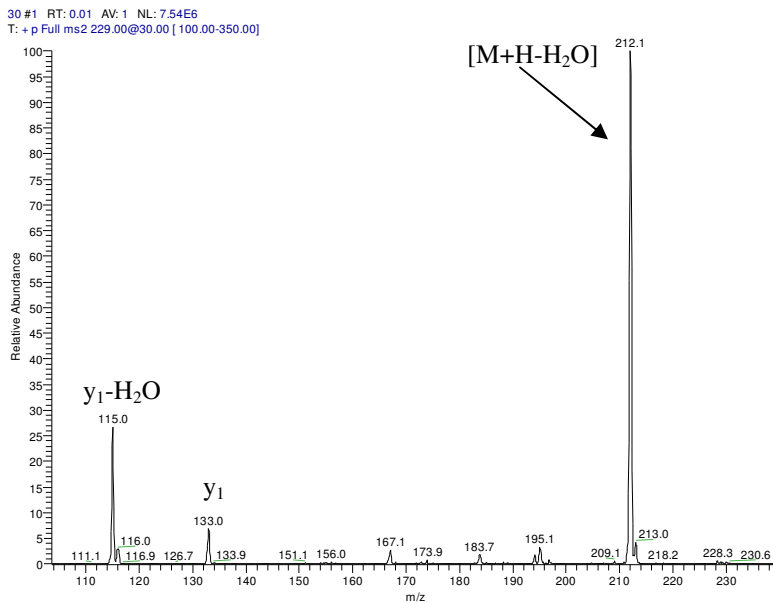
**Figure 3.8:** CID Spectrum of Ornithine-Arginine

The Pro-Lys dipeptide, with a protonated mass of 244 m/z, was synthesized and fragmented as per the other dipeptides studied. As was expected due to its basicity, the lysine on the C-terminal position produced y ions without any significant b ions. The most abundant ion was that of the parent with loss of water, which is a common peak in all of these dipeptides, with Ala-Orn as the other with a very high dominance. Overall it was a messy fragmentation with several small fragments at 181, 116, 101, and 70 m/z among the more abundant  $y_1$ ,  $y_1-H_2O$ , and the parent with loss of water and ammonia. These peaks can be seen below on Figure 3.9:



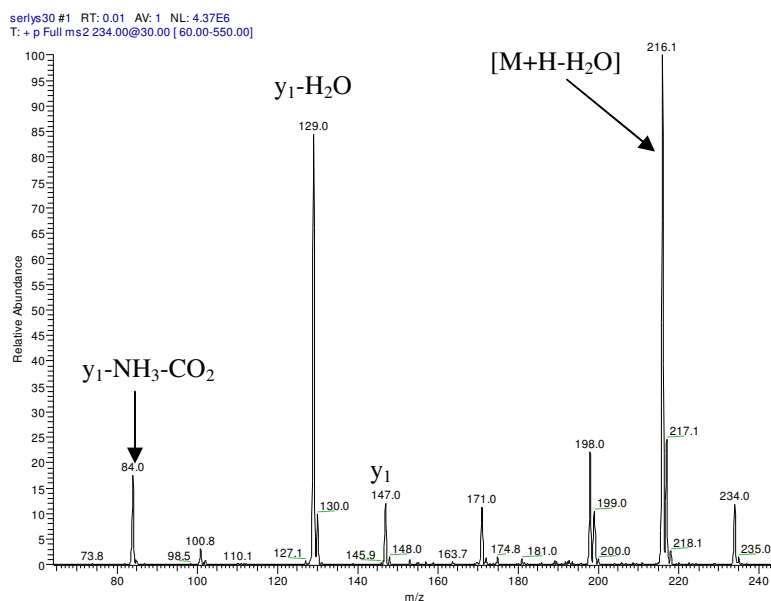
**Figure 3.9:** CID Spectrum of Proline-Lysine

Fragmenting Pro-Orn, with mass 230 m/z, gave yet another characteristically barren fragmentation spectrum. The only peaks of real interest are the most dominant fragment of the parent minus water, the  $y_1$ -H<sub>2</sub>O fragment, and  $y_1$  ion. Other fragments include 195, 184, and 167 m/z but with barely any abundance. It appears a clear pattern is developing in the spectrum of ornithine-containing dipeptides when comparing them to the lysine-containing peptides.



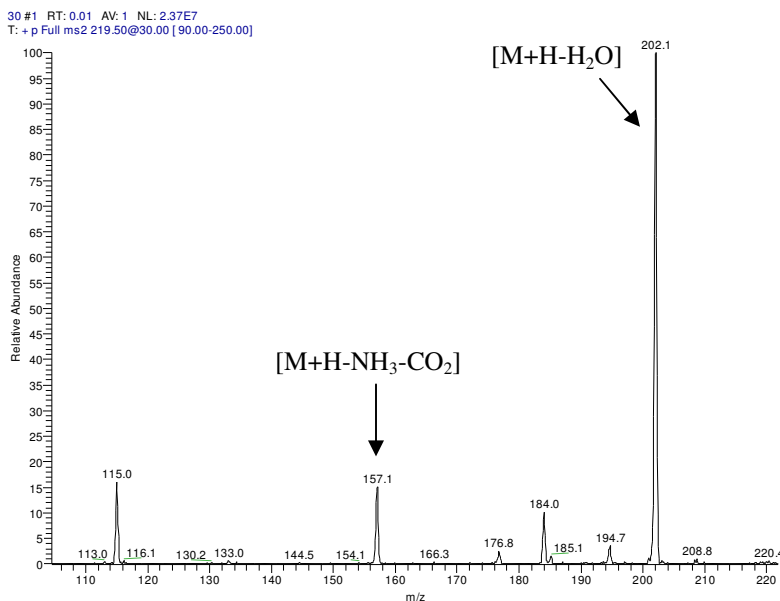
**Figure 3.10:** CID Spectrum of Proline-Ornithine

The final dipeptides combination synthesized and fragmented was the Ser-Xxx dipeptides. In Ser-Lys, with mass of 234 m/z, shown below on Figure 3.11, there is a tendency for y ions as opposed to b ions, due again to the basicity of that amine group on lysine's side chain. Like the Pro-Lys fragment spectrum, the most dominant peak for the parent with loss of water, followed by the  $y_1$  and  $y_1$ -H<sub>2</sub>O peaks also being present. Other peaks to note are the  $y_1$ -NH<sub>3</sub>-CO<sub>2</sub> at 84 m/z, the parent with loss of water and ammonia at 198 m/z, and the smaller fragments of 171 and 101 m/z.



