



**This electronic thesis or dissertation has been
downloaded from Explore Bristol Research,
<http://research-information.bristol.ac.uk>**

Author:

JIANG, XIAOTONG (CANDY)

Title:

Systematic analysis of the fragmentation of natural products by tandem mass spectrometry and computational method

General rights

Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>. This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy

Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

Systematic Analysis of the Fragmentation of Natural Products by Tandem Mass Spectrometry and Computational Method

By Xiaotong (Candy) Jiang

Under the supervision of Dr Paul J. Gates and Dr Chris Arthur



A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science

School of Chemistry
April 2023
Word count: 50336

Acknowledgement

Thanks to my supervisors, especially Paul, for his continuous support and incredibly helpful knowledge.

Also, thanks to my husband, for being the llama of my life, who carried me through a few bumps in life.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: XIAOTONG JIANG

DATE: 09/10/2023

Table of Contents

Contents

1.	Introduction	1
1.1.	History of Mass Spectrometry: Overview	1
1.2.	Ionisation Methods.....	1
1.2.1.	A Brief History.....	1
1.2.2.	Choice of Methods	3
1.2.3.	High Energy Methods.....	5
1.2.4.	Energy Sudden Methods.....	7
1.2.5.	Low Energy Methods	9
1.2.6.	Surface Analysis Methods	16
1.3.	Mass Analyser – The Orbitrap Mass Analyser	17
1.3.1.	Technical Description.....	18
1.3.2.	Recent Applications.....	19
1.4.	Tandem Mass Spectrometry	20
1.5.	Collision Induced Dissociation and Higher-energy Collisional Dissociation.....	21
1.6.	Ion Mobility Spectrometry	23
1.6.1.	Travelling Wave Ion Mobility Spectrometry	24
1.6.2.	Collision Cross Section.....	24
1.6.3.	Ion Mobility Spectrometry Applications	25
1.7.	Computational Methods	26
1.7.1.	The Development of Gaussian	26
1.7.2.	Hartree-Fock and Density Functional Theory.....	26
1.7.3.	Hybrid Functionals	27
1.7.4.	Basis Sets	28
1.7.5.	IMoS.....	29
1.8.	Natural Products.....	31
1.9.	Aims and Objectives	32
2.	Amino Acids.....	33
2.1.	Amino Acids Overview.....	33
2.1.1.	Amino Acids Analysis by Mass Spectrometry.....	35
2.1.2.	Aims and Objectives.....	38
2.2.	Experimental.....	39
2.2.1.	Amino Acids Samples	39
2.2.2.	Instrumentation	39
2.2.3.	Computational.....	40
2.3.	Results and Discussion.....	41
2.3.1.	General Trends	41

2.3.2.	Tandem Mass Spectrometry Analysis	46
2.3.3.	Energy Breakdown Graph.....	63
2.3.4.	Mechanisms	74
2.3.5.	Computational Studies.....	94
2.4.	Conclusion.....	106
3.	Flavonoids	108
3.1.	Flavonoids Overview	108
3.1.1.	Studies of Flavonoids by Mass Spectrometry	112
3.1.2.	Aims and Objectives.....	116
3.2.	Experimental.....	117
3.2.1.	Flavonoids Samples.....	117
3.2.2.	Instrumentation	117
3.2.3.	Computational.....	117
3.3.	Results and Discussion.....	118
3.3.1.	Overview.....	118
3.3.2.	Common Losses	118
3.3.3.	Diagnostic Product Ions.....	120
3.3.4.	Energy Breakdown Graphs	122
3.3.5.	Common Losses Negative Ion Mode	133
3.3.6.	Loss of CH ₄	135
3.3.7.	Competition of CH ₃ [•] and CH ₄ Losses	138
3.3.8.	Proton Affinity	140
3.4.	Conclusion.....	141
4.	Polyether.....	142
4.1.	Polyether Overview.....	142
4.1.1.	Studies of Polyether Antibiotics by Mass Spectrometry	144
4.1.2.	Aims and Objectives.....	146
4.2.	Experimental.....	147
4.2.1.	Analytes	147
4.2.2.	Instrumentation	147
4.2.3.	Computational Studies.....	148
4.3.	Results and Discussion.....	149
4.3.1.	Overview.....	149
4.3.2.	Salinomycin Monosodiated Ion	151
4.3.3.	Narasin Monosodiated Ion.....	156
4.3.4.	Salinomycin Metal Complexes	160
4.3.5.	IMS-MS/MS	170
4.4.	Conclusion.....	178
5.	Conclusions	179

Abstract

This thesis consists of the studies of three types of natural products - *α*-amino acids, flavonoids, and polyether ionophores. A natural product is defined as a natural compound or substance produced in nature by living organisms. A better understanding of natural products from an analytical approach is crucial to many scientific disciplines. This research focused on the systematic analysis of the fragmentation of natural products by mass spectrometry and computational methods in order to develop a robust methodology that can be applied to a broader range of natural products and other research areas.

The first part of this research used tandem mass spectrometry to systematically analyse the fragmentation of all 20 proteinogenic *α*-amino acids in both positive and negative ionisation modes. The combination of computational modelling with the correlation between side chain functionalities and fragmentation pathways is compared and summarised. Isomers such as leucine and isoleucine are differentiated, and the differences are rationalised using the developed methodology. The second part tested and improved the analytical methodology in the analysis of flavonoids. More than 30 flavonoids from six classes were investigated. Isomeric flavonoids were differentiated. Diagnostic product ions from the positive ion mode collision induced dissociation spectra were identified.

Polyether ionophore antibiotics were selected as the final part of the research due to their structural complexity, importance in pharmaceutical development, and tendencies to isomerise in solution. The unique properties of the polyethers provide an optimal ground to test the robustness of the methodology developed in the studies of the smaller molecules. By combining high-resolution ion mobility mass spectrometry data, tandem mass spectrometry, and computational methods, differentiation of isomers of salinomycin and narasin was achieved. The addition of the study of metal cation complexation provided novel fragmentation pathways, which led to a better understanding of the structural elucidation of the selected polyethers. Complete assignment of the product ions was also achieved for salinomycin monosodium cation and its isomer equivalent. Additional computational data offered a degree of improved rationalisation to the experimental observations.

Results generated from this research demonstrated the powerful approach of combining mass spectrometry with computational methods for the analysis of natural products. It provides an excellent foundation for more challenging future work in related research fields.

1. Introduction

1.1. History of Mass Spectrometry: Overview

Mass spectrometry (MS) is a fundamental analytical tool in nearly every discipline within applied and chemical sciences, and it is used for both qualitative and quantitative studies in association with chromatography. A mass spectrometer is an instrument that generates gas-phase ions, separating them based on their mass-to-charge ratios (m/z) using electric, magnetic, or radio frequency fields and recording the abundance distribution of the ions. Analytes are introduced into the ion source by an inlet system, for example, a syringe, a probe, or a plate. Ionisation occurs and produces ions (either positively or negatively charged), which are then transferred into the mass analyser for separation based on their m/z ratio. A detector is used to register and amplify the arriving ions. A data system is also necessary to control the entire process and to store/manipulate the results.

The earliest history of MS began with the discovery of positive gas-phase ions in a low-pressure electrical discharge tube in 1886.¹ It is followed by observations that these ions can be deflected by strong electric or magnetic fields.² This laid the foundation of MS, which allowed Sir Joseph John Thomson to develop the first mass spectrometry instrumentation and demonstrate its important application in chemical analysis.³ The first commercial mass spectrometer was made available in the 1940s due to the high demand in the petroleum industry for its usefulness in detecting organic molecules.⁴ Later, MS was coupled with gas chromatography (GC-MS). This has dramatically broadened the use of MS in other disciplines, such as biochemistry and environmental science.⁵

1.2. Ionisation Methods

1.2.1. A Brief History

Since the 1950s, the application of MS broadened into disciplines such as biochemistry through the coupling of GC to create the hybrid technique of GC-MS.⁶ Gas chromatography can separate thermally stable biological compounds such as carbohydrates. The mass spectrometric instruments of the time could only detect low molecular mass volatile compounds, so analysis in biochemistry laboratories was also limited to thermally stable volatile compounds. In addition, the ionisation method used was electron ionisation (EI).⁷ A high energy or 'hard' ionisation method involves transferring high internal energies

to the analytes during the ionisation process, resulting in fragmentation and degradation of complex biomolecules.

Due to the limitations associated with EI, Chemical ionisation (CI) and Laser desorption/ionisation (LDI) were developed in the late 1960s.^{8,9} In comparison with EI, CI gives the advantage of yielding molecular ions with less fragmentation. Molecular species can be more easily recognized. However, it still has its limitations in the mass range and requires thermally stable and volatile samples. LDI is efficient in producing gaseous ions¹⁰, however, the quality of the spectra obtained depends on the physical properties of the analyte. In addition, fragmentation is always observed with a sample mass above approximately 500 Daltons (Da).¹⁰ In 1974, plasma desorption (PD) was introduced and allowed the observation of biomolecules with a molecular weight of up to 10,000 Daltons.¹¹ However, a radioactive source was needed, and the data collection was slow. Fast atom bombardment (FAB) and Liquid secondary ion mass spectrometry (LSIMS) were introduced in the early 1980s.^{12,13} Larger biomolecules such as peptides and nucleotides could be analysed. But, regarding sample preparation, the analyte must be dissolved in a non-volatile liquid matrix, which requires the analyte to be soluble and stable in the matrix. Moreover, high chemical background noise may be observed, and matrix or matrix/salt cluster ions can dominate the lower m/z range of the spectrum.

As a result, in the late 1980s, it still seemed impossible to obtain an intact molecular ion with masses higher than 10,000 Da from thermally unstable non-volatile compounds. These limitations were addressed by the development of Matrix-assisted laser desorption/ionisation (MALDI) and Electrospray ionisation (ESI) mass spectrometry.^{14,15} With these new ionisation techniques, the analysis of large biomolecules is made viable, intact molecular ions from non-volatile compounds can be obtained as well as multiply charged ions. Both MALDI and ESI are catalogued as ‘soft ionisation techniques’. It was only after the introduction of these techniques that MS moved into the realms of proteomics and biomedical analysis.

The development of ionisation methods in MS is an ever-growing and rapidly evolving area. As a result, this thesis will mainly focus on a discussion of ionisation methods found in the MS analysis of natural products and relevant previous studies. Ionisation techniques such as MALDI are mentioned due to their importance and wide usage, but they are not explained in detail. Discussion of the relevant ionisation methods is divided into four sections: high energy methods, energy sudden methods, low energy methods, and surface analysis methods.

1.2.2. Choice of Methods

An important consideration in ionisation for MS is the transfer of energy that occurs during the process. In particular, the degree of fragmentation of the molecular ion is proportional to the amount of excess internal energy the ion possesses immediately after ionisation. When this excess internal energy is large, i.e., a hard ionisation method is employed, extensive fragmentation can occur. In comparison, a soft ionisation method refers to reduced excess internal energy in the molecular ion. Therefore, much less fragmentation is observed. Compared with EI, CI and FAB are relatively softer ionisations. However, some fragmentation is still observed. MALDI and ESI are catalogued under very soft ionisation methods.

The other important parameters to be considered are related to the physico-chemical properties of the analyte.¹⁶ Specific ionisation methods are only suitable for the ionisation of analytes with specific physico-chemical properties. For example, EI and CI can only analyse samples in the gas phase. So, these methods are restricted to volatile and thermally stable compounds. Alternative ionisation methods must be used for analytes with insufficient vapour pressure or thermally labile. With these analytes, liquid-phase or solid-state ion sources can be employed to extract molecules into the gas phase and ionise them. ESI is a liquid-phase method, whilst LDI, FAB, and MALDI are all solid-phase methods. Table 1 shows a concise summary of the main features of the most commonly employed ionisation techniques.

Ionisation technique	Properties of analytes	Sample introduction	Typical mass range*	Brief description
Electron ionisation (EI)	Volatile, thermally stable	GC, or heated solids/liquids probe	<1000 Da	hard, 70 eV electron beam, considerable fragmentation
Chemical ionisation (CI)	Volatile, thermally stable	GC or heated solids/liquids probe	<1000 Da	soft, protonated reagent gas, less fragmentation than EI
Laser Desorption Ionisation (LDI)	Non-volatile, thermally labile	Solids probe or plate	500-2,000 Da	soft, laser irradiation, limited ion yield
Plasma Desorption (PD)	Non-volatile, thermally labile	Solids probe or plate	500-25,000 Da ¹⁷	soft, ²⁵² Cf fission fragments transfer energy to analyte

Fast atom bombardment (FAB)	Non-volatile, thermally labile	Solids probe	500-3,000 Da	soft, high energy atom beam, matrix driven
Liquid Secondary Ion Mass Spectrometry (LSIMS)	Non-volatile, thermally labile	Solids probe	500- 10,000 Da ¹⁸	soft, high energy ion beam, matrix driven
Matrix-assisted laser desorption ionisation (MALDI)	Non-volatile, thermally labile	Sample plate	100- 1,000,000 Da ¹⁹	very soft, UV laser, matrix driven, singly charged intact molecular ions
Electrospray ionisation (ESI)	Non-volatile, thermally labile	Liquid inlet: LC, syringe pump	no limitation	very soft, electrochemical process, singly and multiply charged intact molecular ions
Atmospheric Pressure Chemical Ionisation (APCI)	Non-volatile, thermally labile, Mostly non polar	Liquid inlet: LC, syringe pump	\leq 1500 Da	very soft ionisation, water/nitrogen plasma, very little fragmentation for analytes with less functionality

Table 1 Summary of the various features of EI, CI, FAB, LSIMS, MALDI, ESI, and APCI. ^{10, 20, 21}

**depends on the instrument.*

1.2.3. High Energy Methods

1.2.3.1. Electron Ionisation

Electron ionisation (EI) is one of the ‘classic ionisation techniques’ for MS, first described by Dempster⁷ in 1918 and further modified by Bleakney, to improve the technique.²² Because EI is a gas phase technique, it requires the analyte to be thermally stable and volatile. It involves transferring a high level of energy to the analyte during ionisation, so more labile analytes will undergo extensive fragmentation in an attempt to reduce that energy. As a result, molecular ions are often not observed, or they are at a low relative intensity. EI is therefore categorised as a ‘hard’ ionisation method.

A schematic drawing of an EI source is shown in Figure 1. A droplet of the analyte in solution in a volatile solvent is delivered by a liquids probe and vaporised. If the analyte is thermally labile, it will fragment at this stage. The analyte vapour enters the ionisation chamber, which contains a tungsten filament. When the filament is heated, it emits electrons, which are then collimated into a beam and steered across the chamber towards the electron trap. During this process, the electrons interact with the gaseous molecules of the analyte. The probability of interactions between the electron beam and the analyte molecules is increased by fixing a magnet with opposite poles across the chamber, creating the electron beam to spiral.

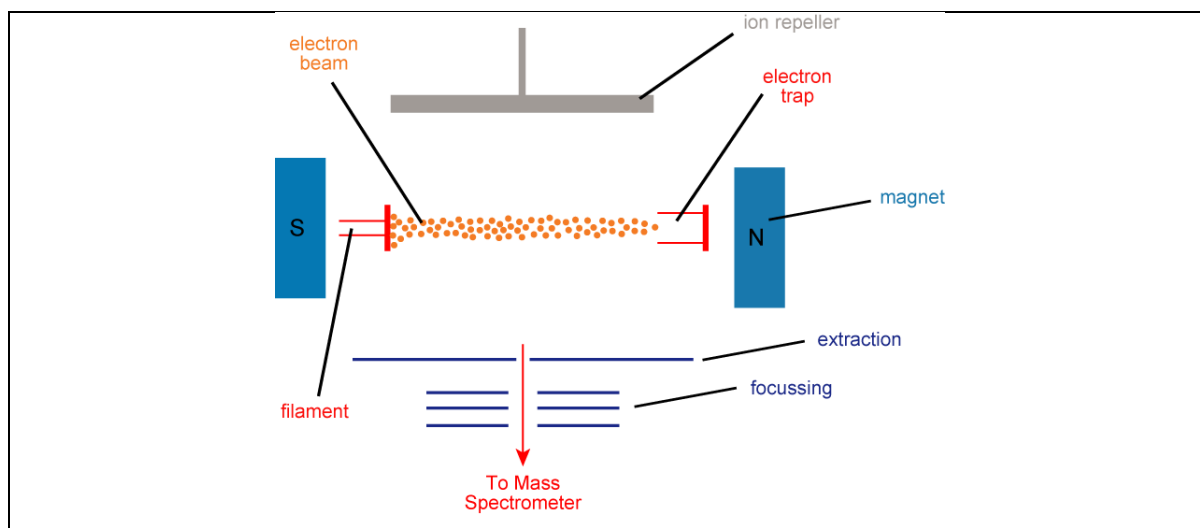


Figure 1. Schematic diagram of an electron ionisation source.²³

The fundamental physics of EI is based on De Broglie’s relation that particles such as electrons can possess wave-like properties.²⁴ He combined Einstein’s famous $E = mc^2$ with the Planck-Einstein relation $E = h\nu$ to come up with a function for wave-particle duality which can be applied to electrons:

$$\lambda = \frac{h}{p} = \frac{h}{mv}$$

Equation 1

Where λ is the wavelength of the electron, h is Planck's constant, p is linear momentum, which equals m , mass times v , velocity. When the kinetic energy of the electron is 70 eV, the wavelength is about 1.47 Å, which is close to the average carbon-carbon bond length of an organic molecule. This results in a disturbance in the electric field of the analyte molecules. If a wave frequency has enough energy $h\nu$ for molecular electronic excitation, an electron can be expelled from the neutral molecule.²⁵ Thus, the energy transferred from the electron to the molecule leads to the ionisation of the analyte molecule. Based on de Broglie's relation, when the electrons' kinetic energy is too small, there will not be enough energy to induce ionisation. At the same time, when the kinetic energy is too high, and the wavelength is too small compared to the typical bond length, the analyte becomes 'transparent' to the electrons, decreasing the ionisation efficiency.²⁶ Therefore, in the analysis of organic molecules, the acceptable maximum of electron energies is normally around 70 eV. Energies higher than 70 eV are not likely to increase ionisation efficiency beyond this threshold.²⁵

In the EI process, only between 10 eV and 20 eV is transferred to the molecule despite the electron beam being at 70 eV. This is more than capable of ionising most organic molecules. The excessive energy can result in extensive fragmentation observed in the resulting mass spectra. Fragmentation can be reduced by lowering the energy but at the cost of decreased ionisation efficiency. Due to this nature of creating extensive fragmentation, EI can be extremely useful for obtaining structural information of unknown analytes or by database matching the fragmentation for known analytes against the database.²⁷

1.2.3.2. Chemical Ionisation

Chemical ionisation (CI) is developed as a lower energy alternative to EI as it proved too challenging to yield intact molecular ions with many analyte classes. The technique of CI was introduced by Munson and Field in 1966.⁸ The analyte molecules react with reagent ions in the gas phase to create analyte ions by the transfer of a charged species. CI produces ions with much less excess energy when compared to EI, resulting in much less fragmentation and more information on molecular weight being obtained from a wider range of analytes. Both EI and CI follow the same basic setup and source design, and it is common in modern applications to have EI and CI sources on the same instrument. Data obtained from CI is complementary to EI, and in combination, they can be a powerful tool for structural elucidation.

Although CI is a much softer ionisation technique than EI, it still requires the analytes to be vaporised before ionisation. Thermal degradation of the analyte can sometimes still lead to considerable fragmentation. So, as a result, as with EI, CI is limited to thermally stable and volatile samples. Furthermore, the sensitivity of this technique depends strongly on the experiment conditions, such as the reagent gas used, the reagent gas pressure, and the source temperature. Reagent gases could be methane, isobutane, or ammonia, and the primary reaction would be a classical EI reaction to generate a cascade of chemical reactions. The analyte would then be ionised by the ions generated by the reagent gas reactions.

1.2.4. Energy Sudden Methods

1.2.4.1. Fast Atom Bombardment

Fast atom bombardment (FAB) was introduced in the early 1980s.²⁸ It introduced the capability of analysing thermally labile and involatile compounds, especially when compared to EI and CI. FAB uses a discharged beam of fast atoms of an inert gas (typically argon or xenon) to bombard the target analyte. It requires the analyte to be dissolved in a non-volatile liquid matrix, most often glycerol or nitro-benzyl alcohol.^{29,30,31} The liquid matrix can maintain a fresh surface for continued bombardment, thus extending the ion yield, sensitivity, and spectral lifetime of the analysis. The mixed sample solution is placed on the tip of a probe and introduced directly into the ion source. A fast-moving particle beam is then shocked onto the surface of the analyte/matrix spot, transferring some of its kinetic energy to the analyte/matrix solution surface and ejecting ions into the gas-phase (Figure 2). These ions are then extracted and focussed before passing into the mass analyser.

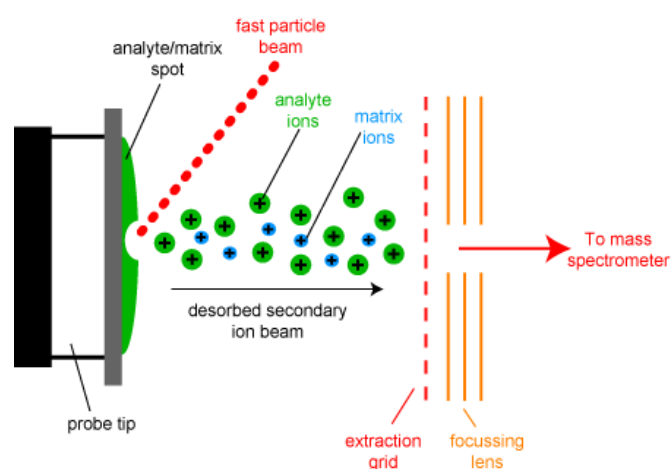


Figure 2 Schematic diagram of a fast atom bombardment ion source.³²

FAB has the benefit of being able to analyse thermally sensitive samples. It can also produce ion beams that can be maintained for as long as tens of minutes, allowing various additional analyses to be carried out, especially tandem mass spectrometric studies.²⁹ However, the addition of the matrix can cause the spectrum to become more complex to analyse, especially for lower massed analytes. The popularity of FAB rapidly decreased after the introduction of the more recent desorption/ionisation methods such as ESI and MALDI.

1.2.4.2. Plasma Desorption

Plasma desorption (PD), developed by Macfarlane and Torgerson,³³ is an ionisation technique in which the analyte is deposited on a small, aluminised nylon foil. It is then bombarded with ions and/or neutral atoms formed from the fission of ²⁵²Cf. The shock waves resulting from the bombardment induce the desorption of neutrals and ions from the analyte.²⁹ Although PD was successfully used for the analysis of natural products³⁴, peptides³⁵, and short-chain polymers throughout the 1980s and 1990s, the difficulties and hazards of handling radioactive ²⁵²Cf made PD a very limited ionisation technique. The introduction of MALDI immediately made PD obsolete, and it was rapidly confined to being a specialist technique with limited applications.

1.2.4.3. Laser Desorption Ionisation

Laser desorption ionisation (LDI) in MS was first reported in the early 1960s and applied in analytical studies throughout the 1970s and 1980s.^{36,37} It was efficient in generating ions in the gas phase. In some cases of LDI, the laser could be used as a heat source only, which could induce the desorption of large organic molecules with very limited volatility from the solid phase into the gas phase thermally.³⁸ In other cases, a laser pulse from the range of $10^6 - 10^8 \text{ Wcm}^{-2}$ is used to focus on a solid sample surface. This depletes a part of the surface, creating a microplasma of ions and neutral molecules in the gas phase. The ions and neutral molecules can then react with themselves in the dense vapour near the sample surface through ion-molecule collision reactions.³⁹ In LDI, the laser pulses are capable of vapourising and ionising the sample simultaneously. LDI is also useful for studying the surfaces and for the elemental analysis of solid samples.^{40,41} However, whether a useful mass spectrum could be obtained depends largely on the physical properties of the analyte. Samples larger than 500 Da nearly always resulted in fragmentation.⁴² One of the most recent studies is using LDI-MS in the *in-situ* detection of amino acids on extra-terrestrial surfaces.⁴³ The application of LDI has dramatically expanded due to the development of MALDI. Modern MALDI instruments allow the use of LDI via matrix free methods.⁴⁴

1.2.5. Low Energy Methods

1.2.5.1. Matrix-Assisted Laser Desorption Ionisation

LDI has its limitations in the analysis of larger organic compounds. The energy required to ablate the analyte surface almost always results in fragmentation. To overcome this obstacle, a new softer ionisation technique was required.

The term matrix-assisted laser desorption ionisation (MALDI) was coined in 1985 by Karas & Hillenkamp et al.⁴⁵ In this study, a mixture of alanine and tryptophan was analysed. It stated that tryptophan functioned as an absorbing matrix, resulting in molecular ion formation of the nonabsorbing alanine. In 1987, another paper reported that mellitin, a polypeptide of 2843 Da, could be ionised when mixed with a matrix.⁴⁶ In 1988, Koichi Tanaka et al. used an 'ultra-fine metal powder plus liquid matrix' method for ionisation. An ion cluster of more than 100 kDa was successfully observed.^{14,47} Further developments included the ionisation of a 67 kDa protein using nicotinic acid as matrix,⁴⁸ and a protein of mass 120kDa.⁴⁹ Further improvements were achieved using a 355 nm laser and more varieties of acidic matrices.⁵⁰ The first commercial instrument introduced in the early 1990s brought MALDI to an increasing number of researchers. MALDI -MS has since become one of the most widely applied analytical tools in applied analytical sciences. In 2002, Koichi Tanaka won a share of the Nobel Prize in Chemistry for his development of MALDI-MS.

During the preparation for a MALDI-MS experiment, the analyte is first dissolved in a suitable solvent and mixed with an excess of an appropriate matrix. This chosen matrix must have strong absorption at the laser wavelength and be highly sublimable. The mixture is then spotted on a MALDI target plate and dried to remove any trace of solvent. The co-crystallised mixture forms a 'solid solution' deposit. The matrix separates the analyte molecules from each other to ensure the 'solid solution' is homogenous so that stable desorption can be achieved.

The second step occurs in the vacuum when the solid solution is vapourised into the gas phase via a pulsed UV laser beam, typically at 377 nm. The laser beam focuses on the surface of the solid solution, and the chromophore of the matrix couples with the laser frequency, inducing vibrational excitation and causing localised disintegration of the solid solution. Clusters of analyte molecules surrounded by matrix and salt ions are ablated from the surface. Matrix molecules then evaporate off, leaving the free analyte in the gas phase.

Ionisation occurs when the excited matrix molecules are stabilised through proton transfer to the

analyte. A range of characteristic $[M+X]^+$ (M = neutral molecule; X = H, Na, K, etc.) analyte ions are formed through proton transfer or cationisation. These ionisation reactions take place in the desorbed matrix-analyte cloud above the surface. The ions are then extracted into the mass spectrometer for analysis. Figure 3 is a schematic drawing of the process of ion formation during MALDI-MS.

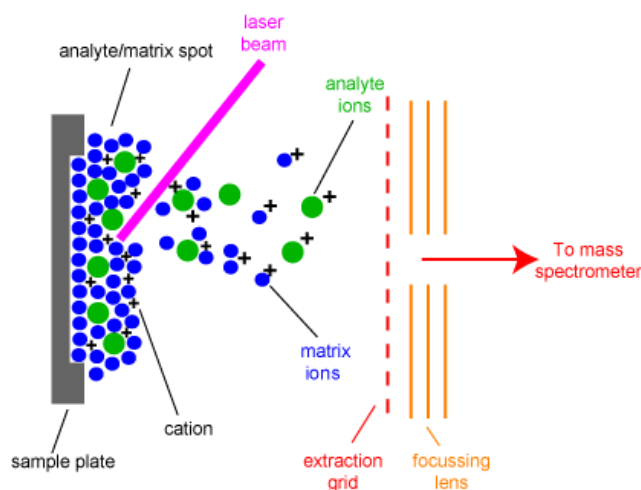


Figure 3 A schematic diagram of the mechanism of MALDI.⁵¹

The exact mechanisms of ion formation in MALDI are a subject of continuing research and discussion.⁵² No single mechanism can explain all the ions observed in a single MALDI experiment. Common matrices used in MALDI experiments are nicotinic acid, sinapinic acid, 2,5-Dihydroxy-benzonic acid etc. MALDI matrices are generally crystalline solids of low vapour pressure. They also have to absorb light of the wavelength used for the experiment. The choice of a matrix is crucial for success in MALDI experiments.⁵² In recent years, researchers have successfully used alternatives to traditional matrices in MALDI experiments, such as graphite, silicon, and fullerenes.^{53,54,55} In addition, MALDI has been used in MALDI imaging where a thin slide of tissue section is inserted into a MALDI mass spectrometer, an image processing software will import the data while the mass spectrometer records the spectrum.⁵⁶

One disadvantage of MALDI is that it has low shot-to-shot reproducibility. Each laser pulse results in a variation of ions observed due to the variations in the make-up of the analyte : matrix spot and the position where the laser is targeted on the spot. Sample preparation is the key here, as the more homogenous the solid solution is, the more reproducible the signal is.

1.2.5.2. Electrospray Ionisation

Electrospray ionisation (ESI) is one of the most widely applied ionisation methods in current MS studies. Its applicability, adaptability, ability to detect multiple charged ions and capability to couple with high-performance liquid chromatography (HPLC) or capillary electrophoresis makes ESI a very popular tool in biological applications, such as the study of proteins and peptides, the study of polymers as well as the study of small molecules in applications such as metabolomics, drug discovery and structural elucidation of natural products.⁵⁷

ESI was first introduced by Dole et al. in 1968⁵⁸, but it was the later studies published by Yamashita and Fenn (in 1984) demonstrating revisions to the technique resulting in its application to larger molecules, including the phenomenon of multiple charging.⁵⁹ John Fenn won a share of the 2002 Nobel Prize for Chemistry for his contribution to the analysis of biological macromolecules. He gave his Nobel lecture a fascinating title, 'Electrospray wings for molecular elephants,' aiming to describe the ability of ESI to produce multiply charged ions of proteins and effectively make the analysis of very large biological molecules achievable.⁶⁰

The general process for ion formation in ESI is divided into three steps: droplet formation, droplet shrinkage, and gaseous ion formation. The sample in solution is usually introduced through a syringe pump or as the eluent flow from liquid chromatography (LC). It passes through a capillary at a high potential difference. The high electric field generates a mist of highly charged droplets, which then pass towards the mass analyser. A schematic illustration of an ESI source is shown below (Figure 4).

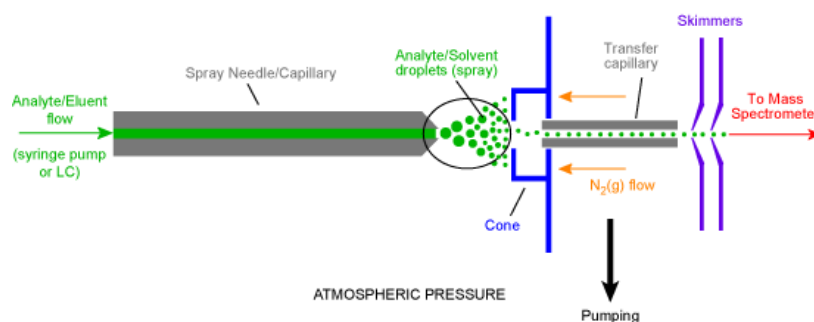


Figure 4 A schematic illustration of an ESI source.⁶¹

To induce the electrospray process, a strong electric field (3-6 kV between the capillary and the counter electrode) is applied to the liquid passing through a capillary tube with a weak flux under atmospheric pressure.¹⁰ The field stimulates charge accumulation on the liquid surface at the end of the capillary,

which then breaks to form highly charged droplets. These droplets then pass through a curtain of heated inert gas (usually nitrogen) or a heated capillary to remove the last solvent molecules. The shapes of the droplets at the tip of the capillary have been investigated, where it was shown that as the voltage increases, the stronger electric field causes more pressure on the accumulated charges at the tip, which elongates the droplets. The surface tension of these droplets is gradually overcome by Coulombic repulsion, and their shapes change to a 'Taylor Cone'⁶², and the process of electrospray begins.⁶³

A few models have been proposed for how ions are formed, and two are the most recognised. Firstly is the 'charged residue model', proposed by Dole et al.⁵⁸ This model describes the formation of smaller droplets by solvent evaporation from the original droplets. This continues until the droplets become so small that only one solute molecule is contained in every droplet. After the remaining solvent evaporates, the remaining solute molecule retains some droplet charge and becomes a free gas-phase ion.⁵⁹ Second is the 'ion evaporation/desorption model' proposed by Iribarne and Thomson.⁶⁴ It agrees with the formation of smaller and smaller droplets but states that instead of droplets reaching a size small enough to contain only one solute molecule, an ion can be lifted by the field strength at the droplet surface. Both models are believed to be valid. However, it has been suggested that larger analytes with molecular weight exceeding 5,000 to 10,000 Da generally undergoes with charged residue model.¹⁰

One of the practical considerations of ESI is that charged ions generated by ESI do not reflect the charge state of the analyte in the solution. They result from charge accumulation in the droplets and the electrochemical processes at the probe tip. This was demonstrated in an experiment by Kelly et al. in 1992.⁶⁵ In addition, the sensitivity of ESI is directly linked to sample concentration rather than the total quantity of solution injected into the source⁶⁶. The study also demonstrated that lower flow rates gave the best sensitivity. As less analyte and solvent were injected into the source, there was less contamination. Due to this dependence of sensitivity on concentration, modifications such as nanospray have been developed to allow much lower flow rates to be used.⁶⁷ Nanospray (nESI) is capable of analysing an extremely small quantity of sample, and it only requires μL of low concentration nmol/mL. This benefit is achieved by reducing the inner diameter of the spray needle to produce smaller droplets. Nanospray also has a higher tolerance for salt contamination, which is an advantage in biological and physiological sample analysis.

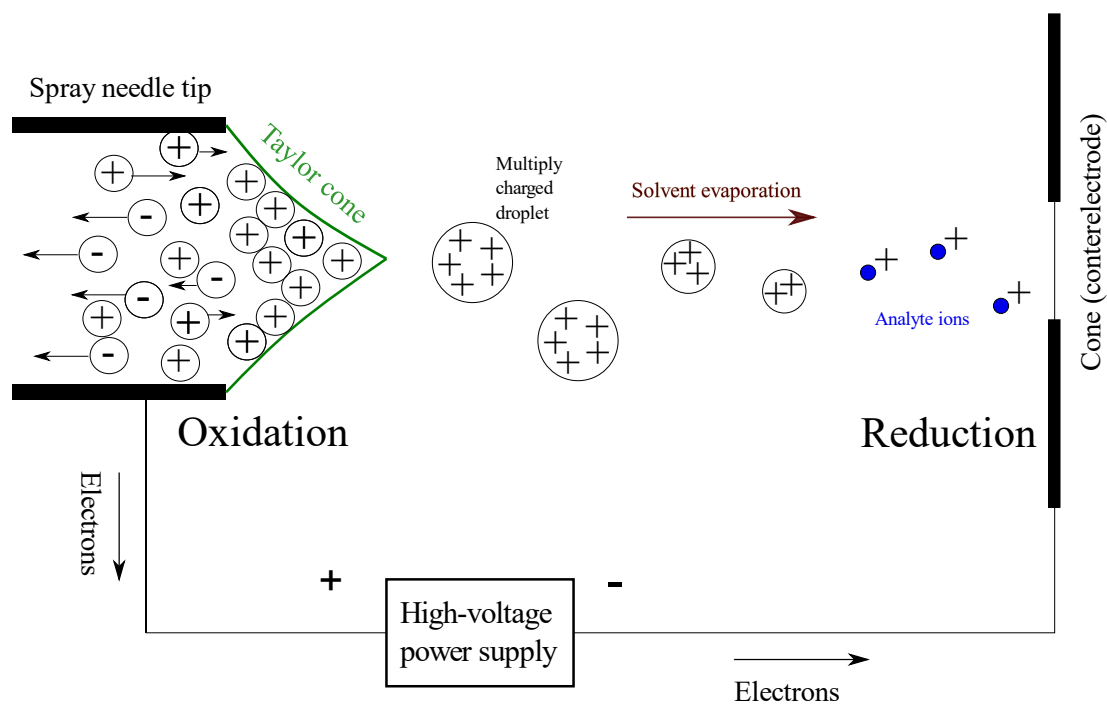


Figure 5 Schematic representation of the electrochemical process in ESI.

Figure 5 is a schematic representation of the electrochemical process in ESI. As mentioned previously, ESI is an electrochemical process that can be affected by ion concentration, ion mobility, and polarisation effects at the probe tip. A study of the negative ion spectrum of myoglobin showed that the signal-to-noise ratio is higher at pH 3 than at pH 10.⁶⁵ This is because protons have high electrochemical mobility. The reduction process happens at the probe tip during the negative ion mode analysis. Based on Figure 5, positive ions are extracted for analysis during the positive ion mode. Electrons provided by the capillary go through the electric circuit. This process is enabled by the oxidation of ions with sufficient mobility at the probe tip. The same number of negative charges must be 'pumped' out of the solution as the positive charges are extracted to the mass analysis. Accumulated positive charges at the probe tip form a 'Taylor cone'. Under a high potential difference, this cone is unstable, and multiply charged droplets enriched with positive charges are emitted from the cone tip.⁶⁸ As the droplets travel from the spray needle tip to the cone, the solvent evaporates until the solute molecule retains some droplet charge and becomes a free gas phase ion.

In a paper published in 1995, Berkel et al. confirmed the analogy of an ESI source as a constant-current electrochemical cell.⁶⁹ Results from this paper explained the occasional observations of radical ions in the ESI spectrum. Assuming the constant ion current in the ESI source is I_M , too many ions from salts in the flow will be able to produce I_M . Ions from the sample will either be at low abundance or not observed. On the contrary, if the solution is at a very low concentration or extremely low flow rate, I_M is not sufficiently sustained. Excessive oxidation will occur at the probe tip to produce additional ions.

This could be from the solvent or the analyte, depending on their redox potentials. Such a process will result in the observation of radical ions in the mass spectrum.

The relation between the ion concentration and the ion abundance can be explained by the following equations mentioned in a book by Fenn et al. and simplified by Hoffmann et al.^{10,70}

Assume ion current I_A is produced by ions of analyte A desorbed from the droplets, then:

$$I_A = k_A[A].$$

Equation 2

where k_A is a rate constant depending on the nature of analyte A, and $[A]$ is the concentration. We also assume that ion B is produced from the buffer, and $[B]$ is the concentration of B then:

$$I_B = k_B[B]$$

Equation 3

No other ions are sprayed in this hypothesis. Therefore, the total ion current I_T for these two types of ions are:

$$I_T = I_A + I_B$$

Equation 4

It was stated previously that this total ion current is limited by the oxidation process at the probe tip in positive ion mode analysis. This limited ion current mentioned earlier is I_M . Therefore, $I_T = I_M$ if only ions A and B are present. The ion current of A and B is proportional to their relative desorption rate, which can be represented as:

$$I_A = I_M \cdot \frac{k_A[A]}{k_A[A] + k_B[B]} \quad I_B = I_M \cdot \frac{k_B[B]}{k_A[A] + k_B[B]}$$

Equation 5

From Equation 5, we assume $[B]$ is constant. Under two extreme conditions of $[A]$, we can deduce that when $k_A[A] \gg k_B[B]$:

$$I_A \approx I_M \cdot \frac{k_A[A]}{k_A[A]} \approx I_M \quad I_B \approx I_M \cdot \frac{k_B[B]}{k_A[A]}$$

Equation 6

And when $k_A[A] \ll k_B[B]$:

$$I_A \approx I_M \cdot \frac{k_A[A]}{k_B[B]} \quad I_B \approx I_M \cdot \frac{k_B[B]}{k_B[B]} \approx I_M$$

Equation 7

These two Equations 5 and 6 showed that when the concentration of A is much higher than the salt ion B from the buffer, I_A remains constant, the signal intensity of B becomes smaller as $[A]$ increases. On the other hand, if $[A]$ is lower than $[B]$, the ion intensity of A is directly proportional to its concentration but inversely proportional to $[B]$. This relationship between ion intensity in ESI and ion concentration has been confirmed by the same study that explored the mechanism of ESI MS.⁶⁸ They demonstrated using protonated Morphine molecular ion as A, a combination of NH_4^+ and Na^+ as B. At a low concentration of A, the intensity of B remains stable. As $[A]$ increases, the signal intensity of A also increases. This experimental observation agrees with Equation 7. As $[A]$ becomes higher than $[B]$, the intensity of A starts to flatten, whereas for B, it reduces considerably, which also agrees with Equation 6.

1.2.5.3. Atmospheric Pressure Chemical Ionisation

Atmospheric pressure chemical ionisation (APCI) was developed in the 1970s by Carroll et al.⁷¹ In the study, they described an atmospheric pressure ionisation (API) method using a corona discharge coupled with a liquid chromatograph – mass spectrometer – computer analytical system. The ionisation process is analogous to CI where ions are produced by a mixture of ion-molecule reactions in the vapour phase.⁷² The analyte in solution from either a direct inlet probe or LC is sprayed through a heated nebulizer into the source at atmospheric pressure. A corona discharge is used to ionise the atmospheric gases, solvent, and analyte molecules to create a mixture of ions. In the positive ion mode, proton transfer, redox reaction and adduction of reagent gas ions can occur.⁷¹ In the negative ion mode, deprotonation is mostly dominant.⁷³ The exact chemical reactions depend on the properties, such as proton affinities of the gaseous analyte molecule and the reagent gases used.

APCI is best suited for analysing small, low-to-medium polarity molecules less than 1500 Da.⁷⁴ It often generates monocharged ions and is useful in studying labile small organic molecules with few functional group varieties. Studies have established successful analysis of amino acids, flavonoids, and other natural products by APCI.^{73,75} Modern APCI ion sources are often incorporated with ESI-MS to enable interchangeable fast ionisation analysis. It has become a popular ionisation method when coupled with HPLC for routine mixture analysis.

1.2.6. Surface Analysis Methods

1.2.6.1. Desorption Electrospray Ionisation

The development of ionisation methods for MS has taken further significant advances since the introduction of ESI and MALDI. There are now several ionisation methods that are performed at atmospheric pressure, which have been discussed in the previous section. There are also methods that can be performed at ambient conditions without any sample preparation. One of these methods is desorption electrospray ionisation (DESI).

DESI was based on ESI and was first described in a paper published in 2004.⁷⁶ It is a technique that uses electrospray to produce ions and charged droplets of selected solvent, which is then directly sprayed onto the surface of an analyte. This ionised beam of solvent desorbs some of the analyte surface materials and extracts them into the gas phase. These desorbed ions of analytes from the surface are then bounced into the atmospheric inlet of the mass spectrometer. According to the same study, DESI was able to ionise nonpolar small molecules such as lycopene and larger polar biological compounds such as proteins.⁷⁶ The surface used in DESI is not limited by its material, it can be free moving and even be a person's finger.⁷⁶ Ions produced in DESI are similar to ESI, they are dominated by monocharged or multicharged molecular ions of the analyte. Changing the spray solvent composition can optimise the resulting mass spectrum.⁷⁷

One of the benefits of DESI is its ability to map the image of the analyte surface while obtaining MS spectra. It has been used to map images of metabolites and lipid tissues such as rat brains.^{78,79} Compared to the more established MALDI imaging, DESI has the advantage of not requiring any sample preparation (in terms of adding a matrix), and it can be performed directly under ambient conditions. DESI imaging has been used in natural product research, such as identifying metabolites in plants⁸⁰ and characterising different bacteria species.^{81,82} The most recent advancement in DESI is to use an integrated microfluidic probe for nano-DESI MS imaging of biological sample tissue. The increased scan rate does not compromise imagery quality and molecular coverage. This development was established to facilitate potential 3D imaging of similar biological tissues.^{83, 84, 85}

1.3. Mass Analyser – The Orbitrap Mass Analyser

The concept of orbital trapping was first introduced by Kingdon in 1923.⁸⁶ In his study, a small filament was put axially through a cylindrical anode with closed ends. Positive ions were described to orbit around the filament until they lost enough energy due to collision with gas molecules, which then fell into the cathode. In 2000, Makarov published a paper in which he thoroughly revised the concept of orbital trapping in utilisation for mass analysis.⁸⁷ A new type of mass analyser was thus demonstrated - the Orbitrap.

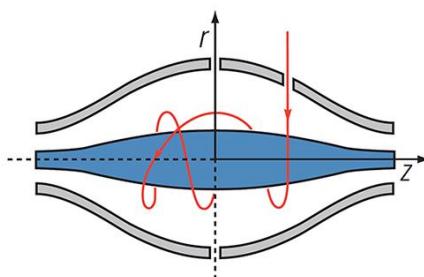


Figure 6 An illustration of orbitrap. It consists of an outer barrel-like electrode and a central spindle-shaped electrode along the axis (r and z). $R1$ and $R2$ are the maximum radii of the corresponding electrodes. Arrows indicate an example of a stable ion trajectory.²⁵

The geometry of the orbitrap is presented in Figure 6. It has an outer electrode with the shape of a barrel and a central electrode with the form of a spindle. Based on ideas from Kingdon, the Orbitrap uses trapping in an electrostatic field, and only direct current voltages are applied. There are no magnetic or radio frequency fields- the ions are stabilised by orbiting around the axial electrode. These orbiting ions also perform harmonic oscillations along the electrode (in the z -axis) with a frequency proportional to their $m/z^{-1/2}$. To obtain mass spectra, ion oscillations are detected by image current detection and transformed using a fast Fourier transformation. Mass resolutions of up to 150,000 were reported in earlier studies.⁸⁷

The benefits of the Orbitrap include high mass resolution, high mass accuracy, high mass stability, high dynamic range, increased upper mass limit, and space charge capacity at higher masses. The reasons behind these advantages have been explained in detail in Makarov's original paper.⁸⁷ Additional study published in 2021 demonstrated that a novel research-only Orbitrap mass analyser could achieve a resolving power over 2,000,000.⁸⁸

Following the introduction of the first commercially available Orbitrap instrument in 2005, the applications of the Orbitrap have broadened to a wide range of scientific fields. A review paper pointed out the ability of the Orbitrap to elucidate the composition of complex samples containing numerous analytes with a wide range of m/z and concentrations. This emphasises the Orbitrap's importance in proteomic and metabolomic studies.⁸⁹

1.3.1. Technical Description

The Orbitrap, also known as an 'electrostatic ion trap', uses Fast Fourier Transform (FT) to obtain mass spectra. FT decomposes functions into individual frequency components and displays the output as a function of frequency. Ions are injected with a specific kinetic energy and oscillate in the trap around the central electrode under the influence of the electrostatic field. This electrostatic field is obtained by the direct current voltage and the unique geometry of the trap.

Equation 8 describes the potential inside the trap¹⁶:

$$U(r, z) = \frac{k}{2} \left(z^2 - \frac{r^2}{2} \right) + \frac{k}{2} (R_m)^2 \ln \left(\frac{r}{R_m} \right) + C$$

Equation 8

Where r and z are the cylindrical coordinates, C is a constant, k is field curvature, and R_m is the characteristic radius. The first section of the equation is a quadratic field, and the second is the logarithmic field of a cylindrical capacitor. The z coordinates only appear in the first part of the first section. Hence the voltage gradient in the z direction can be represented by the partial derivatives with respect to z :

$$\frac{\partial U(r, z)}{\partial z} = kz$$

Equation 9

If an ion of mass m and a charge of q accelerates along z under the influences of the electric field, the electric force $-qkz$ of the electric field is equal to mass m times acceleration (d^2z/dt^2)⁸⁷:

$$F_z = m \left(\frac{d^2z}{dt^2} \right) = -qkz = -q \frac{\partial U(r, z)}{\partial z}$$

Equation 10

And

$$\frac{d^2z}{dt^2} = -\left(\frac{q}{m}\right)kz$$

Equation 11

If E_z is energy, and is defined by⁸⁷:

$$qE_z = \left(\frac{m}{2}\right)\left(\frac{dz_0}{dt}\right)^2$$

Equation 12

The solution to Equation 9 can be given as⁸⁷:

$$z(t) = z_0 \cos \omega t + \sqrt{(2E_z/k)} \sin \omega t$$

Equation 13

With

$$\omega = \sqrt{(q/m)k}$$

Equation 14

Where ω is the angular frequency. This result demonstrates that the frequency is directly linked to the m/q ratio and is independent of the kinetic energy of the injected ions, hence enabling high resolving power and high mass accuracy and stability.¹⁶

1.3.2. Recent Applications

The most common commercial setup for the Orbitrap mass analyser is coupling with an ESI ion source interfaced with HPLC for mixture analysis.⁹⁰ However, recent developments have also seen the Orbitrap interfaced with GC using EI and CI ion sources in the studies of metabolites.^{91, 92} It is also possible to use MALDI, APCI, and DESI ion sources.^{93, 94} Electron transfer dissociation (ETD) can also be performed, and this has been popular in proteomics where it could produce extensive cleavage of peptide backbones while keeping labile bonds intact. ETD allows the identification of primary peptide sequences and post-translational modification, such as phosphorylation.⁹⁵ Modern LTQ (Linear Trap Quadrupole)-Orbitrap allows various ETD configurations to be used to produce quality mass spectra with high resolving power and mass accuracy.⁸⁹

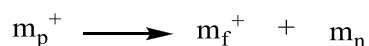
As discussed previously, ambient ionisation methods such as DESI allow direct analysis of unmodified samples in an ambient environment.⁹⁶ Due to the complexity of untreated samples, a high-performance mass analyser is crucial to perform composition studies. Orbitrap coupled with a DESI source has been

demonstrated to be a highly successful tool for characterising peptides, proteins, biological fluids,⁹⁷ plant samples,⁹⁸ synthetic polymers⁹⁹, and active ingredients in pharmaceutical tablets.¹⁰⁰

1.4. Tandem Mass Spectrometry

The development of ESI-MS is crucial to the detection of molecular ions, and the analytical application of ESI is extended further by a combination with tandem mass spectrometry (MS/MS). MS/MS involves at least two stages of mass analysis. A dissociation process or a chemical reaction that causes a change in the mass or charge of an ion is also involved.¹⁰¹

A typical MS/MS experiment uses an initial mass analysis step to isolate a precursor ion - i.e., the ion of interest. This ion then either fragments spontaneously or by activation to generate the product ions and neutral fragments. A mechanism can be shown as follows:



Where m_p^+ is a precursor ion, m_f^+ and m_n represent the product ion and neutral fragments, respectively.