**Figure 3.11:** CID Spectrum of Serine-Lysine

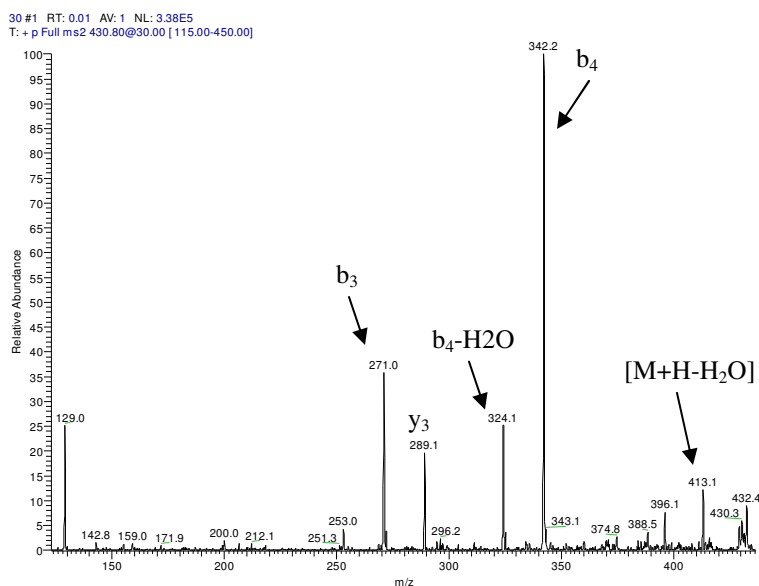
In the Ser-Orn fragmentation spectrum, with a mass of 220 m/z, shown below on Figure 3.12, there is a clear dominance of the parent with loss of water. This spectrum is particularly peculiar in that there is no sign of any y or b ions. There seem to be only fragments of functional groups and the backbone of the peptide seems to be protected from fragmentation somehow. Other than 202, 195, 184, 177, 157, and 115 m/z, there is relatively little background noise. It seems to be a clean fragmentation and with the absence of any usual fragment ions. This leads to the assumption that there appears to be some interaction between the serine and ornithine side chains.



**Figure 3.12:** CID Spectrum of Serine-Ornithine

### 3.2 Lysine, Ornithine, Daba, and Dapa-containing Pentapeptide Fragmentations

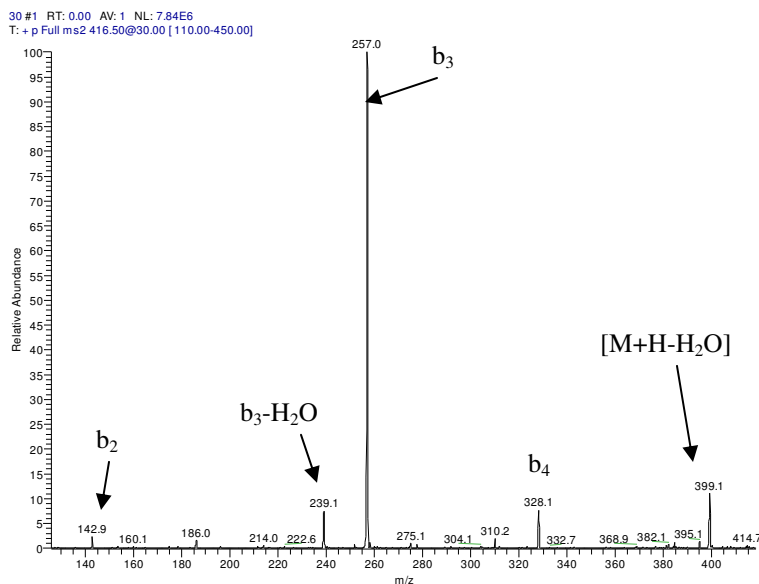
With five amino acids in the chain, a great deal more fragments can be observed, and a better distinction between fragmentations of lysine and its homologs should be observed. MS<sup>2</sup> fragmentation of AlaAlaLysAlaAla, with a mass at 431 m/z, was fragmented with 10, 20, 30, 40, and 50 % collision energy in the same fashion as all the dipeptides studied. The fragmentation at 30 % is shown below:



**Figure 3.13:** CID Spectrum of Ala-Ala-Lys-Ala-Ala

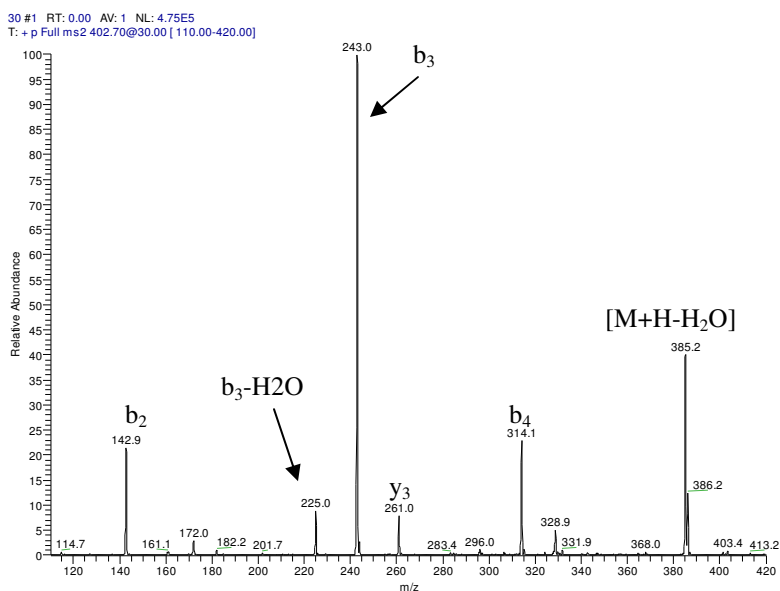
The most abundant peak in this spectrum was the b<sub>4</sub> ion, followed by the b<sub>3</sub>, b<sub>4</sub>-H<sub>2</sub>O, a peak at 129 m/z, which has shown up in several of the fragmentation spectra of the dipeptides studied and has not been identified in any real certainty, and also y<sub>3</sub>. Other smaller peaks include the parent with loss of water, 396, and 253 m/z. Several other insignificant peaks are present that convey a rather messy fragmentation.

The Ala-Ala-Orn-Ala-Ala fragmentation, with a mass of 417 m/z, provides a spectrum with the same characteristic traits of the dipeptides containing ornithine. In this spectrum there is only one very abundant peak, which is  $b_3$ , and only a few other observable peaks. This includes  $b_4$ ,  $b_3$ -H<sub>2</sub>O,  $b_2$ , and the parent minus water, with the lack of any significant y ions. The fact that the most abundant ion is  $b_3$  is very interesting when compared to the Ala-Ala-Lys-Ala-Ala where the most abundant peak was the  $b_4$  ion. These peaks can be observed on Figure 3.14 below:



**Figure 3.14:** CID Spectrum of Ala-Ala-Orn-Ala-Ala

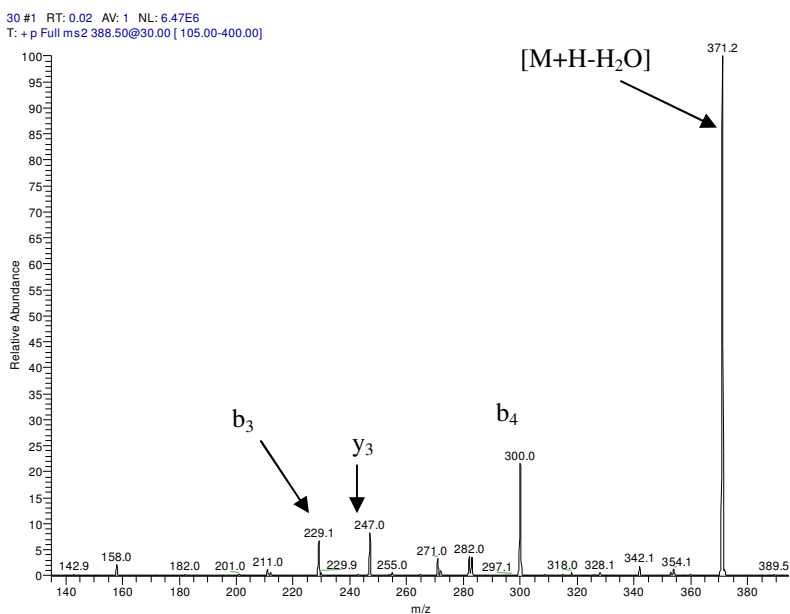
Synthesis and fragmentation at 30 % CE of Ala-Ala-Daba-Ala-Ala, with a mass of 403 m/z, was performed and produced the spectrum shown below on Figure 3.15. The most abundant fragment was  $b_3$ , with the parent with loss of water, followed by  $b_4$  and  $b_2$  as the next most abundant. The  $y_3$  ion was also present as were 329 m/z,  $b_3$ -H<sub>2</sub>O, and 172 m/z. It seems to be quite messier than Ala-Ala-Orn-Ala-Ala, yet it was much cleaner than the Ala-Ala-Lys-Ala-Ala pentapeptide. Like the ornithine-containing peptide, its primary fragment ion was the  $b_3$ , yet there was significantly more  $b_4$  in this spectrum.



**Figure 3.15:** CID Spectrum of Ala-Ala-Daba-Ala-Ala

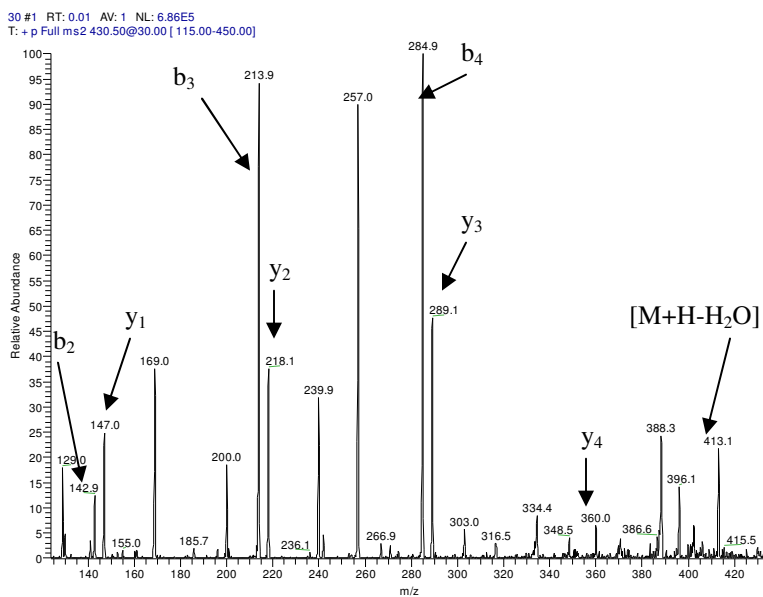


The fragmentation of Ala-Ala-Dapa-Ala-Ala, with a mass of 389 m/z, was done at 30 % CE as with the other pentapeptides. It gave a spectrum with few ions that was much cleaner than Ala-Ala-Lys-Ala-Ala and Ala-Ala-Daba-Ala-Ala. It was slightly cleaner than ornithine, but gave a much smaller ratio of b and y ions compared to the most abundant peak, [M+H-H<sub>2</sub>O]. The b<sub>4</sub> ion was the highest of these fragments, with b<sub>3</sub> and y<sub>3</sub> ions present as well. Other than these fragments, no peaks of interest were present. This spectrum can be seen below on Figure 3.16:



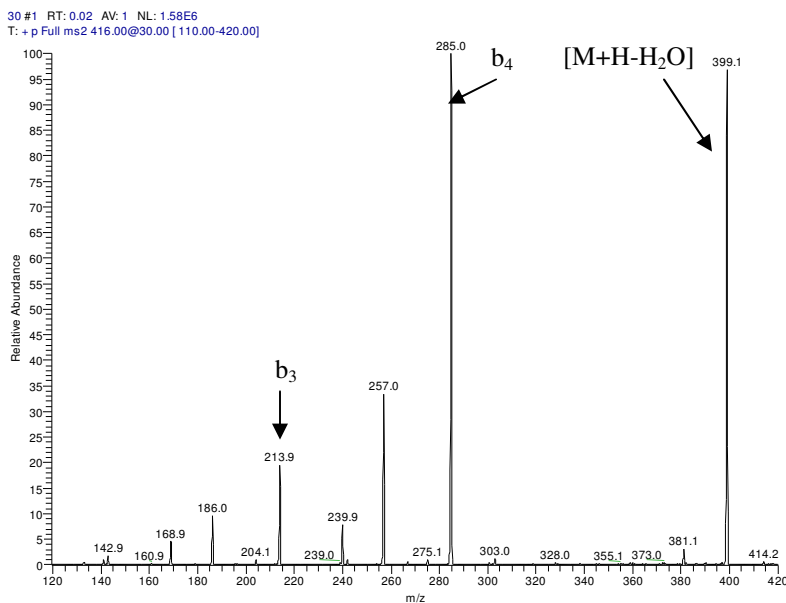
**Figure 3.16:** CID Spectrum of Ala-Ala-Dapa-Ala-Ala

The Ala-Ala-Ala-Ala-Lys peptide, with mass of 431 m/z, gave a fragmentation spectrum with the most abundant amount of fragments of any pentapeptide synthesized. The  $b_2$ ,  $y_1$ ,  $b_3$ ,  $y_2$ ,  $b_4$ ,  $y_3$ , and  $y_4$  are present as are the parent with loss of water, 396, 388, 334, 303, 257, 240, 200, 186, 169, and 129 m/z. It appears that at the end position, lysine may not direct a primary mechanism which would produce a dominant fragment. Furthermore, the fact that more abundant ions are b is interesting since the C-terminal lysine is more basic than the alanines in the chain, which should allow for more y fragments where the proton would end up on the amine side chain of lysine.