A second mass analysis step then analyses the product ions generated (see Figure 7 for a schematic drawing). The number of steps can be increased to carry out sequential mass spectrometry (MS^n). However, the yield of transmitted ions decreases at every step. In addition, as the number of steps increases, the complexity and time costs of the process may also increase. Therefore, a practical maximum number of steps is often applied during the experimental procedures. The maximum number of steps during our study is typically limited to four or five.

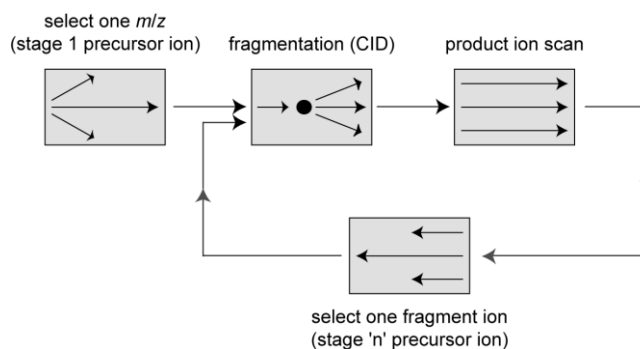


Figure 7 Schematic Drawing of MS/M¹⁰²

Because of the various combinations of mass analysers used in MS/MS, the technique can be performed in two ways: in space and in time. The first describes a coupling of two physically different instruments, and the second involves performing a sequence of events in an ion storage device, for example, an ion trap.

Tandem-in-space mass analysers include sectors, quadrupoles, and Time-of-Flight (ToF).¹⁰³⁻¹⁰⁵ The most common instrument is the triple quadrupole mass spectrometer (QqQ), which was introduced in the late 1970s.¹⁰⁶ The study combined a dual EI/CI source with three quadrupole mass analysers. The lowercase q in the abbreviation QqQ refers to the collision cell, which only operates in the radio frequency mode. Therefore, it can act as a lens for all the ions. Other instruments, such as a combination of a quadrupole with a ToF instrument (Q-ToF), are also commonly used. Generally, the maximum number of mass analysers used in tandem-in-space mass spectrometry is limited to three or four due to the low number of ions transmitted at each step.

Tandem-in-time mass spectrometry involves ion traps and Fourier-Transform mass analysers (e.g. Fourier Transform Ion Cyclotron Resonance (FTICR) or Orbitrap) mass analysers. The instrument is programmed to carry out different steps within the same instrument. In ion trap or orbitrap mass spectrometers, ions can only be observed at the end of the whole process. This leads to faster analysis times – of the order of seconds. When using an FTICR, ions can be observed and measured at each step in the sequential fragmentation process. However, this leads to much extended analysis times of the order of minutes per individual ion acquisition. The maximum number of steps for all of these instruments is typically seven to eight. However, during our research for the analysis of small to medium natural products such as salinomycin, it is realistically limited from three to five steps.

1.5. Collision Induced Dissociation and Higher-energy Collisional Dissociation

Due to the nature of ESI, it is a ‘soft’ technique as the ions observed are generally intact molecular type ions with little fragmentations. To obtain structural information, tandem mass spectrometry (MS/MS) and collision induced dissociation (CID) are often used to generate more ions in the gas phase.¹⁰⁷ Precursor ions of interest are selected to collide with an energised neutral gas in a collision cell. During a collision, a fraction of the ion translational energy is converted into internal energy, promoting the ions to an excited state. The ions are therefore activated to undergo decomposition, resulting in bond breakage and the production of smaller product ions. These ions can then be collected and analysed by a second mass analysis step.

Higher-energy collisional dissociation (HCD) was initially known as higher-energy C-trap dissociation. It was a technique used by Makarov and his co-workers on an LTQ Orbitrap to study peptide modifications.¹⁰⁸ In their study, the radiofrequency voltages in the C-trap were raised to retain the maximum number of product ions, and they proposed the capability of the C-trap to be used as a collision chamber to facilitate triple quadrupole-like fragmentation in the LTQ orbitrap. However, the increased radiofrequency voltage led to a decreased trapping efficiency in the low mass range. Therefore, a second iteration of HCD was developed to overcome this problem. Figures 8a and 8b provide illustrations for both HCD instruments.¹⁰⁹

Instead of having HCD in the C-trap, in the second iteration, HCD fragmentation was performed in a separate octupole collision cell situated at the far end of the C-trap. This collision cell was set to align with the C-trap and to be enclosed in a gas-tight shroud, and it was supplied with a radiofrequency voltage, and a collision gas line was attached. As shown in Figure 8b, ions pass through the C-trap and go into the HCD cell, where fragmentations occur. These ions are then returned to the C-trap and transferred into the orbitrap for mass analysis.

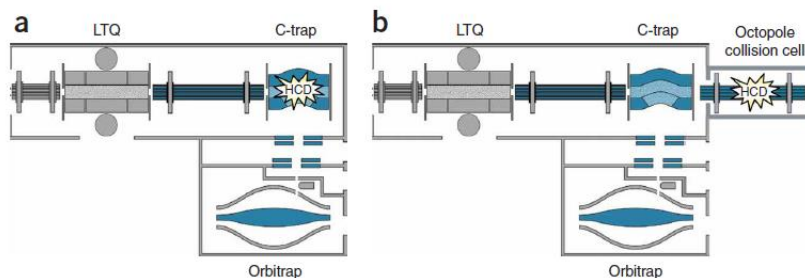


Figure 8 Schematics of the hybrid linear ion trap-orbitrap instrument with additional HCD method. Figure 8a shows that HCD occurs in the C-trap, whereas Figure 8b demonstrates that the HCD procession happens in an additional octupole collision cell.¹¹⁰

Fragmentation in HCD is believed to be very similar to that normally observed in a triple quadrupole instrument.¹⁰⁸ This will be demonstrated in this report's Results and Discussion section of amino acids by presenting the energy breakdown graphs obtained from HCD. These will then be compared with those collected from CID. One published study has compared the performance of HCD to that of CID for large-scale phosphorylation analyses of the murine brain.¹¹¹ It concluded that CID fragmentation with rapid detection in the ion trap could provide substantially richer data sets for large-scale phosphoproteomics. But for back-to-back experiments, it stated that HCD had shown higher search

engine scores with more peptides identified, thus proving its promising applicability in phosphoproteomic analysis.

1.6. Ion Mobility Spectrometry

Ion mobility spectrometry (IMS) refers to the principles, methods, and instrumentation for characterising substances from the speed of gaseous ions derived from a substance in an electric field and through a supporting gas atmosphere.¹¹² Because ion mobility spectrometry measures the mobility of gaseous ions in an electric field through a buffer gas, it is a technique capable of separating ions based on their differences in mobility.¹¹³

When ions are moving in an electric field (\vec{E}), the electric force applied to the ions is:

$$\vec{F} = q\vec{E}$$

Equation 15

where q is the charge of the ion. The higher the charge, the greater the force. Ions will accelerate in the electric field, according to the law of motion, which is:

$$\vec{F} = m\vec{a}$$

Equation 16

where \vec{a} is the acceleration. Ions accelerating in a medium filled with gas will experience collision at certain high gas pressure. The collision slows the ions down, and they act as a friction force in the opposite direction of the electric force. When the collision is frequent enough, and at the correct strength, the friction force will balance out the electric force, reaching a stationary state for the ions. The ions are not accelerating but appear constant, and this is called the drifted velocity \vec{v}_d .

The ion's mobility (K) is given as the proportionality constant between the drift velocity and the electric field:

$$\vec{v}_d = K \cdot \vec{E}$$

Equation 17

From Equation 17, ion mobility spectrometry separates ions based on the ion's mobility, K . This K value depends on the charge of an ion and its friction in the buffer gas. Larger ions typically have higher masses, therefore more friction in the buffer gas, hence lower mobility. However, IMS does not limit itself to measuring ions of different masses. Ions with equal mass and charge can have different

molecular structures. If an ion adopts a more expanded structure, for example, a long linear shape in 3D, it will have more friction and smaller ion mobility. This results in a smaller drift velocity. In comparison, ions with compact structures will experience less friction force and hence have higher mobility and greater drift velocity.

The ability to separate ions based on their 3D structures is the unique characteristic of ion mobility spectrometry. When coupled with MS, ion mobility can provide complementary and valuable information on the analytes. Ion mobility mass spectrometry (IM-MS) is particularly useful for separating isomeric compounds (analytes with the same atoms but different 3D structures) or isobaric compounds (analytes with the same mass but different composition).¹¹³

1.6.1. Travelling Wave Ion Mobility Spectrometry

Travelling wave ion mobility spectrometry (TWIMS) was first investigated and reported in a paper published in 2004 by Giles and his co-workers.¹¹⁴ In the paper, they reported a new ion mode of ion propulsion within an RF ion guide based on a stack of ring electrodes. A travelling wave is produced by passing a voltage pulse of an initial electrode and then moving the pulse along to an adjacent electrode. The ions can ‘surf’ through the stacked ring electrodes with higher speed for smaller, high-mobility ions and slower for larger, low-mobility ions. It was also confirmed that the travelling wave could effectively be a collision cell for MS/MS. Synapt HDMS (Waters Corp., Milford, USA) was made commercially available based on the published technologies, followed by Synapt G2 and G2 Si.¹¹⁵

1.6.2. Collision Cross Section

The collision cross section (CCS) (σ) is defined as: “In simple hard sphere collision theory, the area of the circle with radius equals to the sum of radii of two colliding spherical particles”.¹¹⁶ Collision cross section hence follows a simple equation of the area of a circle (Figure 9):

$$\sigma = \pi(r_i + r_g)^2$$

Equation 18

However, in IMS, atoms and molecules consist of electrons with no well-defined boundary. They cannot be assumed to be ‘hard spheres’. The probability of collision depends on their distance and their electron density polarisability. Electrons move when they approach a charge or a dipole. Polarisability is also affected by the velocity of the particles, the faster they approach, the greater polarisability the

other particle will experience. Therefore, it is difficult to define σ for atoms and molecules. Instead, the momentum transfer between the ion and the gas is considered important in IMS, referred to as the momentum transfer collision integral.

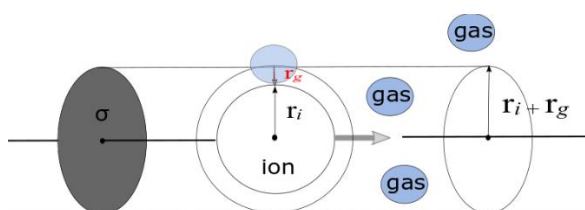


Figure 9 A reproduction of the effective section of collision (σ) between two hard spheres of radii r_i (ion) and r_g (gas). An ion is moving to the right of gas molecules in a collision cylinder. ¹¹³

Momentum transfer collision integral (Ω) is expressed in relation to the ion mobility (K) by Langevin in 1905 ¹¹⁷:

$$K = \frac{3}{16} \sqrt{\frac{2\pi}{\mu k_B T}} \frac{ze}{N\Omega}$$

Equation 19

where N is the number of gas molecules per unit volume with a unit of m^{-3} ; k_B is the Boltzmann constant. T is the gas temperature, z is the ion's charge, e is the elementary charge, and μ is the reduced mass of the ion-gas couple. Equation 19 shows the relationship between the ion mobility and the momentum transfer collision integral, although it is still commonly accepted to use the term 'collision cross section or CCS'. Once ion mobility is measured, one can determine the moment transfer collision integral Ω based on the equation above.

1.6.3. Ion Mobility Spectrometry Applications

Ion mobility spectrometry (IMS) has a wide range of applications, from drug and explosive detections,¹¹⁸ amino acid differentiation,¹¹⁹ metabolites studies¹²⁰ to protein analysis.¹²¹ The fact that it can separate ions based on its structure makes it advantageous in studying natural products, for example, polyether ionophores. In the case of salinomycin, it isomerises in solution and produces multiple fragmentation pathways of isobaric and isomeric product ions. A review paper published in 2021 gave a detailed examination into the application of IMS for natural product studies since 2015.¹²² It summarised existing IMS platforms, displayed a workflow on how to identify analytes using both m/z and CCS data. Most remarkably, it compiled a list of IMS studies in natural product research performed from 2015 to 2020. However, it has not mentioned any analysis on polyether ionophores.

1.7. Computational Methods

1.7.1. The Development of Gaussian

The Gaussian software is a computational software commonly used by all levels of researchers to calculate optimised molecular structures and relative abundances of conformations, investigate large systems, model spectroscopy, and study excited state chemistry.¹²³⁻¹²⁸ Its extensive usage makes it one of the most popular computational software for scientific studies. Gaussian was first released in 1970 by John Pople and his research group at Carnegie Mellon University.¹²⁹ The current version of Gaussian is Gaussian 16. However, when this research was initially conducted, Gaussian 09 was used.¹³⁰ It was used mainly to optimise 3D structures of molecular and product ions, simulate molecular orbitals, and calculate proton affinity and the total energy and frequency of the input structure. Data will be established in later chapters and subsequently discussed.

1.7.2. Hartree-Fock and Density Functional Theory

The Hartree-Fock method (HF), also known as the self-consistent field method (SCF), is an approximation for the solution to the Schrödinger equation. It determines an approximate wave function and the energy of many electron systems.¹³¹ The Hartree-Fock equations describe non-interacting electrons under the influence of a mean-field potential consisting of the classical Coulomb potential and a non-local exchange potential.¹³² One of the main assumptions of the HF method is that individual electrons treat others as an average field. The problem with this is that it neglects electron correlation - the energy arises from interactions between electrons. It is unsuitable for accurately modelling reaction energies and bond dissociation energies.¹³³ Therefore, although the HF method plays an important role in quantum chemistry, other methods have been developed to provide more accurate approximations.

Based on the fundamental work of Thomas and Fermi in 1927,^{134,135} the Thomas-Fermi equations were believed to showcase the first effort in defining a density functional theory (DFT). Combined with Slater's approximation on exchange energy,¹³⁶ DFT has been widely used in solid state physics. However, large errors in molecular calculations resulted in DFT being unpopular in chemistry.¹³⁷ Interestingly, this perception was changed after Hohenberg and Kohn successfully proved two theorems, demonstrating the credibility of DFT as a quantum chemical methodology.¹³⁸

The DFT method is based on the strategy of modelling electron correlations via general functionals (a function of a function) of the electron density. It can be demonstrated by several energy components which add up to form electronic energy (see Equation 20).¹³⁹

$$E = E^T + E^V + E^J + E^{XC}$$

Equation 20

where E^T is the kinetic energy term arising from the motion of the electrons, E^V describes the potential energy of the nucleus-electron attraction and the repulsion between pairs of nuclei, E^J is the Coulomb self-interaction of the electron density, also described as electron-electron repulsion. $E^T + E^V + E^J$ together are the classical energy of the charge distribution. E^{XC} is the exchange-correlation representing the rest part of the electron-electron interactions. It includes the exchange energy from the antisymmetry of the quantum mechanical wave function, and the dynamic correlation in the motions of the individual electrons. Hohenberg and Kohn established that E^{XC} is a functional of the electron density ρ . It can be expressed by equation 21:

$$E^{XC}(\rho) = E^X(\rho) + E^C(\rho)$$

Equation 21

Here $E^X(\rho)$ is the exchange functionals and $E^C(\rho)$ is the correlation functionals. Both terms can be calculated by local methods, which depend only on electron density ρ (i.e. local density approximation LDA), or gradient-corrected methods, which depend on both ρ and its gradient (i.e. generalised gradient approximation GGA). Pure DFT methods are a combination of an exchange functional and a correlation functional.¹³⁹ As an example, the popular BLYP theory set consists of Becke's gradient-corrected exchange functional with Lee, Yang and Parr's gradient-corrected correlation functional (LYP).^{140,141}

1.7.3. Hybrid Functionals

Hybrid functionals are based on the idea of mixing exchange energies calculated from Hartree-Fock and DFT, with DFT correlations to improve calculation accuracy. One of the most employed B3LYP exchange-correlation functionals is based on a Becke-style three-parameter functional¹⁴² and gradient-corrected correlation functional by LYP. It has an expression as follows in Gaussian 09 software (see Equation 22):

$$E_{B3LYP}^{XC} = 0.2E_{HF}^X + 0.8E_{LDA}^X + 0.72\Delta E_{B88}^X + 0.81E_{LYP}^C + 0.19E_{VWN3}^C$$

Equation 22

In this expression, E_{HF}^X represents the exchange energy calculated by HF. E_{LDA}^X is the exchange energy arising from local density approximation. E_{B88}^X is Becke's gradient correction to LDA local exchange energy.¹⁴³ E_{LYP}^C is gradient-corrected correlation functional by LYP. And E_{VWN3}^C is functional III described by Vosko, Wilk and Nusiar,¹⁴⁴ it provides the excess local correlation required in the calculation.

Hybrid functionals are popular choices in modern research. In this study, both B3LYP and B3PW91 were initially employed, but B3PW91 was selected as the final functional choice. B3PW91 uses the PW91 (Perdew-Wang 91) correlation function developed by Perdew and Wang¹⁴⁵ instead of LYP. But it keeps Becke's three parameters functional for the exchange energy.

Multiple previous studies have concluded that B3PW91 gives better results (closer to experimental data) when compared to B3LYP. A comparison made on the atomisation energy of carbon dioxide using B3LYP, B3PW91, and other DFT functionals stated B3PW91 values were in excellent agreement with experimental observations.¹³³ Another study concluded that while B3LYP failed to locate an intermediate of the Smiles rearrangement (intramolecular nucleophilic aromatic substitution), B3PW91 succeeded.¹⁴⁶ It was also stated in a paper comparing DFT methods for studies of amino acids, that although the performance of B3LYP and B3PW91 on describing different side chains and carboxyl orientations were similar, B3PW91 was the best performer for proton affinity calculations.¹⁴⁷ In addition, it was reported that B3LYP gave unsatisfactory results in overall error analysis. Therefore, B3PW91 was chosen to be the main method of computational modelling for amino acids and flavonoids in this study. In the case of salinomycin, due to its relatively larger structure and complexation with a metal ion, the computational cost was too high to use the same setting as amino acids and flavonoids. Instead, basic B3LYP with fewer basis sets was used.

1.7.4. Basis Sets

A basis set is a mathematical representation of molecular orbitals within a molecule. It restricts each electron to a specific region of space and assigns a group of basis functions to each atom within a molecule to approximate its orbitals. Larger basis sets are believed to produce more accurate molecular orbital approximations as they impose fewer restrictions on the location of the electrons in space. For standard basis sets in electronic structure calculations, linear combinations of Gaussian functions are used to construct the orbitals.¹⁴⁸

A minimal basis set contains the minimum number of basis functions required for each atom. In Gaussian 09 software, the STO-3G is a minimal basis set, it uses three Gaussian primitives (3G) per basis function, and STO stands for 'Slater-type orbitals'. The nomenclature of some widely used split-valence basis sets (3-21G, 6-31G, and 6-311G... etc.) is a guide to the contraction scheme. The first number is the number of primitives used in the contracted core functions. Numbers after the hyphen demonstrate the number of primitives included in the valence function. The more numbers mean there are more basis functions per atom, this increases the size of a basis set, which is known as a split valence basis set. For example, 6-311G represents a triple split valence basis set, and it adopts three sizes of contracted functions for each orbital type.¹⁴⁹

Polarised basis sets add orbitals with angular momentum beyond what is needed for the ground state to describe each atom. It can add d functions to carbon atoms, p functions to hydrogen atoms, and f functions to transition metals. For example, polarised basis set 6-311G(d,p) describes additional d functions on heavy atoms and adding p functions to hydrogen atoms. Sometimes 6-311G(d,p) can also be written as 6-311G**.

Another commonly used method to increase the basis set size is to add diffuse functions. These are larger versions of s- and p-type functions. Diffuse functions allow orbitals to occupy a larger region of space. They are important in systems where the electrons are relatively far from the nucleus.¹⁴⁸ In Gaussian 09 software, they are represented by the '+' symbol. For instance, 6-311+G(d,p) adds diffuse functions to the heavy atoms, and the double plus in 6-311++G(d,p) means additional diffuse functions to the hydrogen atoms and the heavy atoms.

The basis set chosen for this study for amino acids and flavonoids is 6-311++G(d,p). It has six primitives in the contracted core functions, three sets of valence region functions, diffuse function on both heavy atoms and hydrogens, and multiple polarisation functions: d functions on heavy atoms and p functions on hydrogen atoms. It was chosen because of relative references and their satisfactory performances.^{147,150} The basis set for salinomycin and its metal ion complexes is 6-31G based on a balance between computational costs and data satisfaction.

1.7.5. IMoS

During the study of salinomycin and narasin, we discovered that both readily isomerise in a water-methanol solution after about two weeks. Their tendency to isomerise agrees with previous literature.¹⁵¹ In addition, their fragmentation pathway produces multiple product ions, which are structural isomers

of each other. Some of the structural isomers had never been published before. These intriguing observations led to an interest in calculating the ion mobility and collision cross section of the targeted structural isomers. The results are then compared with our experimental observation and available references.

A limited number of calculators were available to determine ion mobility and CCS. The most commonly used are IMoS¹⁵²⁻¹⁵⁶ and MOBICAL^{157, 158}. In an analysis of 47 organic ions, both calculators gave excellent agreement in He and N₂.¹⁵⁹ However, IMoS has lower computational costs and has a Windows user-ready application. Hence, it was chosen to be the calculator suitable for this research. IMoS stands for Ion Mobility Spectrometry Suite, and it is a set of parallelised tools that can be used to infer the CCS and mobilities of gas-phase charged molecules. The authors of this software have produced detailed explanations and video tutorials on their website¹⁶⁰ to assist in understanding what the software is calculating. The software calculates the ion mobility and CCS based on Equation 19. It can read output files from Gaussian and also takes into account different buffer gas in the spectrometer.

1.8. Natural Products

A natural product is defined as a natural compound or substance produced in nature by living organisms.¹⁶¹ In organic chemistry, natural products are organic compounds produced by primary and secondary metabolic pathways. These are called primary and secondary metabolites. Primary metabolites are vital to the survival of the organisms, whereas secondary metabolites are not. However, secondary metabolites are usually needed for evolutionary advantages.¹⁶² In this research, three types of natural products were selected – amino acids, flavonoids, and polyether ionophores. The first group is a primary metabolite, and the other two are/or extracted from secondary metabolites. These were chosen due to the following considerations:

- Structural Similarity

The natural products chosen for this study have basic backbones consisting of cyclic, phenyl, or heterocyclic rings and carboxylic terminals. The degree of complexity in structure can be shown as amino acids < flavonoids < polyether ionophores. This unique property allows this research to be conducted like a step-ladder, starting with the relatively simplest alpha amino acid group and progressing to the more complicated polyether ionophores.

- Isomers

Alpha amino acids and flavonoids have various structural isomers with the same mass but different functional groups. Salinomycin and narasin, the polyether ionophores chosen in this research, also produce multiple structural isomers during fragmentation. In addition, salinomycin and narasin also isomerise in solution. What is more, when complexed with metal ions, salinomycin chelates with the metal to produce unique **ionofoms** (charged metal-molecule ions with different geometry depending on the location of the metal chelation). These characteristics of the chosen natural products provide a solid foundation to develop and test methodologies in this research.

- Existing Publications

Besides salinomycin and narasin, amino acids and flavonoids have been extensively studied in previous literature by mass spectrometry. These will be discussed in the later chapters of this thesis. The existing publications help this research compare and validate the results produced. Mass spectrometry studies of salinomycin and narasin are limited. Therefore, it is useful to present data obtained from this research.