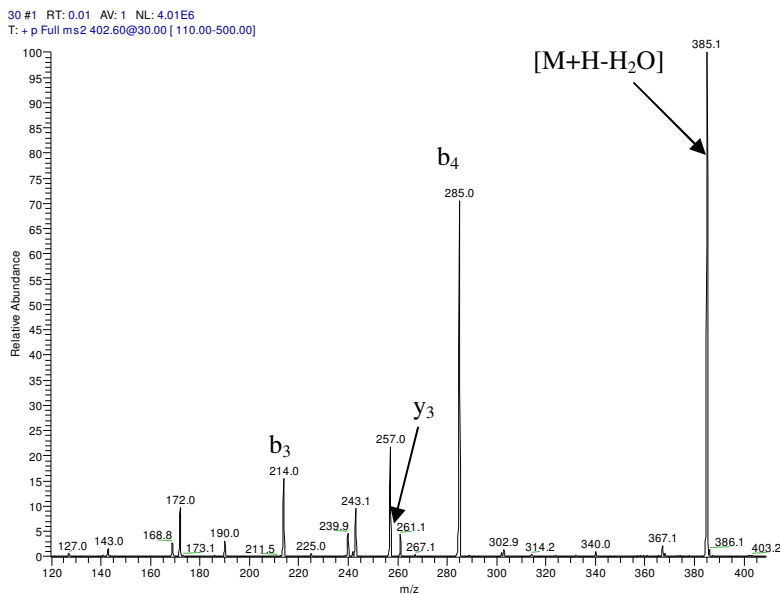
**Figure 3.17:** CID Spectrum of Ala-Ala-Ala-Ala-Lys

The Ala-Ala-Ala-Ala-Orn pentapeptide, with a mass of 417 m/z, was fragmented and recorded at 30 % CE is seen below on Figure 3.18. A much cleaner spectrum is observed in this fragmentation compared to Ala-Ala-Ala-Ala-Lys. The  $b_4$  ion is still the most abundant, but the  $b_3$  is the only other significant typical fragment seen upon CID. The other fragments include the parent with loss of water, which has a very high abundance, 257, 240, 186, 169, and 143 m/z. A few of these peaks, 257, 240, and 186 were seen in the previous pentapeptide, but other than these there are no traces of the several other peaks present in the lysine-containing. Perhaps the length of the lysine side chain does indeed have some effect on directing several different mechanistic fragmentations, instead of the suggestion that its C-terminal position keeps it from swaying the fragmentation in any way.



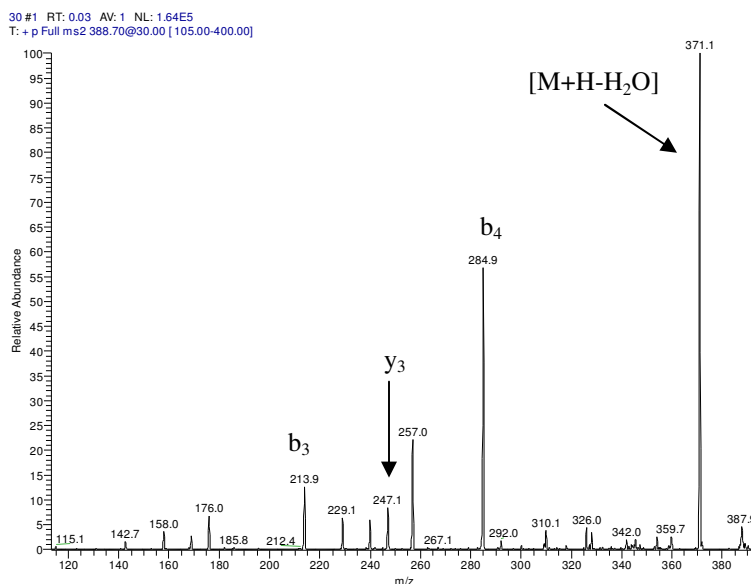
**Figure 3.18:** CID Spectrum of Ala-Ala-Ala-Ala-Orn

The Ala-Ala-Ala-Ala-Daba peptide, with a mass of 403 m/z, was fragmented at 30 % CE was performed and gave a spectrum with few differences to the Ala-Ala-Ala-Ala-Orn pentapeptide. The only real difference was the domination of the parent with loss of water, yet the  $b_4$  ion is still in high abundance with the  $b_3$  at a reasonably significant level, with the  $y_3$  ion at 261 m/z at a very small abundance. Other peaks include the 257 m/z fragment, and 243, 240, 190, 172, and 143 m/z. This spectrum wasn't as clean as the Ornithine containing pentapeptide, but it did have a very similar fragmentation pattern. This spectrum can be seen below on Figure 3.19:



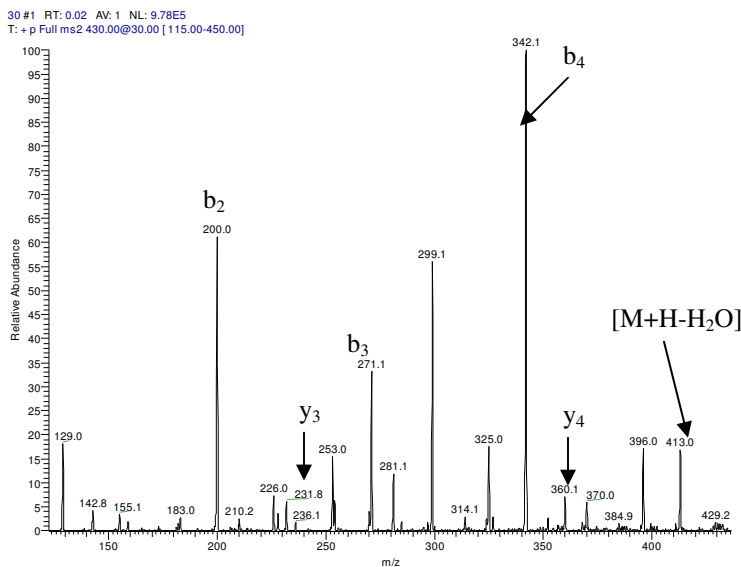
**Figure 3.19:** CID Spectrum of Ala-Ala-Ala-Ala-Daba

The Ala-Ala-Ala-Ala-Dapa peptide, with a mass of 389 m/z, was fragmented and its spectrum closely resembles both the ornithine-containing and Daba-containing pentapeptides. The parent with loss of water is the most abundant, followed by the  $b_4$  ion, which is the same as in the Ala-Ala-Ala-Ala-Daba pentapeptide. There are however several more small peaks in this spectrum, most of which between 285 m/z ( $b_4$ ) and 371 m/z ( $[M+H-H_2O]$ ). Other peaks include 257, 247, 240, 229, 176, 158, and 143 m/z. This rather abundant fragmentation spectrum is somewhat out of trend with the other results seen in the other Dapa-containing peptides.



**Figure 3.20:** CID Spectrum of Ala-Ala-Ala-Ala-Dapa

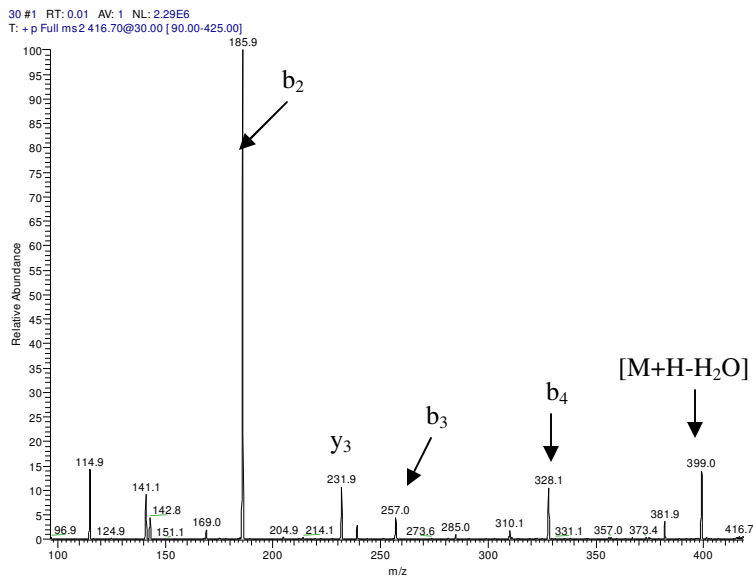
The last set of pentapeptides containing either lysine or its homologs are the alanine chain with the amino acid of interest in the second spot on the chain, AXAAA. Below is the fragmentation of Ala-Lys-Ala-Ala-Ala, with a mass of 431 m/z, at 30 % CE, seen on Figure 3.21:



**Figure 3.21:** CID Spectrum of Ala-Lys-Ala-Ala-Ala

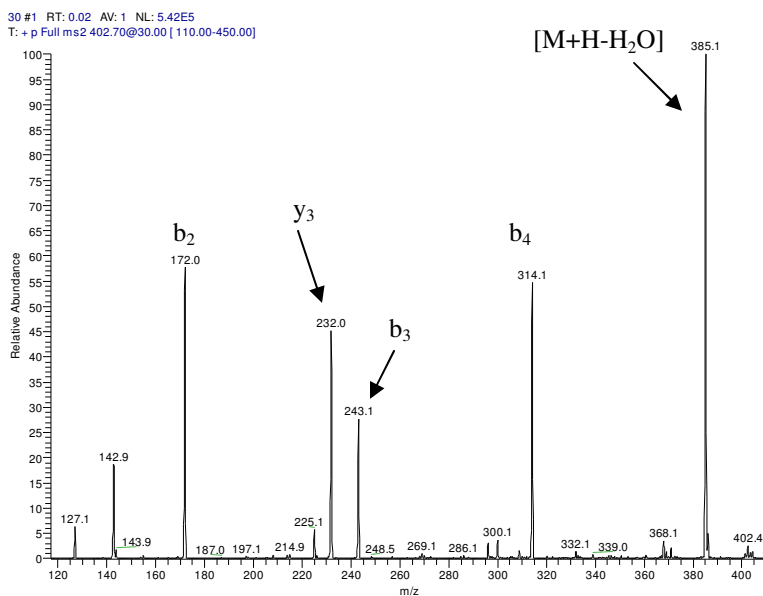
As with the two other lysine containing pentapeptides, there are a wide spectrum of fragment ions resulting from the CID fragmentation. This includes the most dominant b<sub>4</sub> ion, and also the b<sub>2</sub>, b<sub>3</sub>, y<sub>3</sub>, and y<sub>4</sub> fragment ions. Other peaks present include the parent with loss of water, 396, 370, 325, 299, 281, 253, 226, 183, 155, 143, and 129 m/z.

A very different spectrum resulted from the fragmentation of Ala-Orn-Ala-Ala-Ala, at mass 417 m/z, at 30 % CE. The most abundant ion was the  $b_2$  instead of the  $b_4$ , yet the  $b_4$ ,  $b_3$ , and  $y_3$  ions were all still present, but at lower abundances. Moreover, it was again a much cleaner spectrum with the other significant peaks at  $[M+H-H_2O]$ , 382, 141, and 115 m/z. The fact that the  $b_2$  was most dominant seems very peculiar in that the chain might not have been long enough to result in a large abundance of  $b_4$  which is two amino acid residues away from the Ornithine side chain. Furthermore, there may be a specific ring formation between the ornithine side chain and the first alanine that makes the  $b_2$  dominant, as could have been the case with the Ala-Ala-Orn-Ala-Ala where  $b_3$  is the dominant fragment ion while  $b_4$  was the dominant with Ala-Ala-Lys-Ala-Ala as well.



**Figure 3.22:** CID Spectrum of Ala-Orn-Ala-Ala-Ala

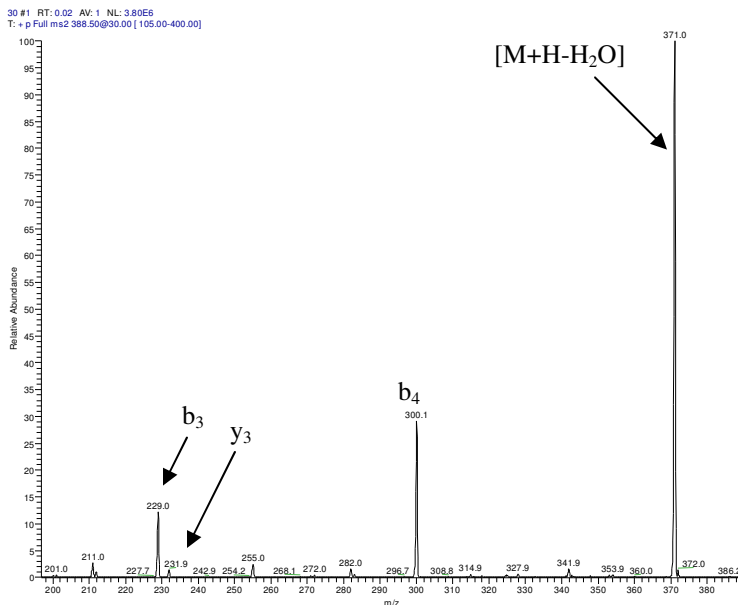
In the second position, the Daba group, with a mass of 403 m/z, fragments with a wide selection of different product ions. The spectrum below, on Figure 3.23, gives a dominant ion of [M+H-H<sub>2</sub>O] followed by the b<sub>4</sub> ion. The b<sub>2</sub> is next abundant, followed by the y<sub>3</sub> and b<sub>3</sub>. Other peaks with some significance include 225, 143, and 127 m/z without very much noise otherwise. As with the AAXAA cases, the Daba-containing pentapeptide gave more prominent signal ions than did the Ornithine-containing, yet it was quite cleaner a spectrum than the lysine-containing. It is interesting that the parent with loss of water was dominant, as that was the same case in Ala-Ala-Ala-Daba, but it should be noted that it was not the case in Ala-Ala-Daba-Ala-Ala.



**Figure 3.23:** CID Spectrum of Ala-Daba-Ala-Ala-Ala



The last pentapeptide synthesized, Ala-Dapa-Ala-Ala-Ala, with a mass of 389 m/z, was fragmented at 30 % collision energy as identical to the other peptides. It only had three significant peaks, in which the parent with loss of water was dominant, followed by the b<sub>4</sub> ion and the b<sub>3</sub> ion. Other than these three ions and a tiny y<sub>3</sub> peak at 232 m/z, there were small signals at 282, 255, and 211 m/z. Overall, this was a very clean fragmentation spectrum. Once again, as with the other Dapa-containing pentapeptides, the parent with loss of water was the dominant fragment ion. The ratio of b<sub>4</sub> to b<sub>3</sub> stayed pretty much the same as well, with y<sub>3</sub> much smaller in this second position. Clearly, by shortening the side chain to only one methylene, there is a limited amount of pathways that the peptide can fragment as compared to the lysine, ornithine, and Daba-containing peptides. Perhaps the short amine side chain favors the loss of water, whose signal is drowning out the other possible fragments in this spectrum.