- Commercial Availability

All three types of natural products selected were easily accessible through commercial suppliers. This allows the research to be conducted cost-effectively and increases the reproductivity of experiments.

1.9. Aims and Objectives

Over the past few decades, studies of natural products by ESI-MS/MS have attracted a vast amount of interest due to the instrumentation advancement and their robust analytical power. Understanding the fragmentation reactions is crucial to elucidate the structure and identify natural products successfully. There are still challenges in this area due to the lack of systematic analysis of the gas-phase fragmentation reactions by ESI-MS/MS.¹⁶³ Isobaric and isomeric peaks in the mass spectra can pose difficulty in correctly assigning the chemical identities of the product ions.¹⁶⁴ Extensive presence of structural isomers within the same class of natural products also exhibit a challenge in their identification and quantification.¹⁶⁵⁻¹⁶⁷ Isomerism plays a vital role in drug development.¹⁶⁸ Being able to differentiate isomers confidently is crucial for the safety and efficacy of any pharmaceutical treatment.

In addition, ESI-MS/MS has no reliable fragmentation libraries like those present for GC-EI-MS. So an increased understanding is needed if we require a reliable/usable library for known/unknown identification by LC-ESI-MS/MS.

Therefore, this research has four main objectives:

1. Systematically compile gas-phase fragmentation reactions of the three groups of natural products to increase our understanding of the way that molecules fragment.
2. Use high-end ultra-high resolution accurate-mass techniques to perform the experiments in order to obtain reliable data and accurate results. Apply MSⁿ to extend the range of product ions observed to lower masses.
3. Improve methodologies for identifying and differentiating structural isomers. Especially combining routine mass spectrometry methods with IMS when separating isomers and isobars.
4. Investigate novel or misannotated product ions of selected natural products by combining mass spectrometry and computational methods.

2. Amino Acids

2.1. Amino Acids Overview

Amino acids are organic compounds that consist of a carboxyl group, amine, and a side chain. The side chain is usually represented as an 'R' group and is specific to each amino acid. By 1983, more than 500 naturally occurring amino acids had been discovered through, for example, a search for new antibiotics.¹⁶⁹ In this initial part of the study, the focus will be on the 20 'protein building' or proteinogenic amino acids, i.e., the amino acids found in natural proteins. These amino acids combine to form peptide chains, which then become the fundamental building blocks of proteins. These are all α -amino acids, and all are L-stereoisomers. The structure of a general α -amino acid is represented in Figure 10. According to the Cahn-Ingold-Prelog sequence rules, all L- α -amino acids have a configuration of 2S on the carbon stereocentre, except L-(2R)-cysteine. This leads to minor confusion, allowing the L-D terminology to survive.¹⁷⁰ L stands for laevorotatory, and it is measured by the ability to rotate plane-polarised light, with reference to L-(-)-glyceraldehyde.¹⁷¹

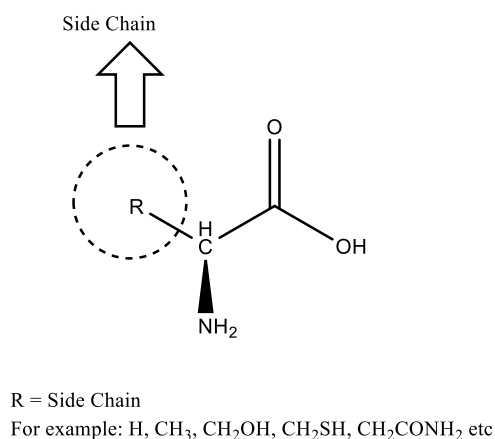


Figure 10 General structure of L- α -amino acid. 'R' is the side chain and it has different functional groups depends on the specific amino acids.

For the convenience of this study, the 20 α -amino acids have been categorised into 7 groups by their side chain functionalities. These are :

- Aliphatic (alanine, glycine, isoleucine, leucine, proline and valine)
- Acidic (aspartic acid and glutamic acid)
- Hydroxylic (serine and threonine)
- Basic (arginine, histidine and lysine)
- Amidic (asparagine and glutamine)
- Aromatic (phenylalanine, tryptophan and tyrosine)
- Sulfur containing(cysteine and methionine)

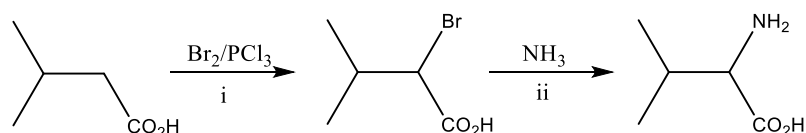
General information useful for this study, including abbreviation (3-letter and 1-letter), formula, molecular structure, residue mass, exact mass, protonated mass, sodiated mass and deprotonated mass, are summarised in Appendix: 2. Lists of Analytes. A colour-coded scheme has been used to distinguish different categories of these α -amino acids.

One of the most important biological significances of α -amino acids is their role in nutrition. Nine are essential to humans because they cannot be synthesised from other compounds by the human body and must be ingested through food. These are histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine. In addition, some α -amino acids are widely used as food additives and flavour enhancers (such as glutamic acid and its monosodium salt). Taking amino acid supplements is unnecessary for people with an adequate and balanced diet¹⁷², and toxicity can arise even from the essential amino acids. In 1989-1990, an outbreak of eosinophilia myalgia syndrome (EMS) resulted from the use of tryptophan as a dietary supplement, affecting more than 1500 people in the USA. Even though later, it was reported that the problem resulted from contamination during the production of the amino acids.¹⁷³

Other applications of α -amino acids include the addition of methionine in paracetamol products¹⁷² due to its ability to counteract the serious side effect of paracetamol overdose. Methionine helps to restore glutathione levels in the body, which can react with oxidised paracetamol to produce a stable adduct. Some α -amino acids are also useful in archaeology and forensic science by measuring the degree of racemisation.^{174, 175} It can also give useful information on the age of art specimens and the dating of oil paintings.

In neuroscience, glutamate (the anion of glutamic acid) is an important neurotransmitter used for excitatory function in the vertebrate brain.¹⁷⁶ The N-acetyl derivate of aspartic acid has been reported as a putative marker of neurones. N-methyl-D-aspartate has attracted much interest for its role in treating Alzheimer's disease.¹⁷⁷

Early discovery and isolation of the proteinogenic α -amino acids from natural sources have been covered in exceptional detail in the literature.¹⁷⁸ Therefore, it will not be explored in detail here. The most traditional way to obtain L- α -amino acids is by hydrolysis of proteins and separation of the resulting mixture. However, as the demand for these amino acids increases and with the advancement of technology, various methods of industrial synthesis are now being used, especially fermentation methods and chemical synthesis.¹⁷⁹ The target is to achieve a homochiral product, so resolution must be performed if the target is assembled without regard to α -carbon chirality. Otherwise, an asymmetric synthesis can be employed, or an enantiospecific conversion of a homochiral compound to the required α -amino acids must be performed. These procedures have been outlined and discussed in the literature, and Scheme 1 is one example of the general synthetic methods to produce α -amino acids.¹⁷⁹ Chiral resolution is required to separate the resulting diastereoisomers by either derivatising with an optically active reagent or enzymes.



Scheme 1 Displacement reactions on α -halo acids, an example of the synthetic methods to produce α -amino acids.

2.1.1. Amino Acids Analysis by Mass Spectrometry

Because of the biological significance of amino acids, there is an extensive body of literature on their analysis by MS. Different ionisation methods and computational calculations have been explored extensively regarding the detection and fragmentation of amino acids. In this part of the introduction, previous studies are examined and used as a comparison to results obtained from this research.

A 1972 study showed the dependence of the relative abundances of product ions from collisional decomposition in the gas phase on the ionic structure of small organic compounds.¹⁸⁰ In the same year, another study using electron-impact induced decomposition showed a variation of fragmentations with the ring sizes in cyclic α -amino acids.¹⁸¹ Although the loss of the carboxyl radical formed the most significant product ion, ring opening fragmentation was also observed depending on the ring size. A comprehensive study of 14 α -amino acids by H₂ CI was carried out in 1976.¹⁸² It compared the results to those produced by methane CI, showing that the loss of carboxylic acid and ammonia are the most abundant, and demonstrated the observation of extensive intramolecular proton transfer in the protonated molecular ion prior to fragmentation. Most importantly, it also stated that leucine has no loss of NH₃ (-17) from product ion $[(M+H) - HCOOH]^+$ but a loss of C₃H₆, which gives product ion m/z 44, unlike isoleucine which has a significant peak corresponding to $[(M+H) - HCOOH - NH_3]^+$. This observation agrees with our results and is discussed later on pages 63 and 103-104.

FAB-MS has been extensively employed for investigating amino acids and peptide sequencing. It is useful in protein chemistry combined with MS/MS and complements the results obtained from Edman degradation for the amino acids sequencing of peptides.¹⁸³ In the positive ion FAB-MS of 24 α -amino acids, the production of the immonium ions was the most common observation. In addition, isomeric amino acids such as leucine and isoleucine could be identified unambiguously from their collisional activated spectra.¹⁸⁴ A later study examined low-energy CID reactions of α -amino acids by FAB-MS as a function of collision energy and energy breakdown graphs. The energy dependence of the fragmentation reactions is established for various protonated amino acids.¹⁸⁵ It has been stated that the fragmentation reactions of these α -amino acids strongly depend on the identity of the R group in the side chain. This agrees with the results presented in this thesis.

Other ionisation methods have also been investigated for the study of α -amino acids. A 1992 study demonstrated that PD-MS was suitable for studying natural α -amino acids.¹⁸⁶ In this study, the formation of the immonium ion was observed to be the most abundant, and multiple protonation sites were proposed for amino acids with multifunctional side chains. Characteristic secondary ions were also observed, allowing the identification of amino acids with different side chain functional groups. Leucine and isoleucine were also shown to produce different product ions in PD-MS.

ESI-MS is one of the most recognised and powerful tools in analytical chemistry, and as a result more research has been done on small molecules such as amino acids than any other MS based technique. A study in 2003 analysed the fragmentation of 79 biologically significant molecules, including all 20 protogenic α -amino acids with ESI-MS/MS in positive and negative ionisation modes.¹⁸⁷ Tables of the ionisation and fragmentation results are presented, which can be used as a comparison to our research. A few years later, another study was published investigating the fragmentation of protonated arginine, lysine, and their methylated derivatives by ESI-MS/MS.¹⁸⁸ It reported the competitive losses of the side chain amino groups and similarities between the fragmentation chemistries of protonated arginine, its methyl derivatives, and protonated methylated lysines. Also, a compilation of the fragmentation data of single amino acids was published in 2011 using ESI-MS.¹⁸⁹ It stated that the ESI-MS/MS spectra resembled those obtained by CI-MS, with loss of formic acid as the most predominant fragmentation, followed by losses of NH_3 and H_2O .

Although most studies on amino acids were performed in positive ion mode, deprotonated amino acid fragmentation has also been established.¹⁹⁰ For example, APCI-MS/MS was used to generate the spectra of 20 α -amino acids ions. The three most common product ions observed were $[\text{M}-\text{H}-\text{NH}_3]^-$, $[\text{M}-\text{H}-\text{H}_2\text{O}]^-$, and $[\text{M}-\text{H}-\text{CO}_2]^-$.

During the literature review, many papers have combined theoretical data from computational modelling with experimental results obtained from mass spectra.^{188,191-194} Gaussian software is found to be one of the most popular computational chemistry software in use. One paper examined the various conformations of protonated amino acids and the fragmentation of amino acid radical cations.¹⁹⁴ It concluded that ionisation is mainly localised on the $-\text{NH}_2$ group and loss of COOH^\bullet is the most favourable. However, for aromatic amino acids, ionisation can occur on the side chain, with fragmentation leading to the cleavage of the side chain and competition with the loss of the carboxyl group.

2.1.2. Aims and Objectives

From all the previous studies mentioned in the literature review section, it is clear that α -amino acids have been investigated extensively by MS and computational methods. Results obtained by others agreed with most of the observations in this research. There is a strong connection between fragmentation pathways and side chain functionalities of these amino acids. However, a systematic study on the fragmentation of proteinogenic amino acids has not been performed using the same methodology nor cross-referenced with theoretical data from computational modelling.

Therefore, this part of the thesis aims to establish a systematic compilation of the fragmentation reactions of all 20 proteinogenic α -amino acids. The results are successfully used to differentiate the various structural isomers of amino acids¹⁶⁵ and present a correlation between their fragmentation and side chain functionality. Mechanisms are presented to better understand the fragmentation pathways of selected amino acids. Our results also demonstrated the differences between HCD and CID modes for MS/MS to analyse aromatic amino acids.

2.2. Experimental

2.2.1. Amino Acids Samples

All 20 L- α -amino acids were obtained from Sigma Aldrich (purity \geq 98%).

Additional analytes were used to expand the study. Sources of these samples were listed in brackets. All of them have a purity of more than 98%. L-allo-isoleucine (Sigma Aldrich), Fmoc-lysine(Alloc)-OH (Sigma Aldrich), Fmoc-tyrosine(tBu)-OH (AGTC Bioproduct). All analytes were dissolved in 1:1 water: methanol to make up a standard solution of 1 mg/mL. All solvents were HPLC gradient grade obtained from Fisher Scientific.

2.2.2. Instrumentation

Positive and negative ESI-MS analyses were performed on an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific) using the heated ESI source. The analyte solutions were diluted to 0.1 mg/mL in Water:Methanol (1:1) solvent prior to analysis and were delivered using the autosampler module of an RS3000 UHPLC system (Dionex) at a 10 μ L/min flow rate. Tandem mass spectra were recorded in CID-MS/MS mode on isolated precursor ions (100-200 m/z window) with collision energies ramped from 0 to 50 eV. Energies were increased by 1 eV every 12 seconds. The resulting runs were 10 minutes per sample. A full scan (50 – 250 m/z) was obtained, and then the selected precursor ion was automatically selected in data-dependent mode, isolated, and underwent CID-MS/MS using dry N₂ as the collision gas. A few of the amino acids had to be recorded manually. In this case, a syringe pump was used to directly infuse the analyte solution into the Orbitrap. When additional studies (e.g., amino acids with protecting groups) were performed, the settings on orbitrap were slightly adjusted (e.g., expanded mass range) to give the best quality spectra.

For all experiments, the acquisition time was 0.25 mins per scan, with a flow rate of 10 μ L/min, mass analyser FTMS mode at a resolution of 240,000, and 2 microscans were summed with a maximum ion accumulation time of 200 ms.

2.2.3. Computational

Gaussian 09 software¹⁹⁵ was used to perform all optimisation, frequency, and Intrinsic Reaction Coordinate (IRC) calculations. All structures were optimised to give a stationary point (minima) on the potential energy surface. Frequency calculations were performed to obtain zero-point energy (a correction to the electronic energy of the molecule, taking in the effects of molecular vibrations which persist even at 0 K) and to ensure a structure is a minimum with no imaginary frequency, and a structure is a true transition state when there is only one imaginary frequency. IRC calculations were used to verify a transition state does connect the two minima assumed. In other words, it confirms the proposed transition state connects to its reactant and product. Density functional theory (DFT) was used, with the hybrid exchange-correlation functional and basis set at B3PW91 6-311++G(d,p).

2.3. Results and Discussion

2.3.1. General Trends

2.3.1.1. Positive Ion Mode

Data collected from the positive ion ESI-MS/MS of all 20 amino acids are presented below (Table 2). Common neutral losses observed are summarised to include loss of HCOOH, NH₃, H₂O, CO, and HCONH₂. Product ions in purple are observed in HCD mode. Alanine and glycine are highlighted in red, their lower molecular weights take them outside of the Orbitrap's *m/z* range, and as a result, their MS/MS could not be measured. Losses of formic acid (HCOOH) are given as reasonable assumptions according to product ions of other amino acids in the same aliphatic group and from which is seen in the literature.¹⁹⁶

From the table, the first trend is that the loss of HCOOH is predominant and non-discriminating. 18 out of the 20 amino acids (including predictions for alanine and glycine, and tryptophan, which is observed in HCD) lose HCOOH directly from their precursor ions [M+H]⁺*. The other two amino acids, arginine and lysine, belong to the basic amino acid group with side chain pKa above 10.^{197,198} Interestingly, for both of these amino acids, the loss of HCOOH was only observed after the initial loss of NH₃. This observation could be the first confirmation of the hypothesis established at the beginning of this research.

Other trends to be highlighted are that loss of NH₃ is the second most frequent loss, followed by losses of H₂O, HCONH₂, and CO. To summarise, the general trend for neutral losses in positive ion CID-MS/MS analyses of the 20 amino acids is loss of HCOOH > NH₃ > H₂O > HCONH₂ > CO. All neutral losses recorded here occur directly from the protonated precursor ion. Other product ions are grouped into the 'Other Losses' section. By doing this, it is more obvious that some neutral losses only occur after an initial loss of another neutral. For example, serine and threonine lose two water molecules and formic acid. However, the second water loss only occurs after the initial water loss. Another example is cysteine, which has a product ion at

* M stands for neutral amino acid, [M+H]⁺ is the protonated molecular ion.

m/z 87, which results from a loss of H_2O 18 from m/z 105, and the product ion m/z 105 is produced from an initial loss of NH_3 from its precursor ion.

Detailed fragmentation pathways will be discussed later, but it is apparent from this summary table that the side chains have a significant effect on the characteristics of the MS/MS spectra. This is, therefore, a step forward in proving the hypothesis. It is also worth pointing out that HCD generated more product ions, especially with the aromatic amino acids tryptophan and tyrosine. This observation can be explained by HCD being a more focused technique than CID. Secondly, both aromatic amino acids have higher molecular weights and better stabilisation, enabling them to delocalise the CID energy more effectively and thus making them less likely to fragment.

Neutral Losses of 20 α -amino acids in Positive Ion Mode										
Name	Precursor Ion [M+H] ⁺	NH ₃ (17)	H ₂ O (18)	CO (28)	HCONH ₂ (45)	HCOOH (46)	Other Losses	Suggested Fragmentation Pathway for Other Losses		
Alanine	90					44				
Arginine	175	158	157		130		116, 112	116 (175-59)	112 (158-46)	
Asparagine	133	116		105	88	87	74	74 (133-59)		
Aspartic Acid	134		116			88	74	74 (116-42)		
Cysteine	122	105				76	87, 59	87 (105-18)	59 (76-17)	
Glutamic Acid	148		130			102	84	84 (102-18)		
Glutamine	147	130			102	101	84	84 (130-46)		
Glycine	76					30				
Histidine	156					110	95	95 (156-61)		
Isoleucine	132					86	69	69 (86-17)		
Leucine	132					86				
Lysine	147	130	129				84	84 (130-46)		
Methionine	150	133				104	102, 61, 56	102 (150-48)	61 (133-72)	56 (104-48)
Phenylalanine	166	149				120	131, 103	131 (149-18)	103 (120-17)	
Proline	116					70				
Serine	106		88			60	70	70 (88-18)		
Threonine	120		102			74	84	84 (102-18)		
Tryptophan*	205	188				159	170, 146, 144, 142, 132, 118	170/146/144/142/118 (188)	132 (205-73)	
Tyrosine	182	165				136	147, 123, 119	147 (165-18)		
Valine	118					72				

Table 2 Compilation of all neutral losses of the 20 α -amino acids in the positive ion mode. * MS³ of m/z 188 for tryptophan gives product ions: m/z 170, m/z 146, m/z 144, m/z 142 and m/z 118. Red: alanine and glycine are too small to be studied by the orbitrap. Product ions in purple are observed in HCD mode.

2.3.1.2. Negative Ion Mode

Table 3 summarises the neutral losses observed from the negative ion mode ESI-MS/MS of all 20 amino acids. Again, alanine and glycine (highlighted in red) fall outside the Orbitrap mass range, so no data was collected for these two amino acids. All precursor ions were the deprotonated molecular ion $[M-H]^-$.

Unlike in the positive ion mode, in the negative ion mode, the loss of NH_3 is the most abundant loss observed and is therefore considered to be non-discriminating. The only exceptions are alanine, glycine, glutamic acid, methionine, and proline, which did not show product ions corresponding to the loss of NH_3 . The second most frequent loss is the loss of H_2O , followed by losses of CO and CO_2 . Losses of $HCOOH$, H_2N_2C , and $HCONH_2$ are also observed. The trend can be illustrated as $NH_3 > H_2O > CO/CO_2 > HCOOH > H_2N_2C/HCONH_2$. All of these are due to initial losses from the $[M-H]^-$ precursor ion of the amino acids. Secondary losses, such as an additional loss of H_2O , have also been observed, and these are grouped underneath ‘Other Losses’ in the table.

Most of the amino acids lose NH_3 and H_2O during their fragmentation. However, when you take the side chain groups into account as well, the variation is more significant compared to that observed in positive ion CID. There is less uniformity within the same amino acid group. For example, cysteine and methionine have completely different fragmentation routes. Cysteine loses NH_3 , H_2O , and CO , whereas only one 48 mass unit loss is observed with methionine. Despite the fact they both contain sulfur on their side chains. In addition, the acidic amino acids aspartic acid and glutamic acid do not have the same primary fragmentations either. Aspartic acid loses NH_3 , H_2O , and CO , but glutamic acid loses only H_2O and CO .

Apart from this, there are still some clear trends in the fragmentation of the amino acids that belong to the same group. For example, hydroxylic amino acids serine and threonine have the same fragmentation. Aromatic amino acids tryptophan and tyrosine also produce identical product ion peaks in their spectra. Therefore, it is reasonable to suggest that in negative mode, side chain functionalities of these 20 amino acids still influence their respective fragmentation pathways. A detailed comparison will be made later in this chapter, along with proposed mechanisms.

Neutral Losses of 20 α -amino acids in Negative Ion Mode										
Name	Precursor Ion [M-H] ⁻	NH ₃ (17)	H ₂ O (18)	CO (28)	H ₂ N ₂ C (42)	CO ₂ (44)	HCO NH ₂ (45)	HCOOH (46)	Other Losses	Suggested Fragmentation Pathway
Alanine	88									
Arginine	173	156			131					
Asparagine	131	114	113						95	95(113-18)
Aspartic Acid	132	115	114			88			71	(115-44)
Cysteine	120	103	102	92						
Glutamic Acid	146		128			102				
Glutamine	145	128	127						109	109(127-18)
Glycine	74									
Histidine	154	137	136				109		93	93(137-44)
Isoleucine	130	113								
Leucine	130	113						84		
Lysine	145	128	127					99	97, 83	97(145-48) 83(128-45)
Methionine	148								100	100(148-48)
Phenylalanine	164	147				103				
Proline	114			86						
Serine	104	87	86	76						
Threonine	118	101	100	90						
Tryptophan	203	186				159			142,116	142(186-44) 116(186-70)
Tyrosine	180	163				136			119,93	119(163-44) 93(163-70)
Valine	116	99		88						

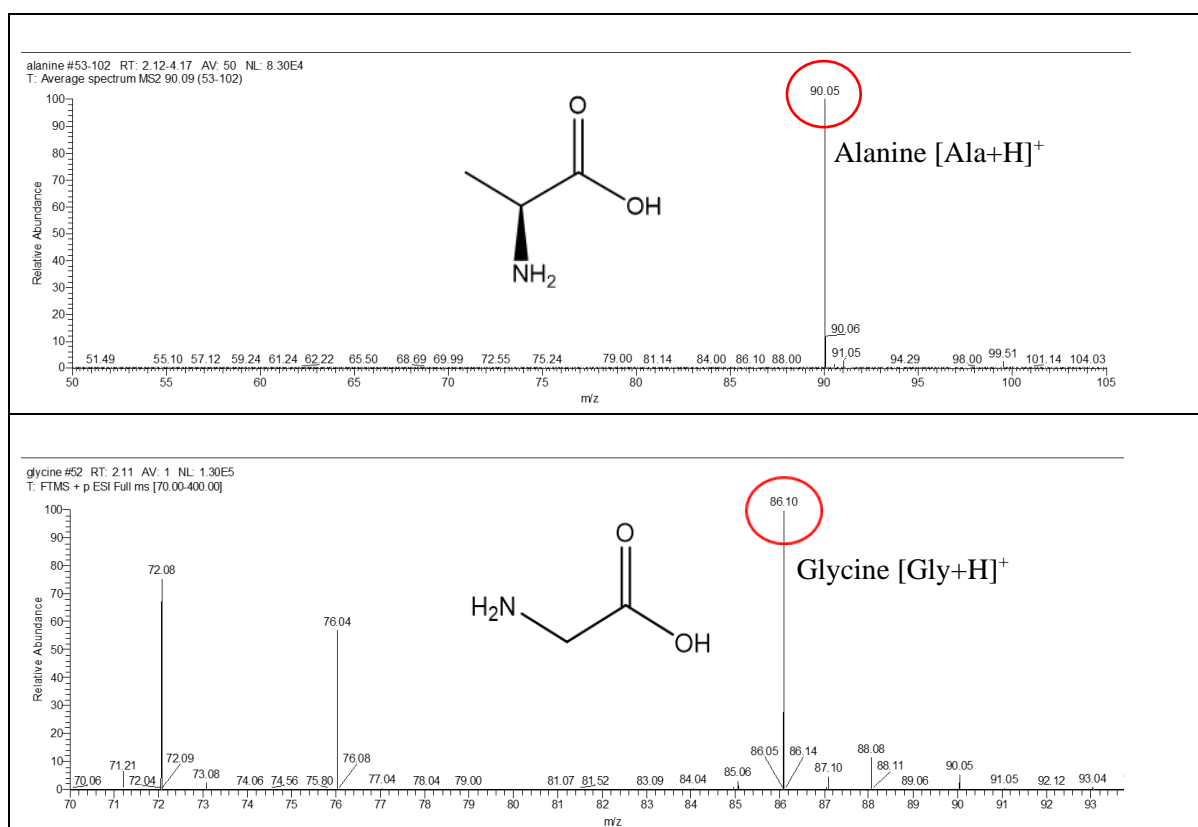
Table 3 Compilation of all neutral losses of the 20 α -amino acids in the negative ion mode. Red: alanine and glycine are too small to be studied within the orbitrap mass range.