**Figure 3.24:** CID Spectrum of Ala-Dapa-Ala-Ala-Ala

## Chapter 4: Discussion Section

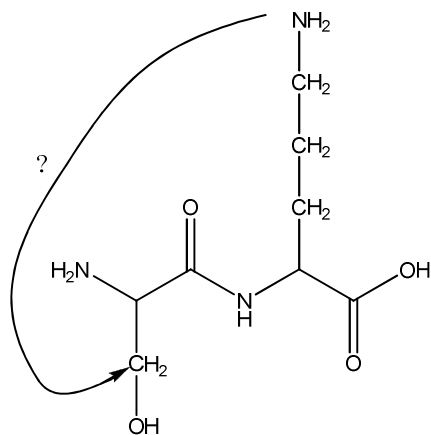
### 4.1 Important Features of Dipeptide Fragmentations

As was mentioned in the first section, ornithine and lysine have very similar proton affinities that are only about 3 kJ/mol apart from each other, not factoring in the experimental error. This means that the differences in their fragmentation spectra result directly from the lengths in their side chains and very little from the thermochemical characteristics that effect the movement of the mobile proton. It seemed most reasonable that dipeptide pairs be compared first in this analysis. In studying dipeptide molecules, there was a smaller chance of many varying fragmentations being present. With fewer competing fragmentation pathways, a smaller amount of factors would need to be considered in the differences between the dipeptide pairs.

The most obvious feature of all the dipeptide pairs synthesized was that the lysine fragmentation spectrum of each was much more abundant in peaks than the ornithine dipeptide. As this type of analysis is a novel approach in understanding the effect of lysine and its homologs on fragmentation, only a number of speculative ideas can be proposed for this behavior. As of yet, no actual mechanism has been proposed for lysine driven peptide fragmentation, but the peak differences do convey some important information.

The first explanation to consider in having such fewer fragment ions in the ornithine deals with the possibility of a very stable structure forming by interactions with the side chains, which could be making several fragmentation possibilities less abundant. The side chain of ornithine could be just the right length to form a ring with the adjacent amino acid with which it is coupled. The most convincing example of this possibility is within the Ser-Orn dipeptide. There appear to be no b or y ions in this fragmentation, but rather a loss of the water and

ammonia functional groups, with the loss of water as the most abundant fragment. Perhaps a stable cyclic ion is being formed between the amine of ornithine and the side chain of serine, where serine loses its OH and the amine loses a hydrogen molecule to cause the loss of water. For clarity, this dipeptide is modeled below on Figure 4.1:



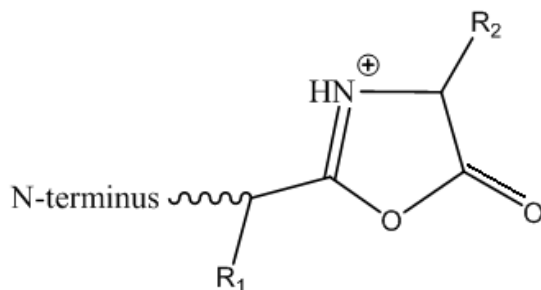
**Figure 4.1:** Serine-Ornithine Dipeptide

Lysine also has the loss of water as its most abundant peak, so it may be forming this structure as well, but it has several other peaks including the  $y_1$  ion, which leads to the idea that the lysine side chain is too long to make a very stable ring structure. The weakness of a structure would result in further fragmentations occurring. This same dominance of the parent with loss of water occurs in Ala-Orn, while the lysine-containing again has other very dominant fragment ions including  $y_1$  and  $y_1-H_2O$ . However, without a hydroxyl group on the alanine side chain, there seems to be a less likely reason for a cyclic structure to be formed from loss of water. However, a pattern does seem to be developing in these dipeptide fragmentations.

Another possibility in the apparent disparity between the amount of fragments in lysine and ornithine containing dipeptides is the advantage lysine gains in being able to move around the peptide structure. With one more methylene in length than the ornithine, it seems reasonable

that the side chain can reach places in the dipeptide to promote specific fragmentations that the ornithine can not. The important question to consider is whether ornithine is too small to reach all the important places to promote specific fragments in a dipeptide molecule. This question is better dealt with when discussing the pentapeptides fragmented in the next section. With these longer molecules, clearer boundaries can be seen in the fragmentation patterns with regard to side chain involvement, while more possible fragments will make it easier to observe the differences.

The last important feature of the dipeptide fragmentations involved somewhat anomalous peaks in the Arg-Lys, Arg-Orn, and Lys-Ala dipeptides. When b ions are formed, they typically become cyclic, and create what are known as oxazolone ions [26,27]. A picture of this structure can be seen below on Figure 4.2:



**Figure 4.2:** Oxazolone  $b_2$  ion

This has become the widely accepted model for b ions, yet this structure can not explain the formation of the  $b_1+H_2O$  and  $b_1$  ions observed with these three dipeptides. Hiserodt et al. have observed this same formation of  $b_1+H_2O$  and  $b_1$  ions in dipeptides that contain arginine, lysine, or histidine on the N-terminus. They maintain that the basicity of these groups allow them to take the charge after fragmentation and this group presents a number of possible mechanisms for these two ions occurring, not shown here [28]. The problem with these fragment ions occurring,

as was discussed by Hiserodt et al., was that these fragments are identical to the  $y_1$  and  $y_1-H_2O$  of the dipeptides containing the same two amino acids, but with the lysine, arginine, or histidine on the C-terminus. These similarities were observed in the spectra Ala-Lys, Lys-Arg, and Orn-Arg in this study as well.

## 4.2 Important Features of Pentapeptide Fragmentations

Whereas only ornithine and lysine were synthesized in the dipeptide fragmentation studies, having added Daba and Dapa, a comparison of shortening the side chain even further can be performed. The additional length of the pentapeptides allow for a more interesting variety of fragments, which should convey even more information in these early analyses of the effects of side chain length on fragmentation.

As with the dipeptide fragmentations, the ornithine-containing pentapeptides gave much cleaner spectra than the lysine-containing for all three positions, AXAAA, AAXAA, and AAAAX. Another important feature in the ornithine-lysine comparison included the dominant fragments seen in each case. As was mentioned in the previous chapter, for the AAXAA position, the dominant fragment for the lysine pentapeptide was the  $b_4$  ion, while for ornithine the most abundant was the  $b_3$  ion. In fact, very little  $b_4$  ion is seen in the ornithine-containing, with 8% relative abundance, while an abundance of 36% for the  $b_3$  ion is seen in the lysine-containing. What this may mean is that the length of the lysine side chain really does allow it to move further out and assist in the fragmentation pathways that occur under CID. With one less methylene, perhaps the ornithine side chain can not reach the bond between the 4<sup>th</sup> and 5<sup>th</sup> alanines well enough produce abundant  $b_4$  ions.

Another possibility is that the length of the ornithine side chain is just the right length to cause a jump in the amount of  $b_3$  ions produced, making the other fragmentations, which are still possible, insignificant. This makes much more sense than the previous idea, since both AADabaAA and AADapaAA both contain the  $b_4$  ion in their spectra. If ornithine was too small to produce an abundant amount of  $b_4$  ions, it would likely be the case that neither Daba nor Dapa would have  $b_4$  ions any more abundant. However, the Daba-containing had an even more abundant  $b_4$  ion peak (23%) than did the ornithine-containing (8%), which makes the idea of the ornithine side chain being too small very unlikely. This does not mean that the length of the lysine side chain does not give it a possible advantage in producing  $b_4$  ions, only that the shortness of the ornithine side chain is not the reason the  $b_4$  is less abundant in its fragmentation spectra.

It should also be noted that this same phenomenon occurs in the AXAAA pentapeptide comparisons. In the lysine containing, the most dominant fragment is still the  $b_4$  ion with the  $b_2$  at 60% relative abundance, while the ornithine-containing has a dominant peak at  $b_2$  with the  $b_4$  at only 11%. As with the AAXAA case, it appears to be a result of the lysine being able to reach further along the side chain to promote the peptide bond between the 4<sup>th</sup> and 5<sup>th</sup> alanines, yet that bond is quite far away from the 2<sup>nd</sup> position amino acid. Furthermore, the Daba and Dapa both have a reasonable signal at  $b_4$ . Another possibility, as was proposed in the AAXAA case, is that the ornithine somehow forces a mechanism which cleaves the peptide bond on its C-terminus when it is somewhere in the middle of the chain. Whereas when it was on the third position, the most dominant fragment was the  $b_3$  ion, here, since it is in the second position, the  $b_2$  ion is the most dominant. In future studies, the ornithine should be placed in larger peptides at various positions to see whether this trait persists.

Another important feature of the AAXAA comparisons was that the ornithine-containing pentapeptide was the only one of the four that did not have a significant abundance of the  $y_3$  ion. While it was present, there was barely a signal, with a relative abundance of only 1%. Compared to this, the AAKAA, AADabaAA, and AADapaAA peptides had relative abundances of 20%, 8%, and 8% respectively, for the  $y_3$  ion. This could mean that the length of the ornithine side chain somehow makes it difficult to break the 2<sup>nd</sup> position alanine to ornithine C-N peptide bond. Clearly, this is not the case for the lysine-containing and it is clearly not unlikely for the Daba and Dapa-containing. Something exceptional may be happening in the fragmentation of ornithine that keeps specific bonds from breaking.

As for when lysine or its homologs are at the end of the peptide, like the AAAAX pentapeptide, fragmentation spectra do differ from the other two positions in specific ways. This peptide chain is of particular interest in that it resembles what a trypsin digested peptide chain would look like, with lysine on the C-terminus. First of all, the AAAAK pentapeptide gave the most abundant amount of fragment ions, by far. For some reason, at this position many fragmentation pathways are competing, since the spectrum contains  $b_2$ ,  $b_3$ ,  $b_4$ ,  $y_1$ ,  $y_2$ ,  $y_3$ , and  $y_4$  ions among many other significant peaks. There are two possibilities for why the lysine at this position causes such a variety of peaks. One is that lysine is further away from the bulk of the peptide and that it is less of a factor in deciding fragmentation, which results in more even peptide cleaving along the backbone. The second possibility is that lysine is in an ideal position to interact with all positions of the chain, producing the variety of peaks.

One feature to consider in these speculations is the dominance of b ions over y ions in this spectrum. One would assume that having the basic lysine situated on the C-terminus would result in the charge staying with the C-terminus upon fragmentation. However, the  $b_4$  and  $b_3$  ion

fragments are the two most abundant ions in the spectrum. It is possible that the amine side chain is donating the sequestered extra proton to an oxygen group on one of these residues, but until further study is conducted on tryptic peptides like this one, the mechanism can not be fully understood. One way to learn more about this feature is to use different amino acids in the tryptic pentapeptides. Other side chains may produce a different dominance between the b and y ions.

The AAAAO fragmentation had a dominant peak at the  $b_4$  ion while the parent with loss of water had a relative abundance of 97%. The fact that the  $[M+H-H_2O]$  ion had such a large abundance is very interesting since it coincides with the  $b_2$  ion abundance for AOAAA and the  $b_3$  ion abundance for AOAAA. Technically the  $[M+H-H_2O]$  is the  $b_5$  ion, but it usually isn't referred to in this way. Indeed, a pattern seems to have developed in considering the dominant fragment in the ornithine containing peptides. Whether it is due to some cyclic intermediate formation has yet to be determined. The next step involves putting ornithine in other locations and also putting it in peptide chains with other amino acids besides alanine.

Of particular interesting with the Dapa, and to a lesser extent, Daba-containing fragmentation spectra of all pentapeptide positions, is the overbearing nature of the parent with loss of water. When Dapa is in any three of the positions, this ion is the most abundant peak, while with Daba it is the most abundant for AAAADaba and ADabaAAA and has a relative abundance of 40 %, the second largest peak, in AADabaAA. One possibility of this happening is that these chains are so short that they don't play a large role in any dominant fragmentations. In this case, the CID energy would most likely pull off only a water molecule since it is easier to remove than a C-N peptide bond. A way to test whether these amine side chains are causing any dominant patterns is to observe an AAAAA peptide chain and look at the differences in this



pentapeptide fragmentation to the Dapa and Daba containing pentapeptides. Obviously, the Daba side chain is having some effects on the fragmentation since it shows considerably different fragmentation spectra than the Dapa-containing in each position, but more work needs to be done to figure out what exactly those effect are.

### 4.3 Conclusions and Future Work

The fragmentation of six dipeptide pairs containing either ornithine or lysine and twelve pentapeptides containing lysine, ornithine, Daba, or Dapa have been studied. Comparing these fragmentation spectra resulted in finding some interesting effects on varying the amine side chain length. In the dipeptide analysis, a cleaner spectrum was found when the dipeptide contained ornithine instead of lysine. Furthermore, studying the spectrum of dipeptides with very basic residues are on the N-terminus have given otherwise unseen  $b_1$  ions. This result has also been confirmed by other groups [28]. Having looked at the spectra of AAXAA, AAAAX, and AXAAA pentapeptides with lysine, ornithine, Daba, or Dapa even more interesting features on side chain length effects can be observed. Of most interest is the pattern in which ornithine-containing peptide fragmentations result in the domination of the  $b_n$  fragment where  $n$  corresponds to the position that ornithine resides on the peptide chain. This is countered by the lack of a similar pattern in the lysine-containing pentapeptides. Furthermore, the same difference in the amount of peaks observed in the dipeptide studies between lysine and ornithine persists with the pentapeptides fragmented. For some reason, removing one methylene group from the lysine side chain drastically reduces the number of significant peaks in the spectrum. Finally, the Daba and Dapa spectra don't show any real significant patterns except for the fact that Dapa always has a dominant peak at  $[M+H-H_2O]$ . They also provide evidence that the

ornithine pentapeptides may be forming stable intermediates that are dominating the fragmentation spectra for each position.