2.3.2. Tandem Mass Spectrometry Analysis

2.3.2.1. Positive Ion Mode CID

i. Aliphatic Group +ve CID MS/MS

Amino acids included in the aliphatic group are alanine, glycine, isoleucine, leucine, proline, and valine. Among these, isoleucine and leucine are structural isomers with different positions of the methyl group on their side chains. Proline has a pyrrolidine side chain. The first interest would be to examine the mass spectra of these six amino acids to search for any similarities and characteristic differences.



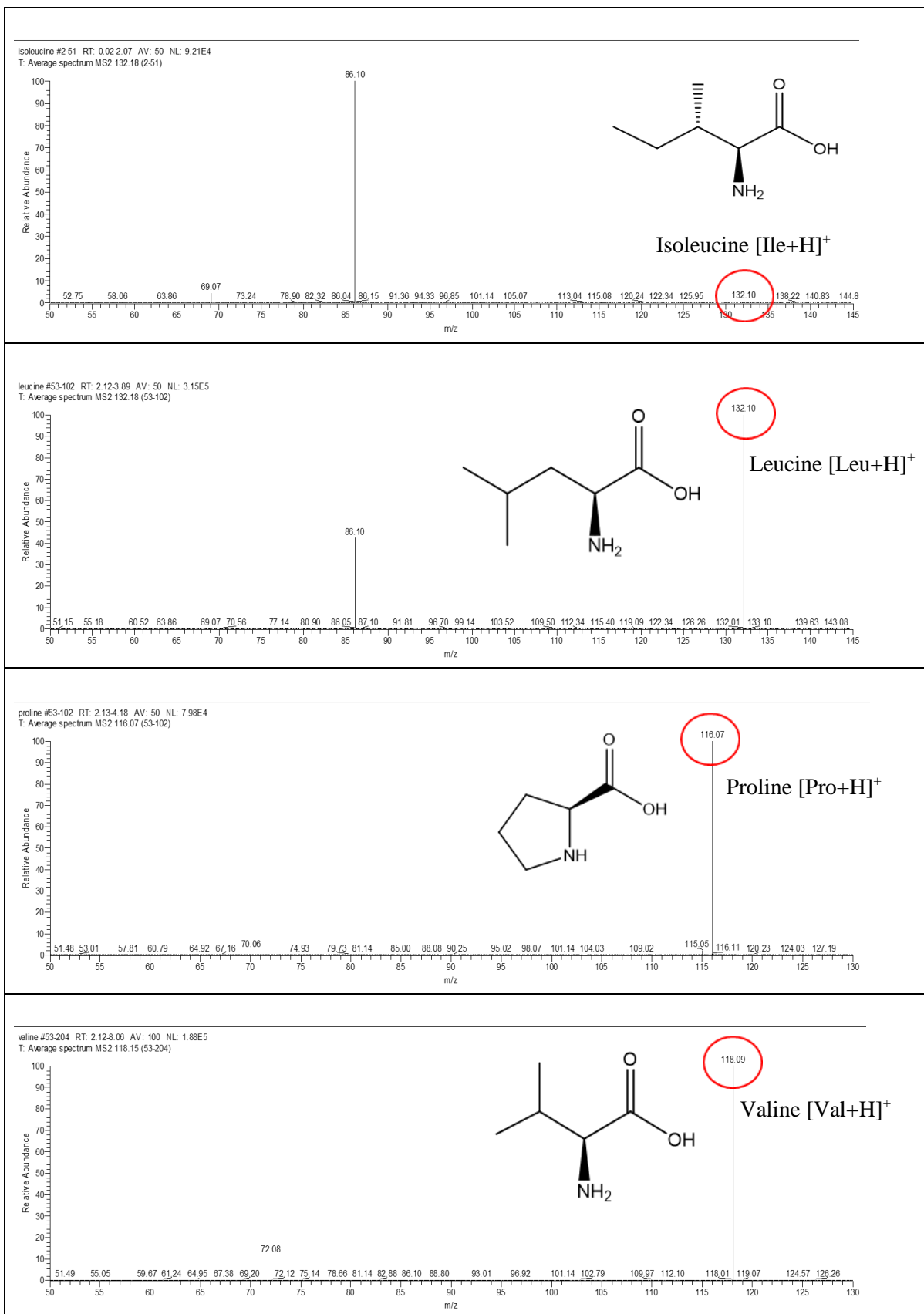


Figure 11 MS/MS spectra of the 6 aliphatic amino acids. The protonated molecular ions, $[M+H]^+$, were the precursor ions, and they are labelled with a red circle.

Figure 11 shows the MS/MS spectra of all 6 aliphatic amino acids. The protonated molecular ions, $[M+H]^+$, were the precursor ions, and these have been labelled. In the spectra of alanine and glycine, no product ions are observed. This is due to the low-mass cut-off of the Orbitrap being 50 m/z . Therefore, the predicted product ions (due to loss of 46, HCOOH) were assigned to these two amino acids (as explained above in Table 1). The other four amino acids showed a neutral loss of 46 mass units. This is the first and most obvious trend observed with the aliphatic amino acids.

Additionally, it must be noted that isoleucine shows an abundant secondary loss of NH_3 from product ion m/z 86, whereas for its structural isomer, leucine, this is significantly reduced in intensity. Therefore, further investigation is needed to determine the possible explanation for this. A detailed discussion of this observation is presented in the later section.

ii. Acidic Group +ve CID MS/MS

Two amino acids belong to the acidic group: aspartic acid and glutamic acid. Structurally, glutamic acid has an extra CH_2 unit in its side chain. Their MS/MS spectra are presented below (Figure 12).

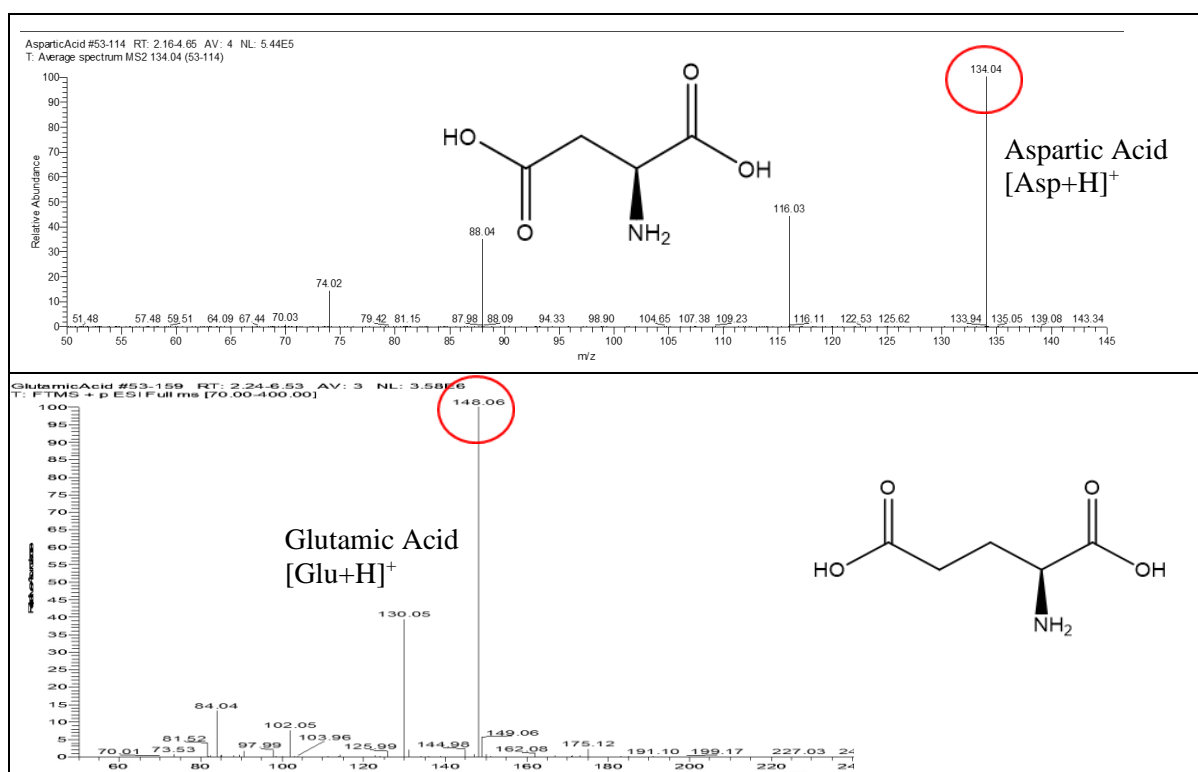


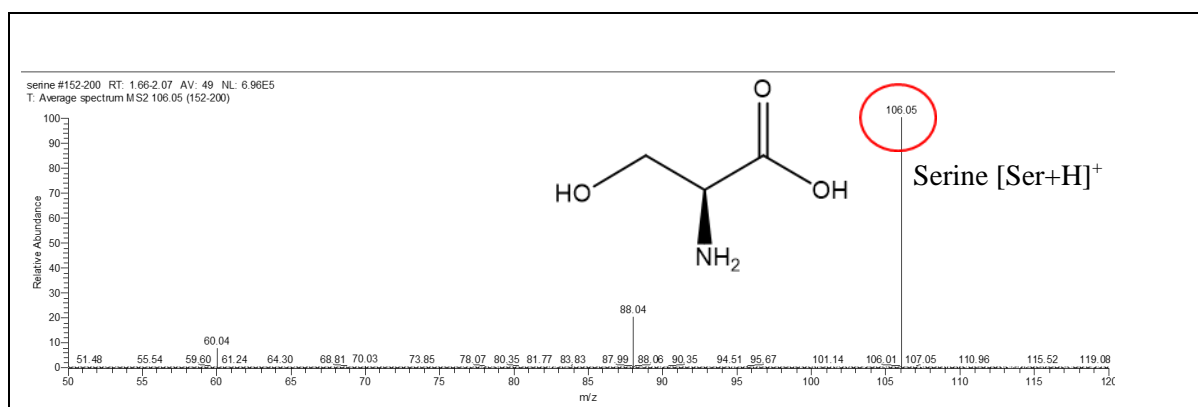
Figure 12 M/MS² spectra of aspartic acid and glutamic acid. The protonated molecular ions $[M+H]^+$ are the precursor ions, and they have been labelled in red circles.

From the MS/MS spectra above, it is clear that both of the acidic amino acids have an initial loss of H₂O (18), followed by a loss of CO (28). At the same time, the peak *m/z* 88 of aspartic acid and *m/z* 102 of glutamic acid could be obtained from a direct loss of HCOOH from the [M+H]⁺. To determine which may be the more favourable route, energy profiles of the respective pathways will be calculated using Gaussian and discussed later.

One difference is that aspartic acid has a loss of 42 from *m/z* 116 (initial loss of water), resulting in the product ion at *m/z* 74. For glutamic acid, a different route is followed, where there is an additional loss of water from product ion *m/z* 102 (loss of formic acid HCOOH), which is not observed with aspartic acid. The possible reason for this is linked to the extra carbon on the side chain, which could lead to the formation of a six-membered ring during the fragmentation in the gas phase. This finding contributes to the hypothesis of this current study.

iii. Hydroxylic Group +ve CID MS/MS

Both serine and threonine have an OH group on their side chain, and they are categorised to be in the hydroxylic group. In their MS/MS spectra (Figure 13), they both have two consecutive water losses and a HCOOH loss from the protonated precursor ions *m/z* 106 and *m/z* 120.



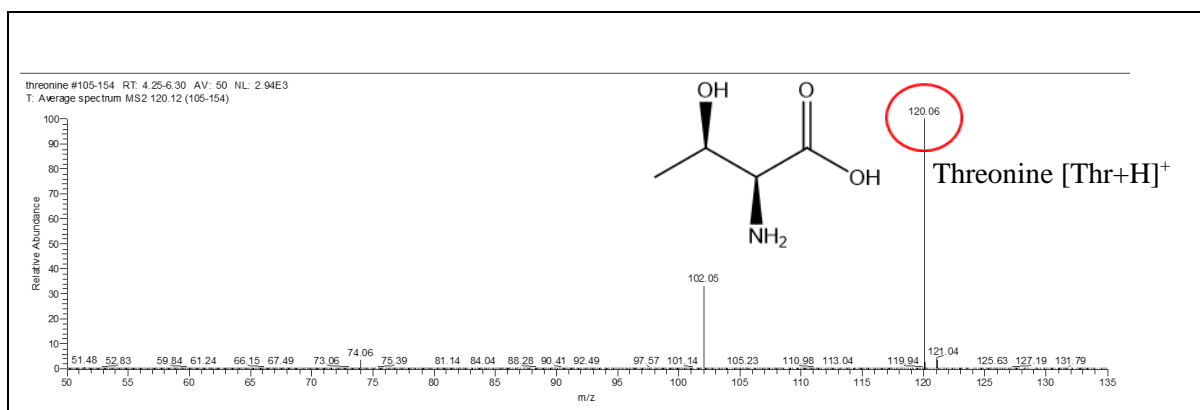


Figure 13 MS/MS spectra of serine and threonine. The protonated molecular ions $[M+H]^+$ are the precursor ions, and they have been labelled.

iv. Amidic Group +ve CID MS/MS

Asparagine and glutamine belong to the amidic group. They both have a side chain carboxamide, with glutamine having an extra CH_2 in its side chain. Their MS/MS spectra are shown in Figure 14.

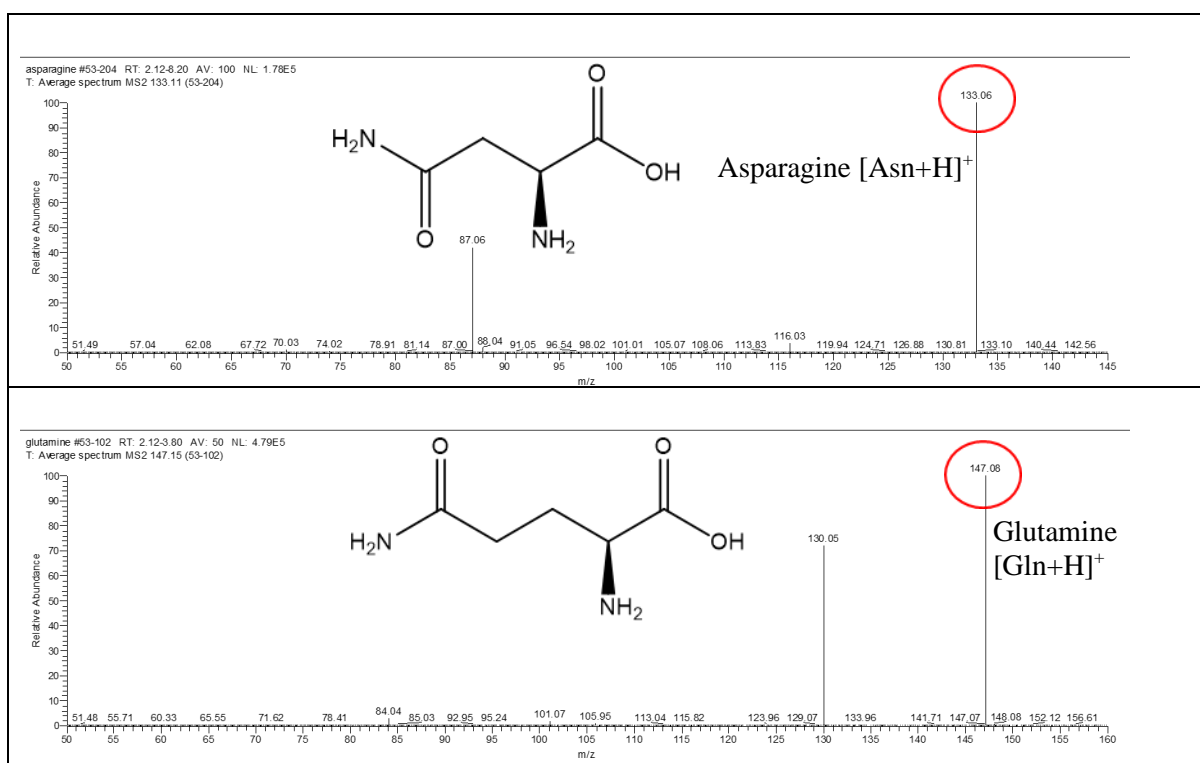
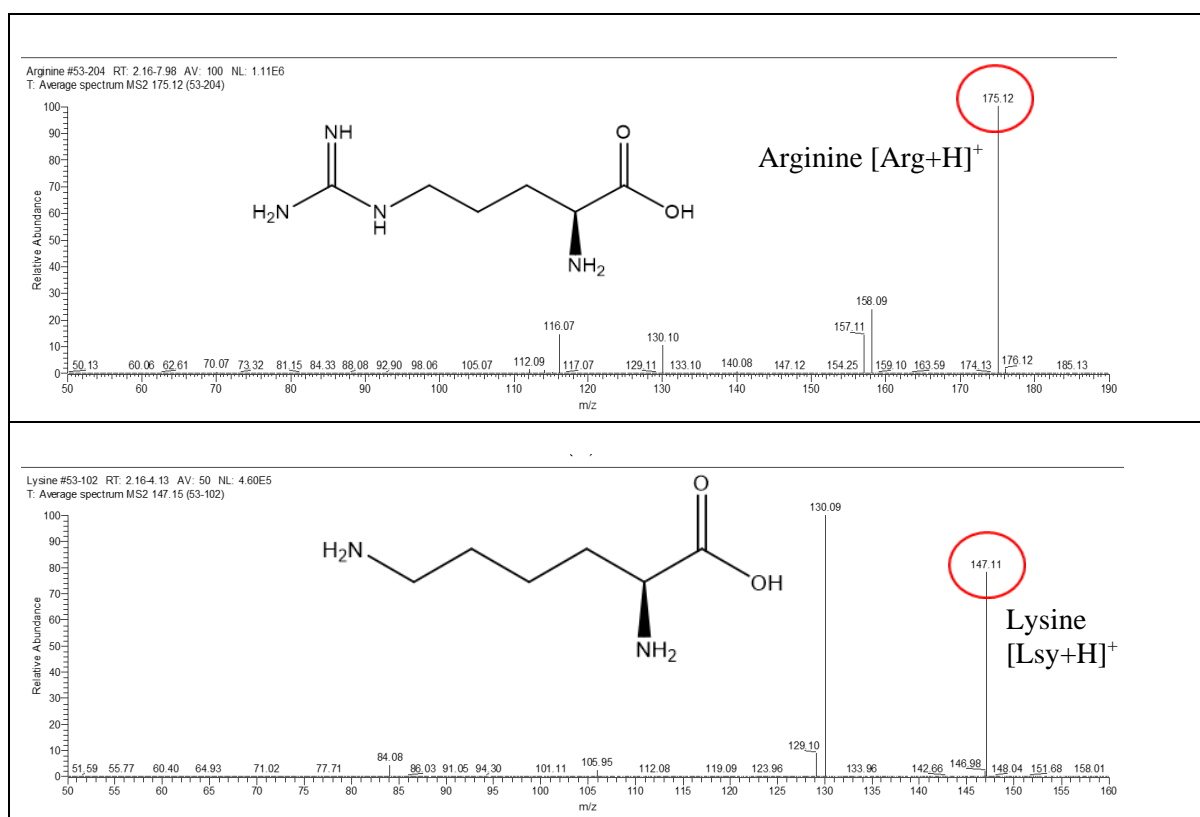


Figure 14 MS/MS spectra of asparagine and glutamine. The protonated molecular ions $[M+H]^+$ are the precursor ions, and they are labelled for clarity.

The first point to be discussed is the different intensities of NH_3 losses of these two amino acids. Compared to the NH_3 loss peak of asparagine, the peak for the same neutral loss at m/z 130 for glutamine is much more intense. In addition, by examining the MS/MS spectrum of asparagine, the loss of 46 (formic acid) is the most dominant, whereas it is the loss of 17 (ammonia) for glutamine. Finally, asparagine has losses of 17, 45, 46, and 59 (Table 2), whereas glutamine only has losses of 17 and 46 from the protonated molecular ion. The 46 loss occurs after the initial loss of 17. An additional loss of 45 also occurs in HCD (see supplementary materials). Even though both asparagine and glutamine possess a carboxamide on their side chain, the differences observed in the fragmentation must result from the extra CH_2 of glutamine.

v. Basic Group +ve CID MS/MS

Arginine, lysine, and histidine are basic amino acids with very different side chains. Arginine has a guanidine functional group on the end of the side chain, lysine has an extra amino group, and histidine has an imidazole side chain. Therefore, they are expected to have very different MS/MS spectra.



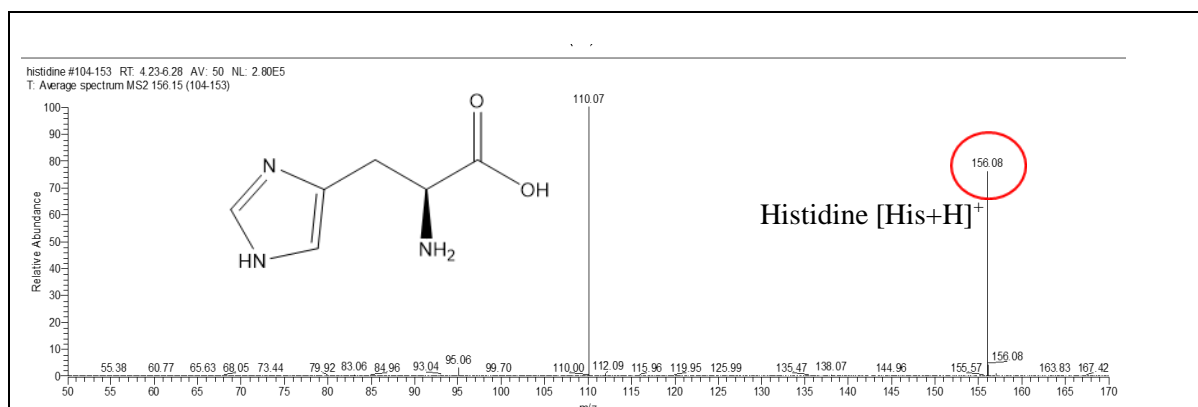


Figure 15 MS/MS spectra of arginine, lysine, and histidine. The protonated molecular ion peaks $[M+H]^+$ are the precursor ions, and they have been labelled.

Figure 15 shows the MS/MS spectra of arginine, lysine, and histidine. As predicted, they are all quite different. Arginine shows losses of 17, 18, 45, and 59, with 59 being the mass of guanidine. Lysine has the most abundant peak from the loss of 17 (m/z 130), which relates to its structure having two amino groups. It then has a loss of 18 from $[Lys+H]^+$ and a loss of 46 from peak m/z 130. The fact lysine does not have a loss of 46 directly from the protonated molecular ion but only 17 indicates that the loss of NH_3 is a much more favourable fragmentation pathway.

Unlike arginine or lysine, histidine has an imidazole side chain. This affects its MS/MS spectrum with a unique loss of 61 at m/z 95. It is not a neutral loss of imidazole itself, but as a nitrogen in the imidazole ring can be protonated, it stabilises a simultaneous loss of CO_2 and NH_3 in addition to a proton transfer to generate the terminal conjugated alkene. On the other hand, histidine still has a neutral formic acid like most other amino acids.

Observations here are sufficient to show that even in the same group, the different side chain functionalities can give rise to very different fragmentation pathways in the MS/MS spectra.

vi. Sulfur Containing Group +ve CID MS/MS

Cysteine and methionine are amino acids with sulfur on their side chains. Cysteine has an HS terminus to its aliphatic side chain, while methionine has a CH_3S- terminus and a longer hydrocarbon chain. Their MS/MS spectra are very different (Figure 16). These results demonstrate how a small change in the side

chain length and the sulfur functional group could significantly affect the fragmentation pathways observed.