Clearly, additional work must be done in establishing more concrete ideas about these side chain length effects. First, the dipeptides synthesized with either lysine or ornithine must also be synthesized with Daba and Dapa. Additional fragmentation studies on dipeptides containing other amino acids with lysine and its homologs would also be a worthwhile endeavor. As for the pentapeptide studies, the other two positions XAAAA and AAAXA should be synthesized and fragmented. Moreover, changing which amino acids the lysine or homolog are with in the peptide chains could produce significant results and rule out possible ideas about the features observed. Lastly, putting these four amino acids in peptides with longer lengths is a necessary step in fully understanding how chain length affects fragmentation side. Pentapeptides are rather short chains when considering the length of many peptide chains that are studied in protein sequencings. If possible, chains of far greater length should be synthesized and fragmented so that a better understanding of the side chain length effects could be had for more complex, and more commonplace, peptide chains.

## References

- <sup>1</sup>McKee, T.; McKee, J. R.; Biochemistry: The Molecular Basis of Life, Third Edition. New York: McGraw-Hill Companies.; **2003**.
- <sup>2</sup>Steen, H.; Mann, M.; The ABC's (and XYZ's) of Peptide Sequencing. *Nature Reviews: Molecular Cell Biology*, **2004**, *5*, 699-711.
- <sup>3</sup>Nesvizhskii, A. I.; Aebersold, R.; Interpretation of Shotgun Proteomic Data. *Molecular and Cellular Proteomics*, **2005**, *4*, 1419-1440.
- <sup>4</sup>March, R. E.; Todd, J. J.; Quadrupole Ion Trap Mass Spectrometry, 2<sup>nd</sup> Edition. New Jersey: John Wiley and Sons, Inc.; **2005**.
- <sup>5</sup>Breci, L. A.; Tabb, D. L.; Yates, J. R.; Wysocki, V. H.; Cleavage N-Terminal to Proline: Analysis of a Database of Peptide Tandem Mass Spectra. *Anal Chem*, **2003**, *75*, 1963-1971.
- <sup>6</sup>Roepstorff P.; Kohlman, J.; Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. *Biomed Mass Specrom*, **1984**, *11*, 601.
- <sup>7</sup>Biemann K.; Mass Spectrometry of Peptides and Proteins. *Annu Rev Biochem*, **1992**, *61*, 977-1010.
- <sup>8</sup>Johnson R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T.; Novel Fragmentation Process of Peptides by Collision-Induced Decomposition in a Tandem Mass Spectrometer: Differentiation of Leucine and Isoleucine. *Anal Chem*, **1987**, *59*, 2621-2625.
- <sup>9</sup>Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A.; Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation. *J Mass Spectrom*, **2000**, *35*, 1399-1406.
- <sup>10</sup>Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V.; Influence of Peptide Composition, Gas Phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model. *J Am Chem Soc*, **1996**, *118*, 8365-8374.
- <sup>11</sup>Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R.; Protein Sequencing by Tandem Mass Spectrometry. *Proc Natl Acad Sci USA*, **1986**, *83*, 6233-6237.
- <sup>12</sup>Paisz B.; Suhai, S.; Combined Quantum Chemical and RRKM Modeling of the Main Fragmentation Pathways of Protonated GGG. II. Formation of b<sub>2</sub>, y<sub>1</sub>, and y<sub>2</sub> ions. *Rapid Commun Mass Spectrom*, **2002**, *16*, 375-389.
- <sup>13</sup>Bojesen G. J.; *Am. Chem. Soc.* **1987**, *109*, 5557
- <sup>14</sup>Gorman, G. S.; Spier, J. P.; Turner, C. A.; Amster, I. J. J.; *Am. Chem. Soc.* **1992**, *117*, 23.
- <sup>15</sup>Li, X.; Harrison, A. G.; *Org. Mass Spectrom.* **1993**, *28*, 366.

- <sup>16</sup>Wu, Z.; Fenslau, C.; *Rapid Commun. Mass Spectrom.* **1992**, *6*, 403.
- <sup>17</sup>Wu, Z.; Fenslau, C.; *Rapid Commun. Mass Spectrom.* **1994**, *8*, 777.
- <sup>18</sup>Olsen, J. V.; Ong, S.; Mann, M.; Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. *Molecular & Cellular Proteomics*, **2004**, *3*, 608-614.
- <sup>19</sup>Akaogi, J.; Barker, T.; Kuroda, Y.; Nacionales, D. C.; Yamasaki, Y.; Stevens, B. R.; Reeves, W. H.; Satoh, M.; Role of Non-protein Amino Acid L-canavanine in Autoimmunity. *Autoimmunity Reviews*, **2006**, *5*, 429-435.
- <sup>20</sup>Stryer L.; *Biochemistry*, 3<sup>rd</sup> ed.; W. H. Freeman and Co.:New York, **1988**.
- <sup>21</sup>Schroeder, O. E.; Andriole, E. J.; Carver, K. L.; Colyer, K. E.; Poutsma, J. C.; Proton Affinity of Lysine Homologues from the Extended Kinetic Method. *J Phys Chem A*, **2004**, *108*, 326-332.
- <sup>22</sup>Chan, W. C.; White, P. D.; *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*. New York: Oxford University Press Inc.; **2000**.
- <sup>23</sup>Breci, L. A.; Thesis. Written under Wysocki, V. H. advisorship.
- <sup>24</sup>Hood, C. A.; Fuentes, G.; Patel, H.; Page, K.; Menakuru, M.; Park, J. H.; Fast Conventional Fmoc solid-phase peptide synthesis with HCTU. *J Pept Sci*, **2008**, *14*, 97-101.
- <sup>25</sup>Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I.; *Anal Biochem*, **1970**, *34*, 595.
- <sup>26</sup>Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. B.; Harrison, A. G.; Why are b ions Stable Species in Peptide Spectra? *J Am Soc Mass Spectrom*, **1995**, *6*, 1165-1174.
- <sup>27</sup>Nold, M. J.; Wesdemiotis, C.; Yalcin, T.; Harrison, A. G.; Amine Bond Dissociation in Protonated Peptides. Structures of the N-terminal Ionic and Neutral Fragments. *Int J Mass Spectrom and Ion Processes*, **1997**, *164*, 137-153.
- <sup>28</sup>Hiserodt, R. D.; Brown, S. M.; Swijter, D. F. H.; Hawkins, N.; Mussinan, C. J.; A Study of b<sub>1</sub>+H<sub>2</sub>O and b<sub>1</sub>-ions in the Product Ion Spectra of Dipeptides Containing N-Terminal Basic Amino Acid Residues. *J. of Amer. Soc. for Mass Spectrom.* **2007**, *18*, 1414.