The experiment found that cysteine did not dissolve very well in the water:methanol solvent. Therefore, peak intensities of cysteine are much lower than those of methionine. From the MS/MS spectrum of cysteine, losses of 17 and 46 can be observed. There are also some low intensity secondary losses from these two fragment peaks. There are no neutral losses involving sulfur. In addition, there are two peaks at m/z at 99 and 98, which could not be assigned to cysteine.

Methionine has a much higher signal-to-noise ratio and a cleaner spectrum than cysteine. Methionine's most abundant neutral loss is that of NH_3 , followed by a loss of formic acid at peak m/z 104 and a loss of 48. The loss of 48 represents a loss of a methanethiol neutral (CH_3SH). Upon closer analysis, there is also some secondary fragmentation. CH_3SH can also be lost after the initial loss of formic acid. There is also a loss of 72, which results from the bond breakage between the S and CH_2 from m/z 133 (loss of NH_3 from the protonated molecular ion). The major difference between methionine and cysteine is the involvement of sulfur in their fragmentation. There is no neutral loss involving sulfur for cysteine.

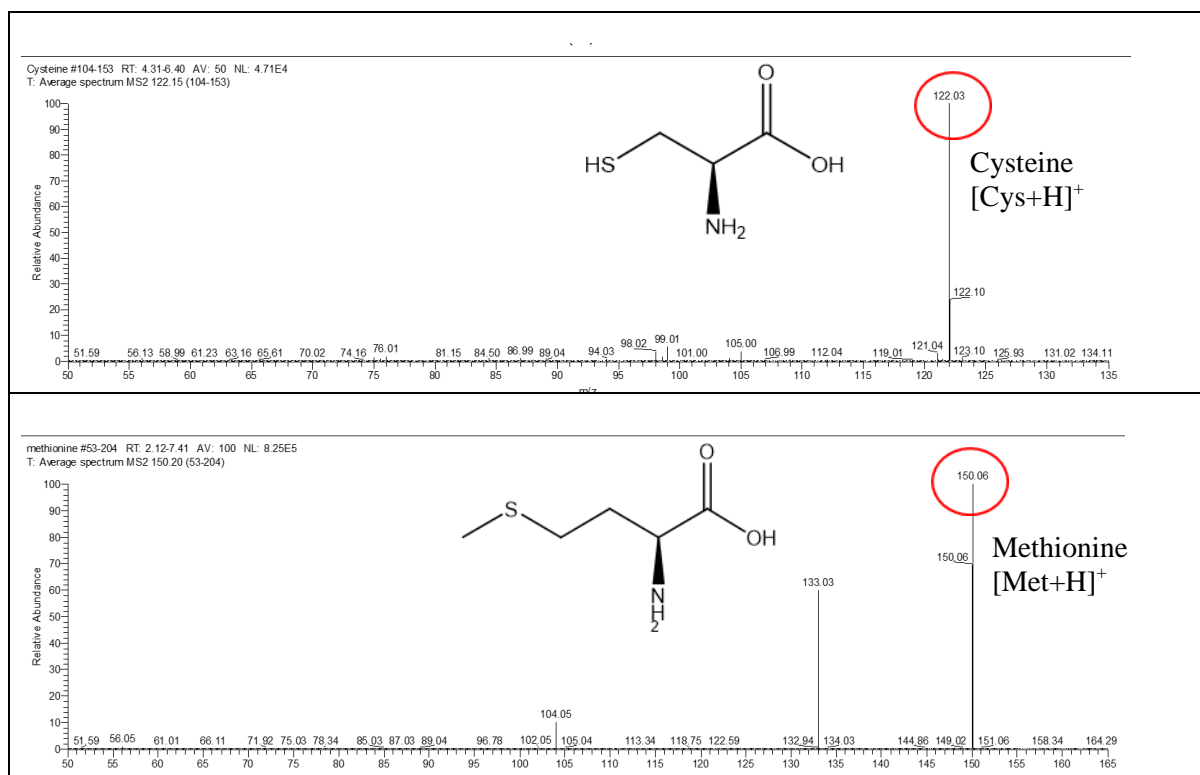
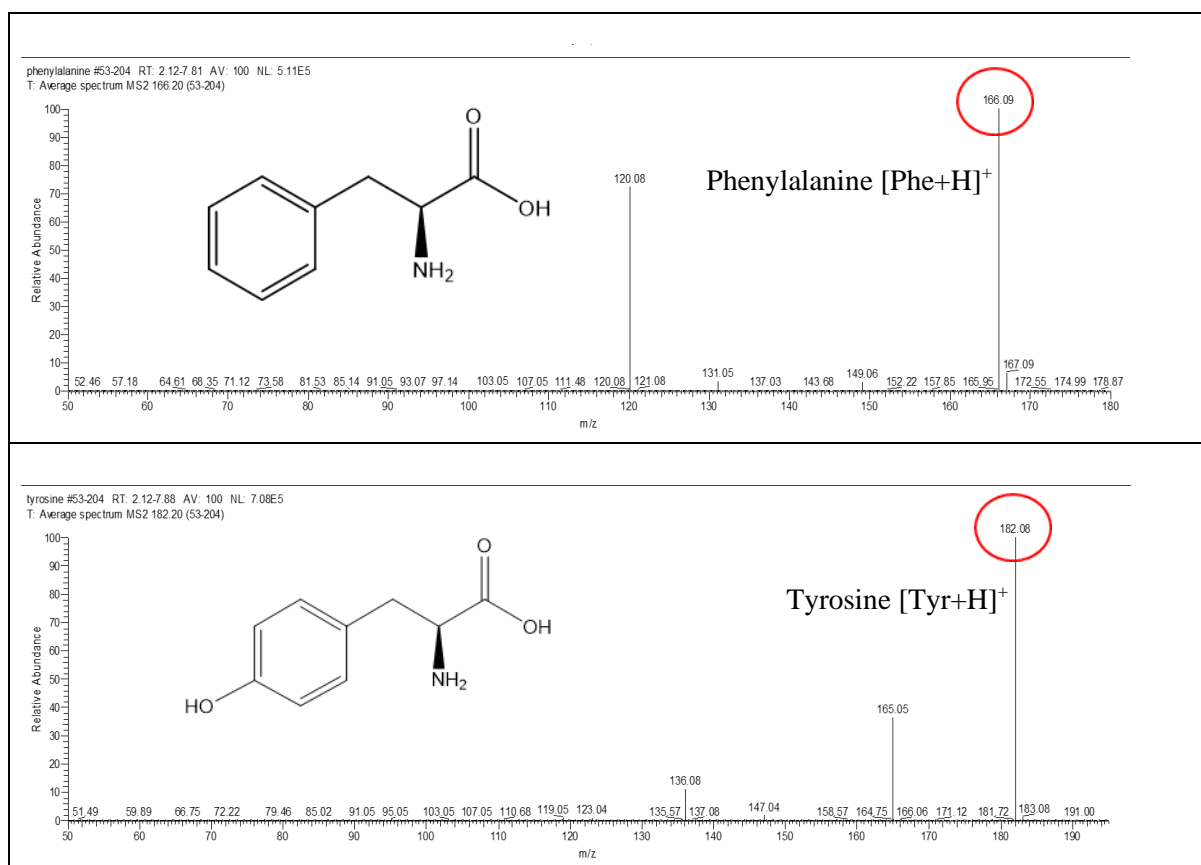


Figure 16 MS/MS spectra of cysteine and methionine. The protonated molecular ion peaks $[M+H]^+$ are the precursor ions, and they are labelled.

vii. Aromatic Group +ve CID MS/MS

Phenylalanine, tyrosine, and tryptophan make up the aromatic amino acid group. Phenylalanine has a phenyl group at the end of the side chain, whereas tyrosine has a phenol. Tryptophan has an indole side chain. From the similarity of the side chains of phenylalanine and tyrosine, it might be expected that they would share common fragmentation pathways (see Figure 17). The MS/MS spectrum of tryptophan is different as the protonated molecular ion could not be selected automatically. As a result, the fragmentation energy ramp was performed manually. Here, the MS² spectrum of tryptophan was recorded at a CID energy of 50 eV.



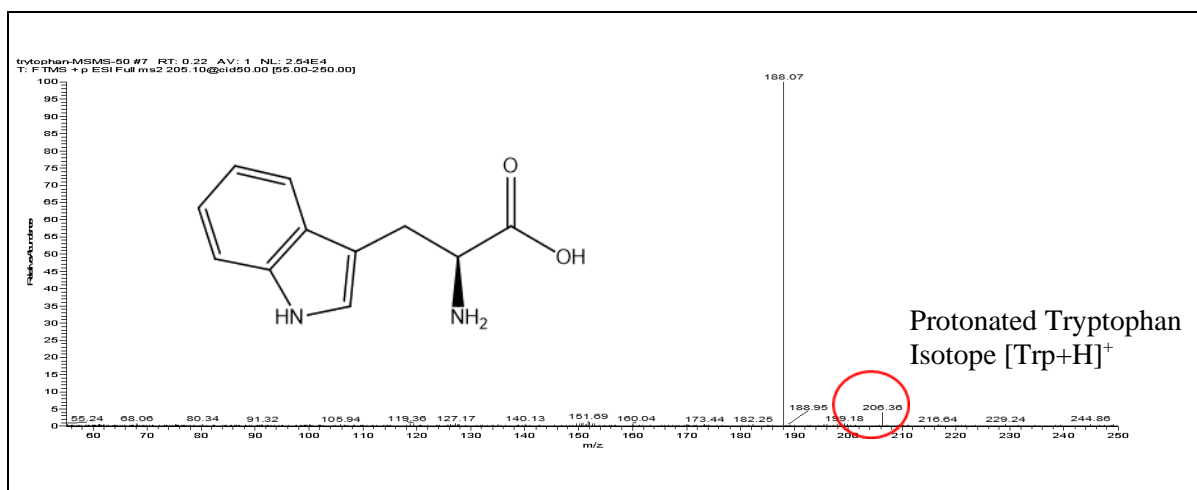


Figure 17 MS/MS spectra of phenylalanine, tyrosine, and tryptophan. The protonated molecular ions $[M+H]^+$ are the precursor ions, and they are labelled except for tryptophan, where only an isotopic peak is present in the spectrum.

The MS/MS spectrum of phenylalanine has a dominant peak at m/z 120, which corresponds to a loss of formic acid. In addition, there are other peaks due to a neutral loss of NH_3 , secondary loss of H_2O (after the initial loss of NH_3), and secondary loss of NH_3 after the initial loss of formic acid. These peaks all have much lower intensities than the predominant loss of formic acid. In contrast, tyrosine has a different spectrum with an intense loss of NH_3 and a much less abundant loss of formic acid. However, tyrosine and phenylalanine share the same fragmentation pathway of secondary water loss after the initial NH_3 loss.

The MS/MS spectra of tryptophan were recorded manually. In the spectrum presented here (at collision energy 50eV), there is only a very intense fragment peak at m/z 188. This is very different from phenylalanine and tyrosine and could result from the stabilising effect of the indole side chain, which is not present with the other two aromatic amino acids. Interestingly, more peaks are observed in the HCD-MS/MS of tryptophan - these will be discussed later. MS^3 has been performed on m/z 188 to confirm some of these fragmentation pathways. Those spectra will also be examined in the tryptophan mechanism section.

2.3.2.2. Positive Ion Mode HCD

i. Aromatic Group +ve HCD MS/MS

As seen in Table 4, comparatively, there are more product ions in HCD-MS/MS spectra than in CID-MS/MS spectra. Here, the spectra of the aromatic amino acids obtained in HCD are presented (Figure 18). These illustrate the major differences between HCD and CID for aromatic amino acids. The most noticeable aspects are that there are more product ion peaks in HCD spectra, and some have higher intensities.

Compared to Figure 17, there are notably more peaks in all three MS/MS spectra in Figure 18. The number of identified product ions is the same for phenylalanine, but the intensity of the neutral loss of HCOOH (at m/z 120) is much higher in HCD than that in CID. Other product ion peaks have also increased in intensity, especially m/z 103, which was hard to distinguish from the background in CID. In the MS/MS of tyrosine, the peaks at m/z 147 and m/z 136 are much more prominent. Additionally, two new peaks can be observed at m/z 123 and m/z 119, which are not seen in CID spectra.

With tryptophan, the tendency to observe more product ion peaks in HCD remains consistent. Compared to Figure 17, the MS/MS of tryptophan in Figure 18 has shown a dramatic increase in product ion peaks. The peak intensity at m/z 188 is reduced. Instead, m/z 146 is the most abundant. In addition, many other peaks have raised intensities — for example, m/z 170, 159, 146, 144, 132, and 118. An MS³ analysis has also been performed to confirm the precursor ions for these additional ions, which will be discussed later.

Name (Protonated Molecular Ion m/z)	CID Product Ion m/z	HCD Product Ion m/z
Phenylalanine (m/z 166)	149, 131, 120, 103	149, 131, 120, 103
Tyrosine (m/z 182)	165, 147, 136	165, 147, 136, 123, 119
Tryptophan (m/z 205)	188	188, 170, 159, 146, 144, 142, 132, 118

Table 4 Comparison of product ions generated by CID and HCD in positive ion modes for the three aromatic amino acids.

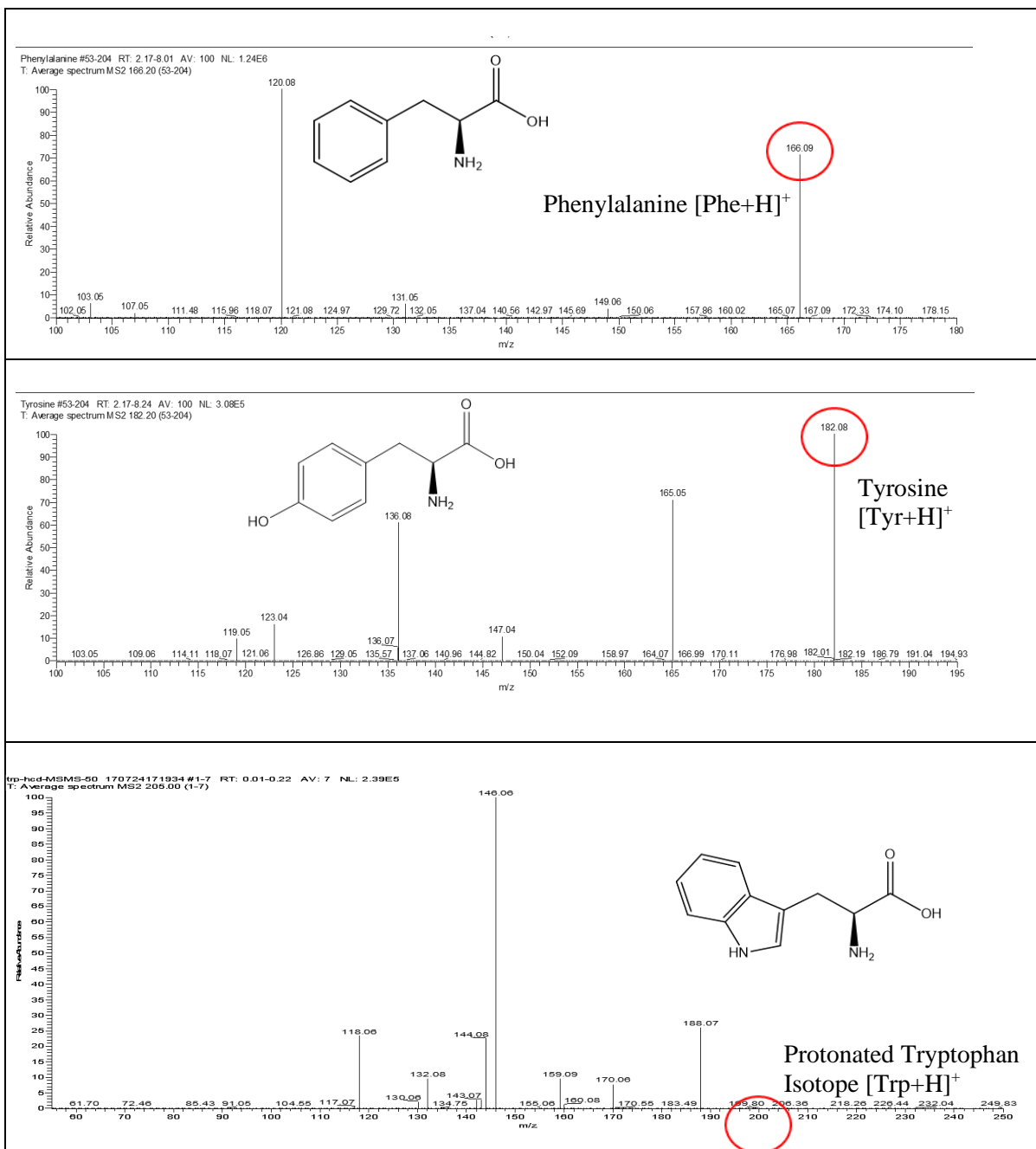


Figure 18 HCD MS/MS spectra of phenylalanine, tyrosine, and tryptophan. The protonated molecular ion peaks $[M+H]^+$ are the precursor ions, and they are labelled except for tryptophan. In the spectrum of tryptophan, only an isotopic ion can be observed.

ii. MS³ of Tryptophan +ve CID MS³

The HCD MS/MS spectrum of tryptophan shows some interesting product ions. An MS³ analysis of m/z 188 has been performed to confirm their fragmentation pathways. An energy ramp was run to see

the changes in the intensities of these ions (spectra were recorded with collision energy from 1 to 50eV), and the spectrum at collision energy 50eV is shown in Figure 19.

As seen in Figure 10, the peak at m/z 146 is the most abundant, confirming it is from the precursor ion 188. In addition, ions resulting from the same precursor ion are 170, 144, 142, and 118. The suggested mechanisms for these ions are presented later. A possible schematic route for these fragments is 170 (loss of H_2O from 188); 146 (loss of 42 from 188, with a formula of $\text{C}_2\text{H}_2\text{O}$); 144 (loss of CO_2); 142 (loss of HCOOH) and 118 (loss of CO from m/z 146).

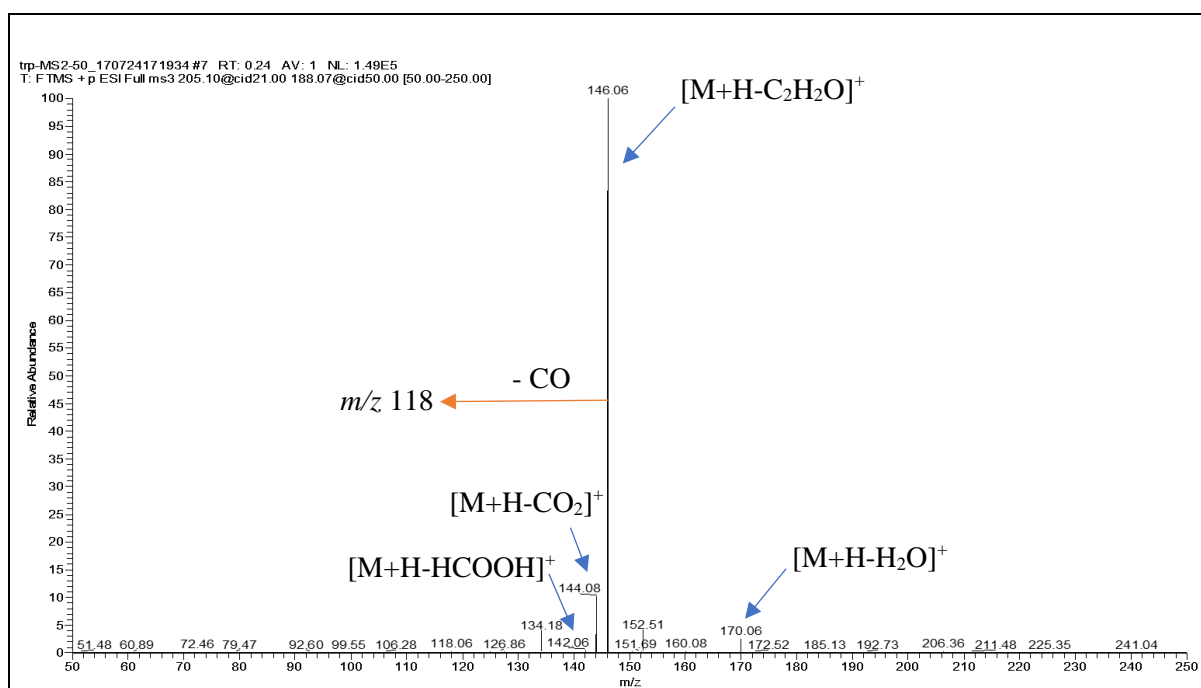


Figure 19 MS³ spectrum of m/z 188 of tryptophan. Route: Full scan \rightarrow m/z 205 \rightarrow m/z 188.

2.3.2.3. Negative Ion Mode

Some amino acids (especially aliphatic ones) did not give satisfactory spectra in negative ion CID mode. Therefore, these will not be discussed in detail in this section. However, their neutral losses are still summarised in Table 2 to ensure this part of the research is comprehensive. Negative ion MS/MS spectra of other amino acids are all presented and discussed below.

i. Acidic Group -ve CID MS/MS

Aspartic acid and glutamic acid share a degree of similarity in their negative ion MS/MS spectra (Figure 20). However, glutamic acid has fewer peaks.

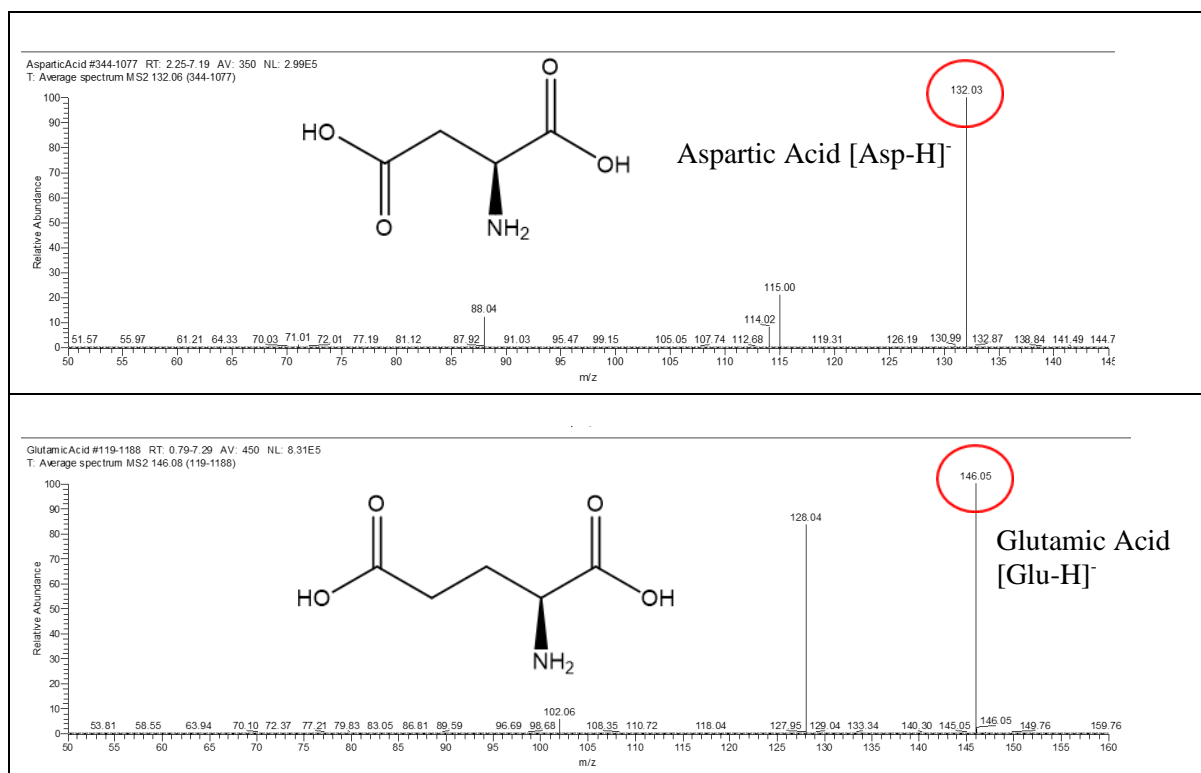


Figure 20 Negative ion CID MS/MS spectra of aspartic and glutamic acid. The deprotonated molecular ion $[M-H]^-$ peaks are the precursor ions, and they are labelled for clarity.

In Figure 20, aspartic acid has losses of NH_3 , H_2O , CO_2 , and a considerably less abundant product ion at m/z 71, which could result from a second CO_2 loss from m/z 115 (after an initial loss of NH_3). The spectrum of glutamic acid has relatively few peaks. The peaks at m/z 128 and 102 correlate to the losses

of water and CO₂, respectively. The additional CH₂ could explain the reason for the different fragmentation patterns in the side chain of glutamic acid.

ii. Amidic Group -ve CID MS/MS

Asparagine and glutamine have very similar negative ion CID-MS/MS spectra (Figure 21). They both have losses of NH₃ and two consecutive water losses. In the MS/MS of asparagine, the ratio of the H₂O: NH₃ (m/z 113: m/z 114) peak abundance is higher than that of glutamine. In addition, the second water loss at m/z 95 is also more abundant than that of glutamine. Interestingly, the fragmentation in the positive mode is quite different, but in the negative ion mode, the fragmentation is very similar. In this case, the addition of the extra carbon in the side chain of glutamine does not seem to alter the fragmentation pathway compared to asparagine.

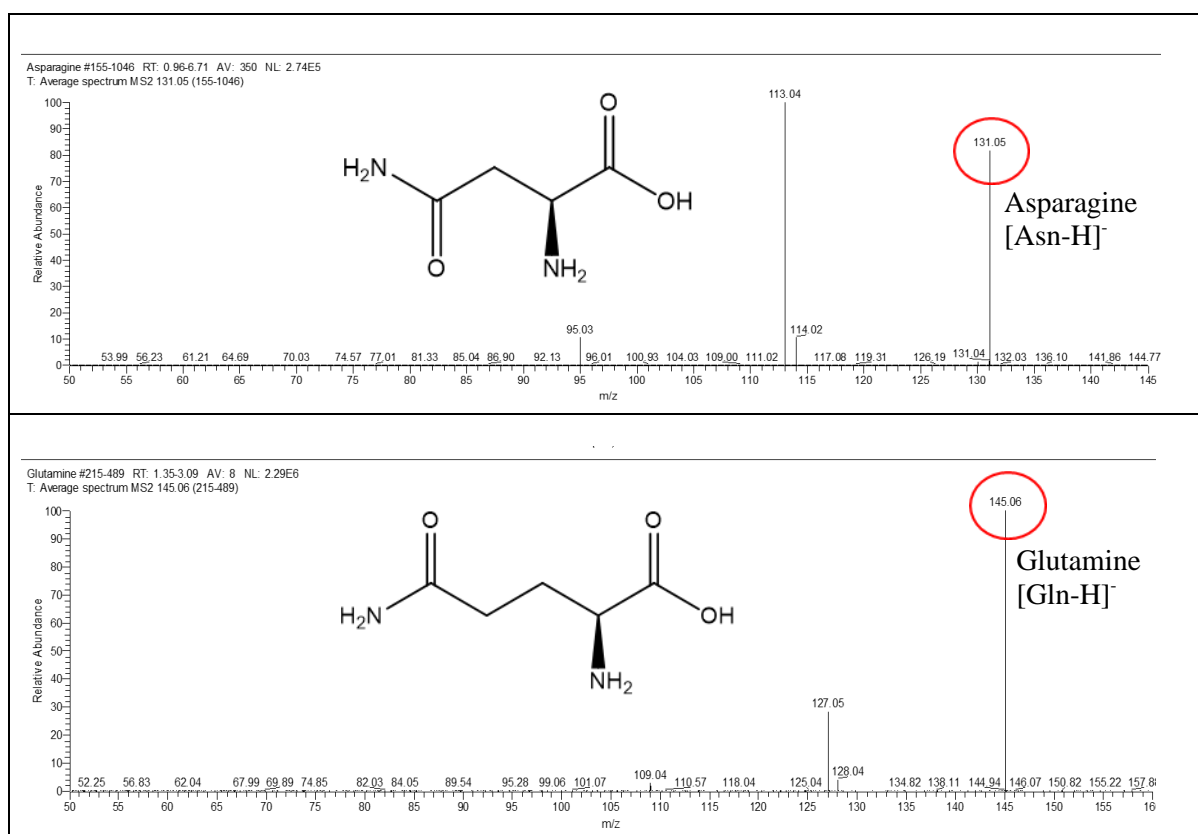


Figure 21 Negative ion CID MS² spectra of asparagine and glutamine. The deprotonated molecular ion [M-H]⁻ peaks are the precursor ions, and they are labelled for clarity.

iii. Basic Group -ve CID MS/MS

Figure 22 shows the negative ion mode CID tandem mass spectra of arginine and histidine. Due to the structural differences in their side chains, it is expected that their fragmentation pathways would be quite different. In the positive ion mode CID, arginine has more product ions than histidine. Here, in the negative ion mode, it is the opposite, and histidine has more ions when compared to arginine, although they are at low peak intensities.

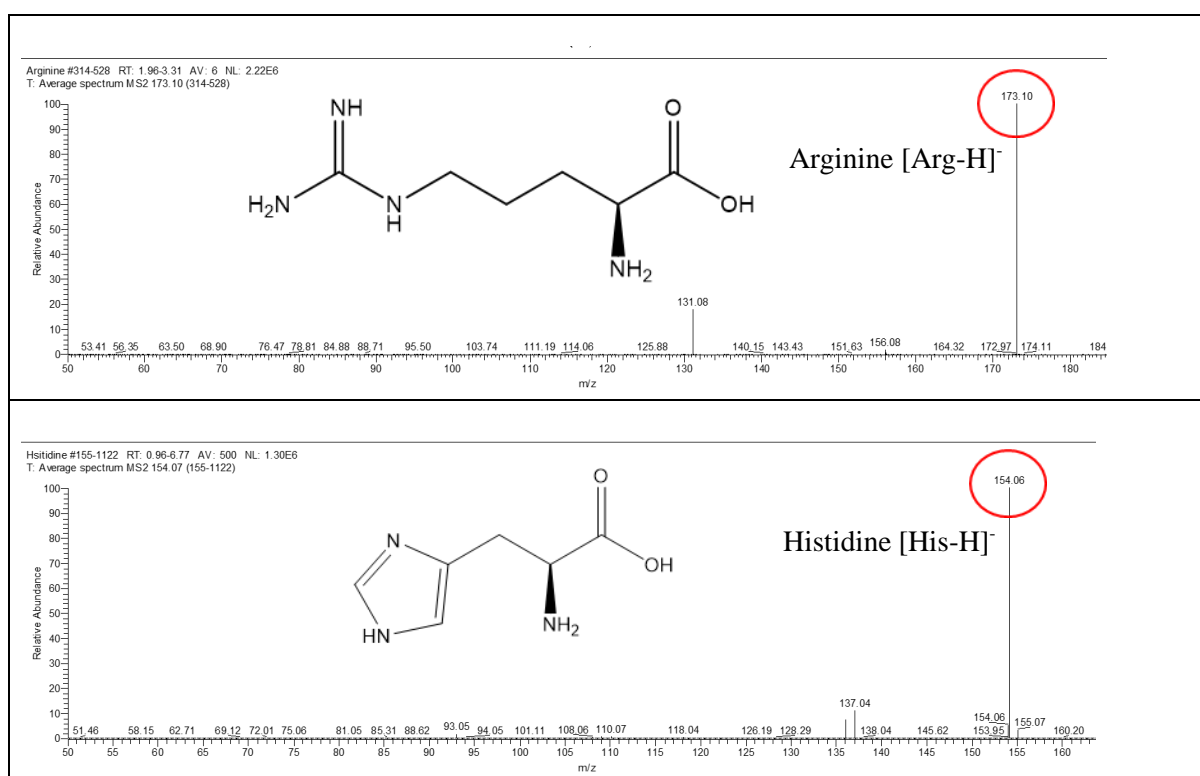


Figure 22 Negative ion CID MS/MS spectra of arginine and histidine. The deprotonated molecular ion $[M-H]^-$ peaks are the precursor ions, and they are labelled for clarity.

In Figure 22, arginine has two product ions at m/z 156 and m/z 131 due to a loss of NH_3 and mass 42 from the guanidine group on the side chain. The spectrum of histidine has an additional water loss peak and a loss of CO_2 from a peak at m/z 137.

iv. Aromatic Group -ve CID MS/MS

In the mass spectra below, tyrosine and tryptophan have more product ions than phenylalanine. However, as they are all aromatic amino acids, these three still share the common neutral losses of NH_3 and CO_2 . Tyrosine has an additional CO_2 loss from the peak at m/z 163 (after an initial loss of NH_3), followed by another cleavage in the aliphatic chain, resulting in the fragment peak at m/z 93. Contrary to what happens in positive ion mode, tryptophan follows essentially the same fragmentation pathway. Here, tryptophan loses CO_2 after an initial loss of NH_3 , followed by a loss of 26 (cleavage of the carbon aliphatic chain). The mechanism corresponding to this route will be discussed later. Phenylalanine follows a different fragmentation pathway.

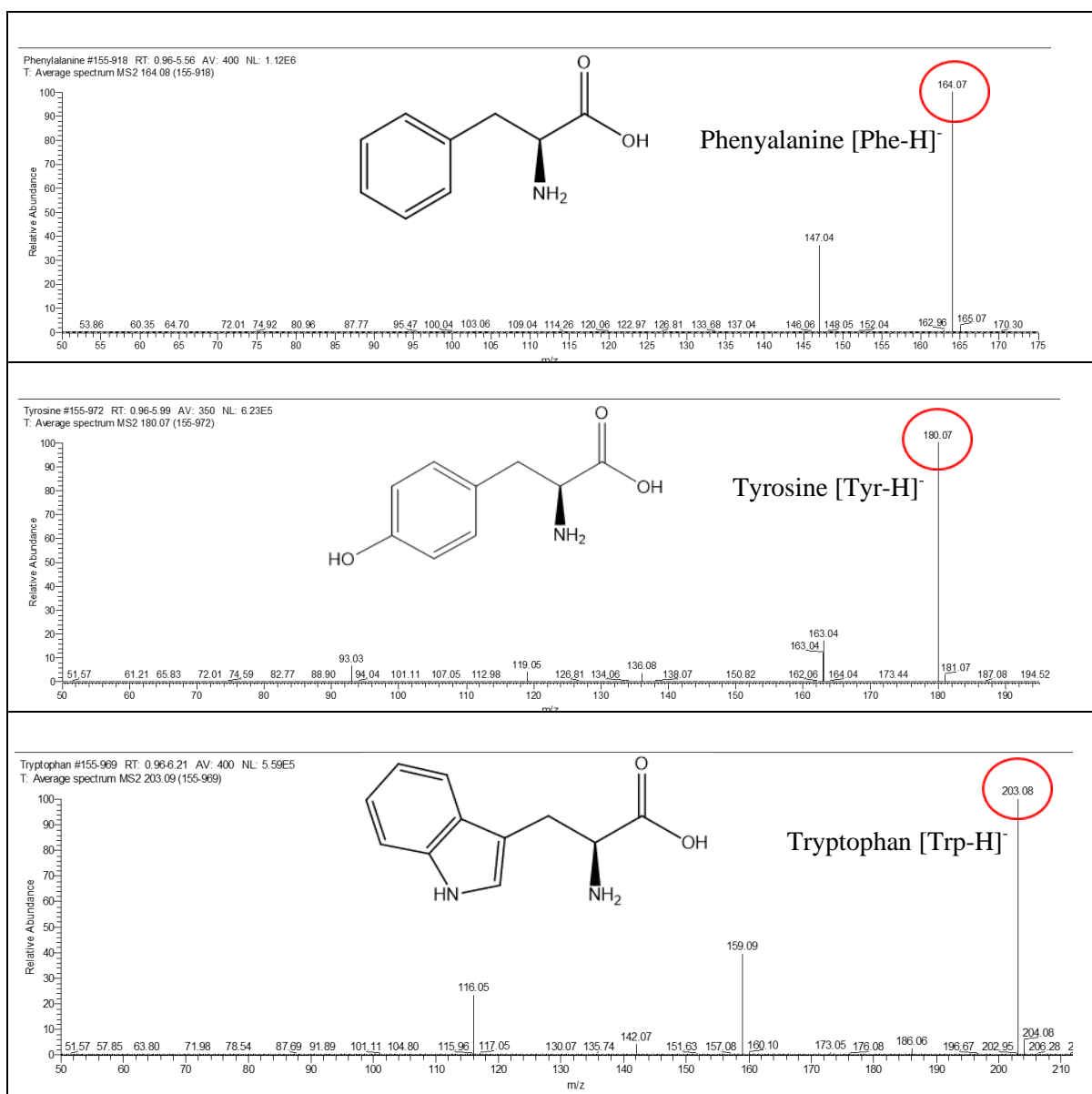


Figure 23 Negative ion CID MS/MS spectra of phenylalanine, tyrosine, and tryptophan. The deprotonated molecular ion $[M-H]^-$ peaks are the precursor ions, and they are labelled for clarity.

2.3.3. Energy Breakdown Graph

The collision energy ramp experiment performed during the CID-MS/MS analysis in the positive ion mode can give rise to an energy breakdown graph for the fragmentation of each protonated amino acid ($[M+H]^+$ precursor ions). As described in the Experimental section, the collision energy used varied from 1 to 50, and individual MS/MS spectra were recorded at each unit of collision energy. The intensity of the product ions obtained from these spectra is then plotted against collision energy.

Due to the similarities of some breakdown graphs, especially those with a similar side chain functionality, only a few are included here to demonstrate the general trends. Energy breakdown graphs correlate peak intensity to collision energy in MS/MS analysis. By examining these graphs, it is possible to see the effect of increasing the collision energy on product ion intensities and the relationships between different ion intensities over the energy range used. These results will be further rationalised by computational modelling and energy calculations from Gaussian 09 later in this thesis.

Another characteristic of these graphs is their shape difference between CID and HCD collision modes. Compared to CID, HCD occurs in a separate collision cell outside the ion trap and is recognised as a more focused collision method with higher levels of energy transfer. Energy breakdown graphs obtained from HCD are more like those in previous studies, which often used triple-quadrupole instruments. These graphs have a general “tailing-off shape”, whereas the CID energy breakdown graphs have peaks with flat ends. These differences will be demonstrated towards the end of this section during the discussion of the aromatic amino acid group.

i. Aliphatic Group +ve CID MS/MS

As discussed previously, the loss of HCOOH is non-discriminating for the aliphatic amino acids. The exception here is isoleucine, which exhibits a secondary NH_3 loss after an initial loss of HCOOH. Hence, breakdown graphs for leucine and isoleucine are displayed as examples (Figure 24).

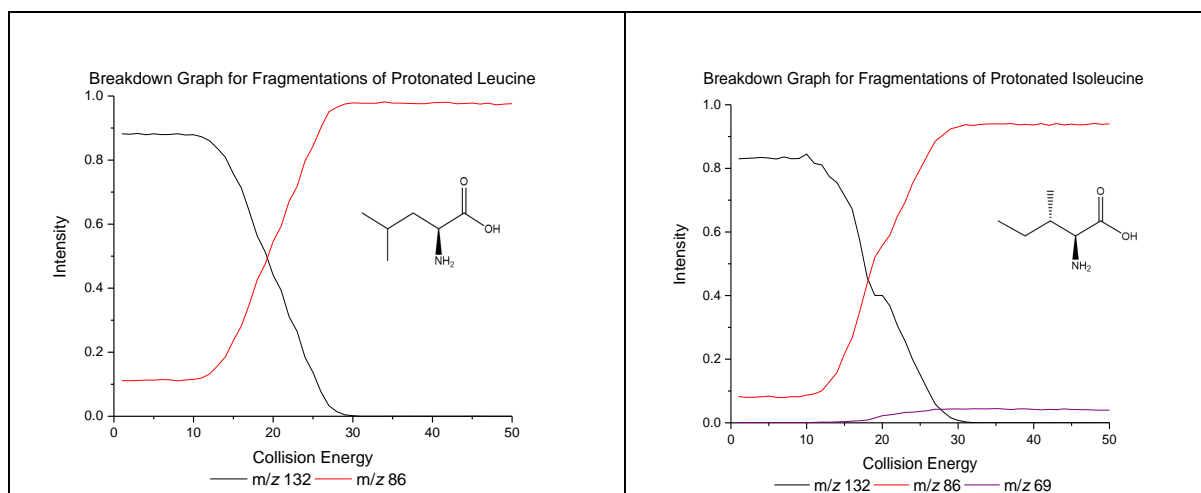


Figure 24 Energy breakdown graph for leucine (left) and isoleucine (right).

For clarity, the line corresponding to the loss of HCOOH is coloured **red**, the loss of NH₃ is coloured **grey**, the loss of H₂O is **blue**, and any secondary loss is generally coloured **purple**. All other colours are for additional neutral losses that only occasionally appear. This colour scheme is applied consistently throughout all breakdown graphs for uniformity.

Figure 24 shows the breakdown graphs with collision energy plotted on the x-axis and normalised peak intensity of the corresponding product ions on the y-axis. Leucine is on the left, and isoleucine is on the right. The breakdown graph of leucine is used here as a representation graph for the whole group.

For leucine, as the precursor ion peak intensity reduces, there is a rise in the intensity of the loss of HCOOH. When the CID energy reaches 30 eV, the precursor ion (m/z 132) disappears, leaving only the product ion at m/z 86, corresponding to 46 (HCOOH) loss from the precursor ion. It is also worth noting that m/z 86 occurs at 0.1 intensity when CID energy is only 1eV.

In comparison, the breakdown graph of isoleucine has an additional loss of 17 from m/z 86, which, as mentioned previously, occurs at about 10% of HCOOH loss. Other features of this graph are very similar to leucine apart from the short region of stabilisation in the intensity of the precursor ion between collision energy 18 eV to 21 eV when product ion m/z 69 starts to appear. This clearly illustrates the correlation between precursor ion intensity and the appearance of product ion peak.

ii. Acidic Group +ve CID MS/MS

As described previously, aspartic and glutamic acid have very different fragmentation pathways. Their energy breakdown graphs are also quite distinctive (Figure 25). Both amino acids have four product ion peaks, but the purple lines represent different neutral losses. In addition, the peak intensities of the same neutral loss of these two acidic amino acids are quite different.

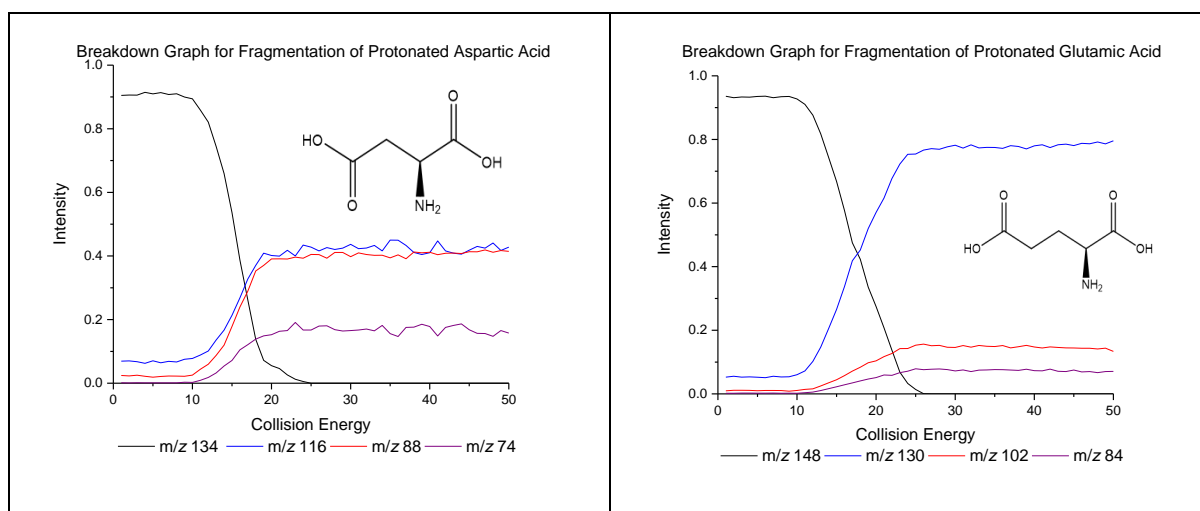


Figure 25 Energy breakdown graphs for aspartic acid (left) and glutamic acid (right).

Both aspartic and glutamic acid have neutral losses of H_2O and HCOOH , represented by the blue and red lines in their breakdown graphs. In the breakdown graph of aspartic acid, the intensity of the water loss is only slightly higher than that of formic acid. This trend is very different from that of glutamic acid, where the ratio of H_2O loss versus HCOOH loss is almost 8:1 above a collision energy of 22 eV. Another point to notice is that for the secondary losses, the m/z 74 product ion of aspartic acid has a much higher intensity than the m/z 84 ion of glutamic acid.

Finally, by studying the rate of increases in the intensity of these ions, although they all start to rise just after collision energy is 10 eV, their highest points correspond to different energies for aspartic acid and glutamic acid. This threshold is just before 20 eV for aspartic acid, whereas for glutamic acid, it is higher. One explanation for this is the obvious structural differences in their side chains, with an additional CH_2 for glutamic acid, which requires more energy for glutamic acid to fragment into its product ions.

iii. Hydroxylic Group +ve CID MS/MS

Figure 26 shows breakdown graphs for both serine and threonine. They have identical fragmentation pathways, and their breakdown graphs are similar. Towards higher energy, there is a slight instability of the intensity of water loss with serine. This could be because the product ion is getting smaller, resulting in inefficient trapping/detection for this specific ion.

One interesting trend to point out is that the reduction in the loss of formic acid is accompanied by a major increase in the intensity of the water loss. The competition between these two fragmentation routes is demonstrated. The presence of the additional OH group on their side chains results in these two hydroxylic amino acids preferring the loss of H₂O over HCOOH. In addition, the second H₂O loss happens at higher collision energy (from CID = 15 eV in both cases) and only after the HCOOH loss starts to drop.

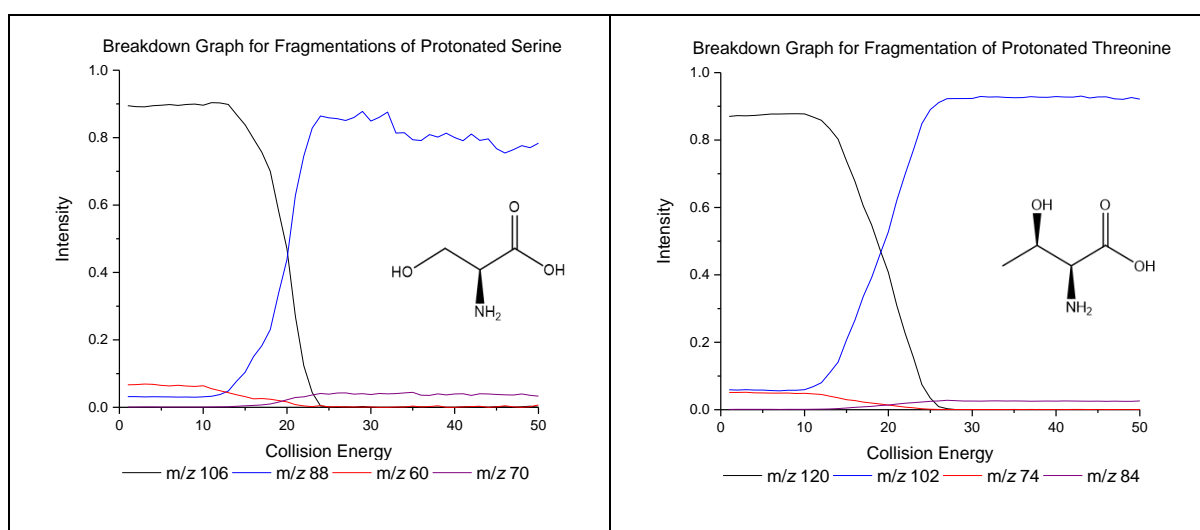


Figure 26 Breakdown graphs for serine (left) and threonine (right).

iv. Amidic Group +ve CID MS/MS

The most noticeable difference in the breakdown graphs of asparagine and glutamine are the intensities of the losses of formic acid and ammonia (Figure 27), represented by the red and grey lines, respectively. For asparagine, the loss of formic acid is the most abundant fragmentation. whereas for glutamine, the

loss of ammonia is the most intense route. This difference demonstrates that it is more favourable for asparagine to lose HCOOH than NH₃, which is the opposite for glutamine.

This observation cannot be reasoned simply by looking at their respective MS/MS spectra. Glutamine has an extra CH₂ in the side chain, which may provide extra stability for the resulting cation after the initial NH₃ loss. In this case, it is possible to form a five-membered ring, resulting in a more stable structure. This observation again links to our initial hypothesis that changing the side chain can not only influence fragmentation pathways being followed but also product ion stabilities observed in their changes in intensities. Investigating the formation mechanisms to explore the possible formation routes for the competing product ions will be useful. Energy calculations performed by Gaussian give insight into their relative stabilities, thus explaining their distinctive energy breakdown graphs.

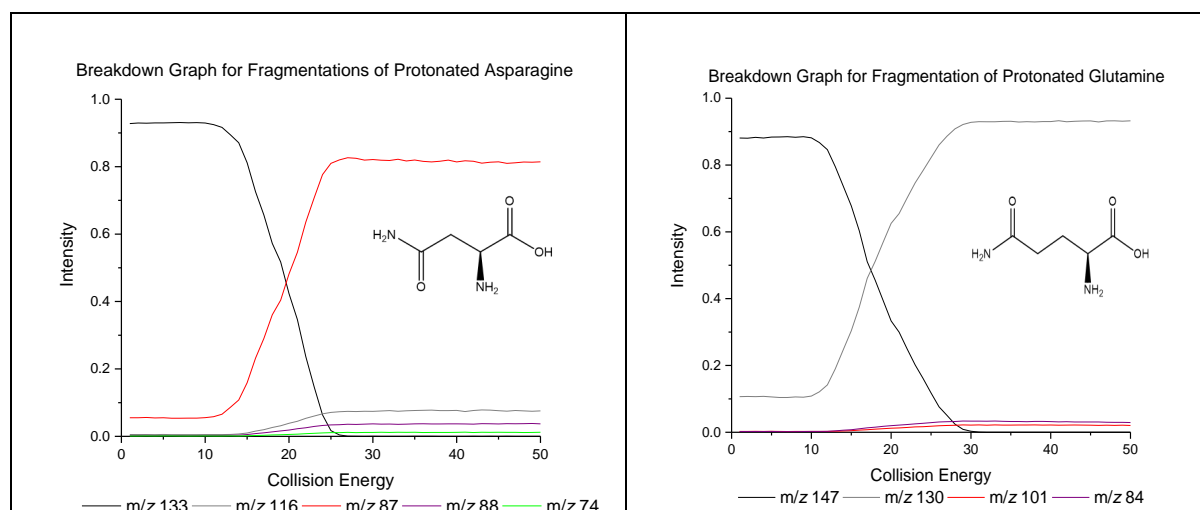


Figure 27 Breakdown graph for asparagine (left) and glutamine (right).

v. Basic Group +ve CID MS/MS

Figure 28 shows the breakdown graphs for arginine, histidine, and lysine. There is a considerable drop in the intensity of the loss of NH₃ in lysine (grey line). The reason for this is not clear but may be an instrumental anomaly. By examining these three graphs, the first thing to notice is that only histidine has a loss of HCOOH. Its breakdown graph is more like a simple aliphatic amino acid, with a predominant loss of formic acid and no loss of NH₃. This is very different from arginine and lysine.

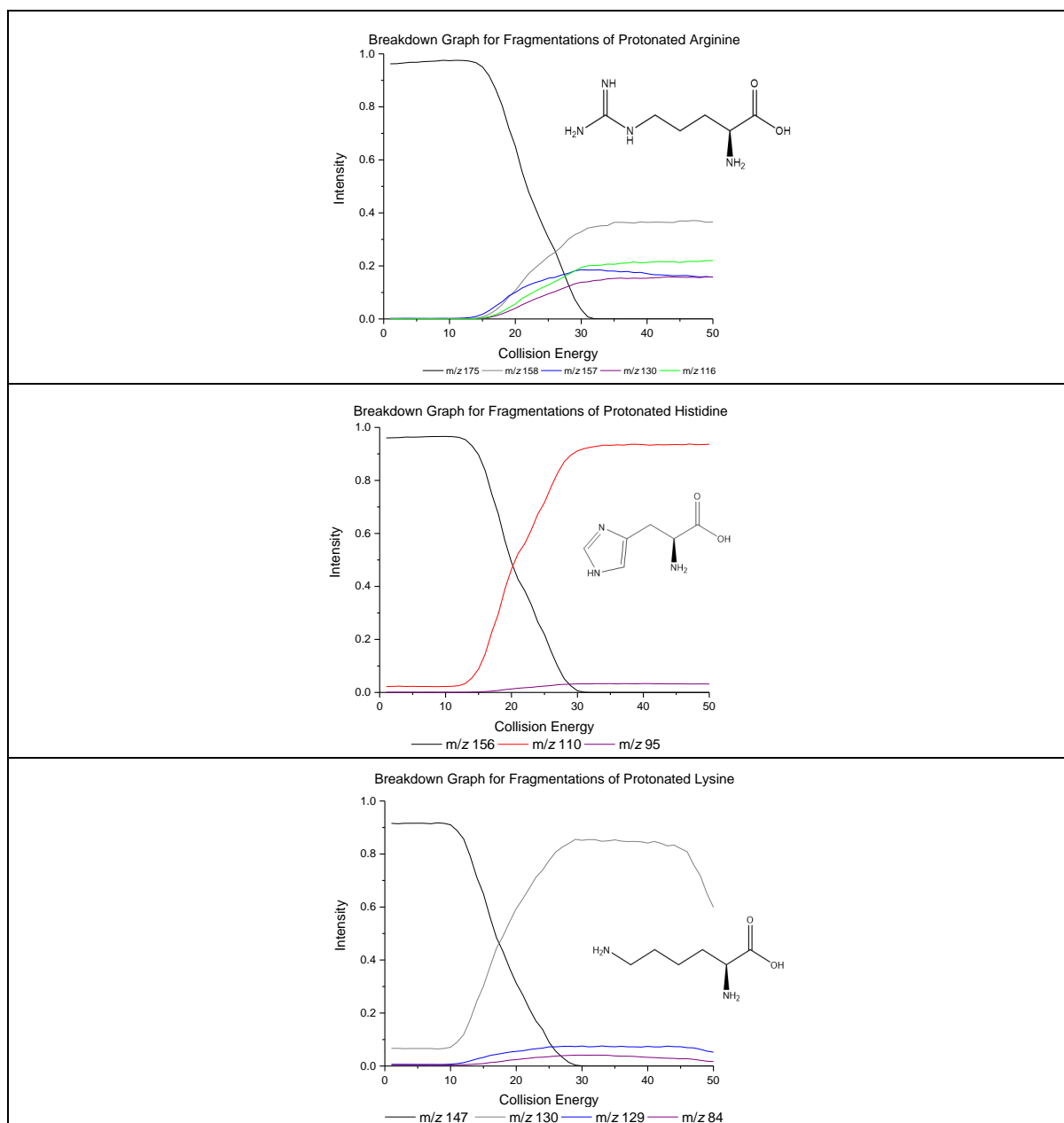


Figure 28 Breakdown graphs for arginine (top), histidine (middle), and lysine (bottom).

Compared to the breakdown graph for arginine, lysine has fewer product ions, but the loss of NH₃ occurs at the highest intensity. Having two NH₂ groups in its structure means loss of NH₃ could result from either, making this loss more intense. In the breakdown graph for arginine, the loss of NH₃ is slightly more abundant than H₂O, HCONH₂, and guanidine. These fragment peaks appear when collision energy is about 15 eV, and they all peak at 30 eV.

vi. Sulfur Containing Group +ve CID MS/MS

Figure 29 shows the breakdown graphs for cysteine and methionine. Due to cysteine's low solubility in a water-methanol solvent mix, the CID-MS/MS spectra have much lower peak intensities, and all spectra had to be obtained manually. Therefore, the peaks are less smooth in the energy breakdown graph.

In comparison, both graphs have the loss of NH_3 as the most abundant route. However, the second most intense peak for cysteine is a secondary loss of H_2O (after the initial loss of NH_3), whereas, for methionine, it is the loss of formic acid. The loss of HCOOH and lower intensity peaks can be seen in the breakdown graph for methionine. These lower intensity peaks are coloured differently and result from neutral loss involving either sulfur or breakage between a C-C bond on the side chain. It is intriguing to notice that instead of losing HCOOH , cysteine prefers to lose NH_3 , and the loss of HCOOH does not seem to increase its intensity as collision energy rises. In addition, compared to methionine, cysteine has fewer product ions and much lower intensities. A close examination of the fragmentation mechanisms will help explain these observations.

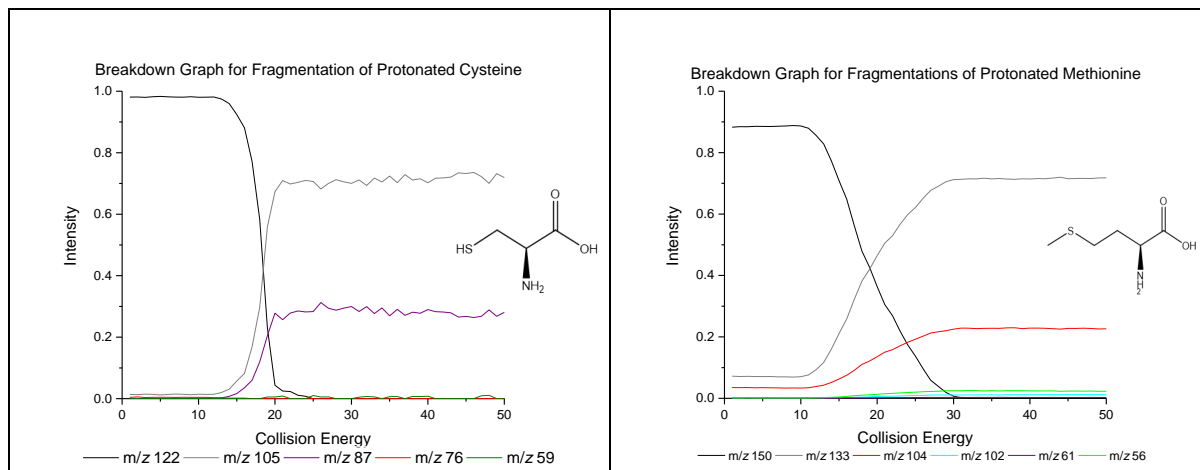


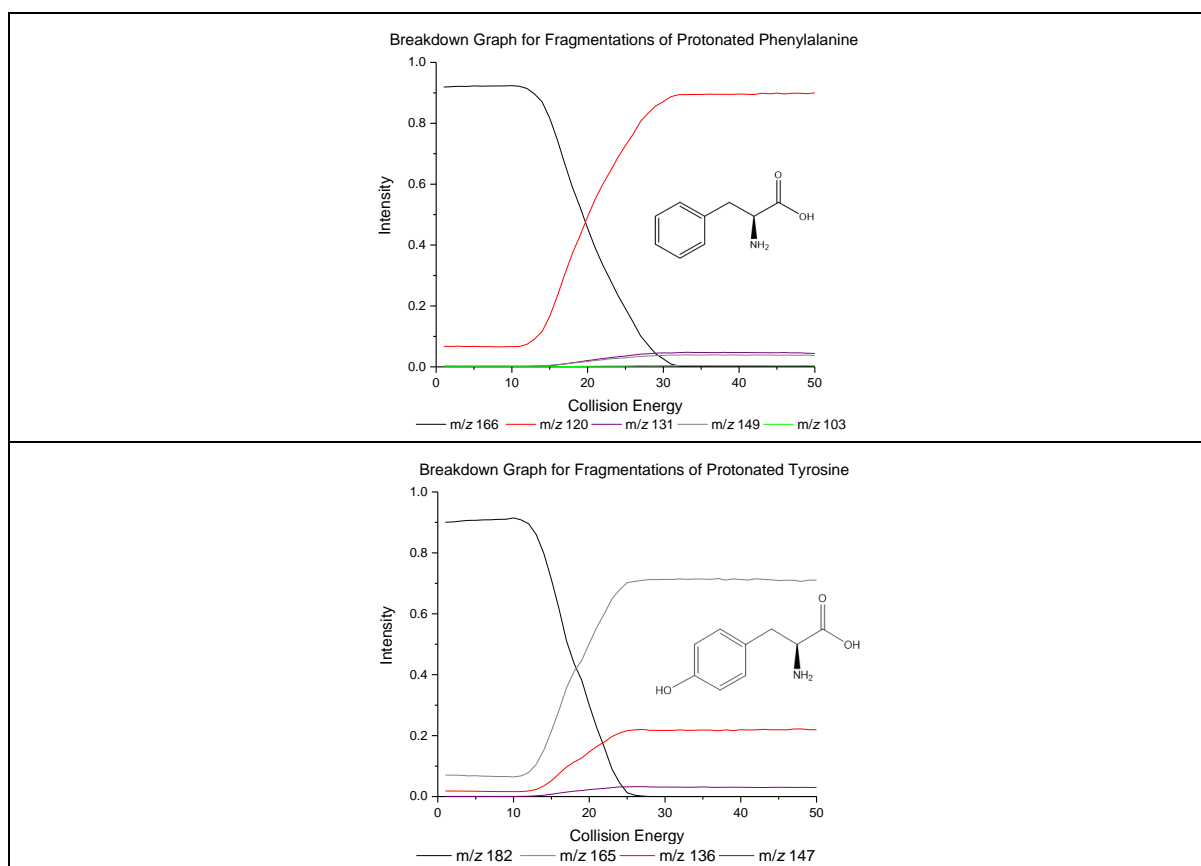
Figure 29 Energy breakdown graphs for cysteine (left) and methionine (right).

vii. Aromatic Group +ve CID MS/MS

Figure 30 shows energy breakdown graphs for phenylalanine, tyrosine, and tryptophan performed under CID. Overall, tryptophan exhibits the simplest breakdown graph, with only the loss of HCOOH (represented by the red line). This can be explained by the stability of the indole side chain. On the other hand, both phenylalanine and tyrosine have more fragment peaks. Phenylalanine has a dominant loss

of HCOOH and a much lower intensity loss of NH₃, with two more low-intensity secondary losses. The breakdown graph for tyrosine is very different, as it has a phenol group on its side chain. The loss of NH₃ is the most intense peak, with the loss of HCOOH occurring at about half its intensity. It also has a secondary loss of H₂O after the initial loss of NH₃.

For tryptophan, the energy ramp had to be performed manually. The breakdown graph is also based on manually extracted data. The breakdown graph shows only one fragmentation, the loss of NH₃. In contrast, as shown in the next section, the HCD energy breakdown graph of tryptophan is considerably more complicated. This distinguishes the HCD results from CID, which is a valuable finding for future studies.



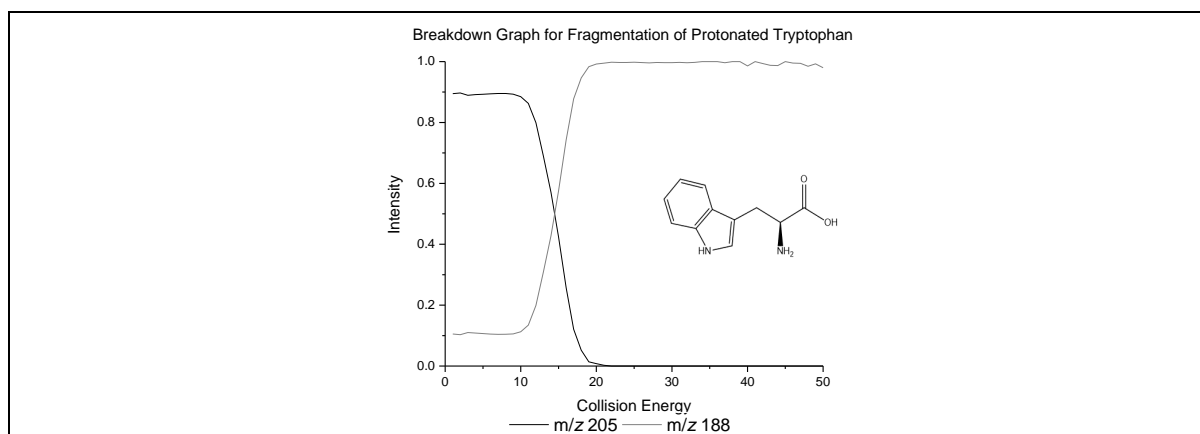


Figure 30 Energy breakdown graphs for phenylalanine (top), tyrosine (middle), and tryptophan (bottom).

viii. Aromatic Group +ve HCD MS/MS

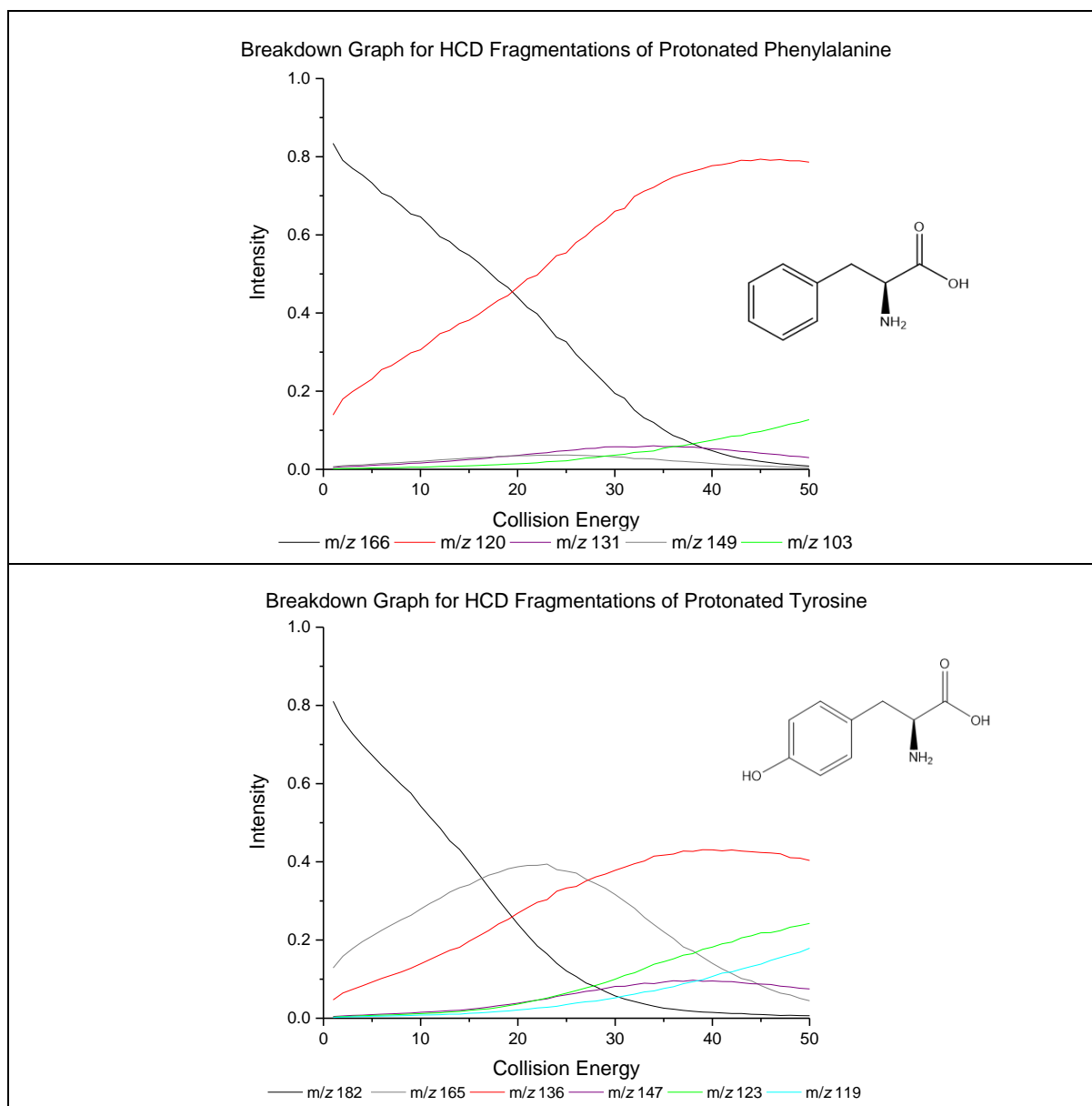
The energy breakdown graphs performed in positive ion mode HCD of the aromatic amino acids are presented in Figure 31. It is necessary to analyse and compare these graphs to those obtained in CID (Figure 21). As mentioned earlier, HCD can produce more product ions, and the breakdown graphs can be more complex than observed with CID.

The most obvious difference is the shape of the graph in HCD mode. In HCD, some lines are bell or Gaussian-shaped, for example, the loss of NH_3 for tryptophan (represented by the grey line). The reason for these differences can be explained by the fragmentation taking place externally in a multipole HCD collision cell.

In the HCD breakdown graph of phenylalanine, there are still four product ion peaks. The loss of formic acid again dominates as it does in CID. The main difference is in the shape of the lines and the increasing intensity of m/z 103 at higher collision energy. This secondary loss is less favourable in CID but seems more stabilised in HCD.

The breakdown graph of tyrosine in HCD produces two extra peaks at m/z 123 and m/z 119. The losses of NH_3 and HCOOH are still more abundant than the others, but unlike in CID, the intensity of the loss of NH_3 drops off quickly after the energy reaches 20 when the intensity of the loss of HCOOH starts to rise. As the loss of NH_3 continues to fall, all three secondary losses increase. The peak at m/z 123 is the most noticeable. HCD can, therefore, be regarded as a more favourable collision technique for multi-step fragmentations.

Tryptophan gives the most complex HCD breakdown graph. Overall, there are 8 fragmentations observed, instead of only 1 in CID. The most abundant route is still the loss of NH_3 , but it decreases dramatically when collision energy reaches about 22eV and the peak at m/z 146 starts to become dominant. All the other ions are also seen to increase their intensities after this point. Again, these ions mostly result from secondary losses (from peak m/z 188). To rationalise these observations, their fragmentation pathways will be analysed in detail.



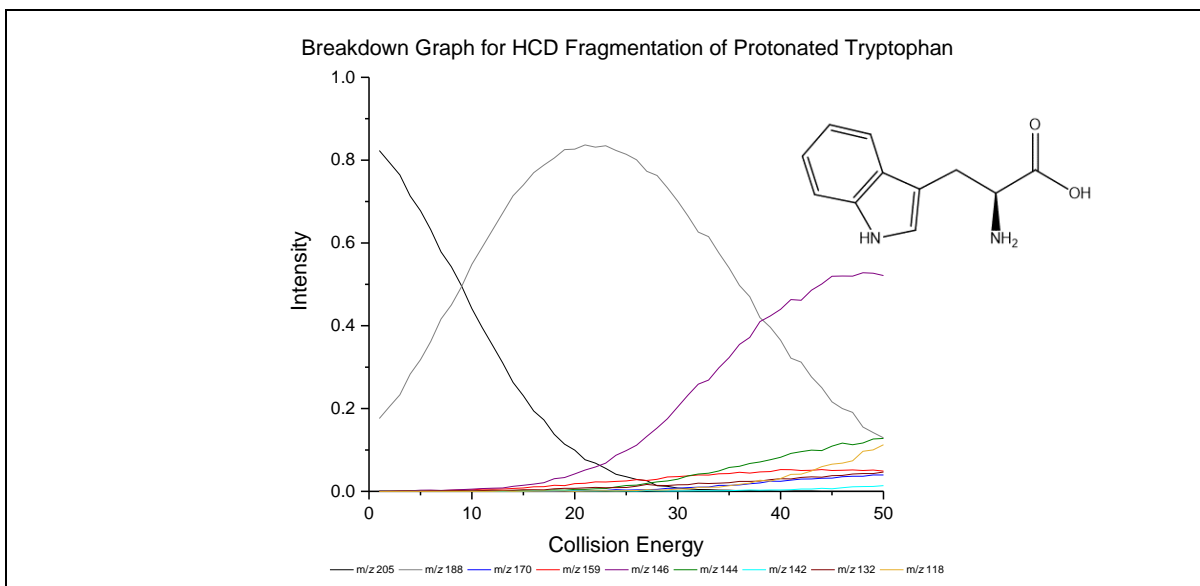


Figure 31 HCD energy breakdown graphs for phenylalanine (top), tyrosine (middle) and tryptophan (bottom).

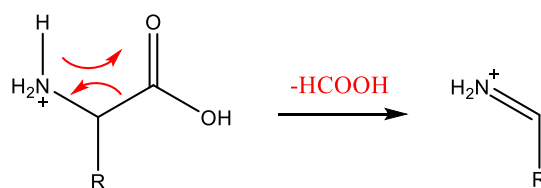
2.3.4. Mechanisms

2.3.4.1. Positive Ion Mode

i. Aliphatic Group +ve CID MS/MS

All aliphatic amino acids have the same loss of HCOOH (mass 46), except isoleucine, which has an additional NH₃ loss. Therefore, the mechanism shown here can be used as a general representation for all formic acid losses. It also applies to most other amino acids unless the loss occurs as a secondary loss or from the side chain (for the acidic amino acids).

The fragmentation mechanisms for leucine and isoleucine will be discussed in detail towards the end of this section. As they are structural isomers with different fragmentation pathways, it was necessary to analyse them further and discuss the results in more detail.



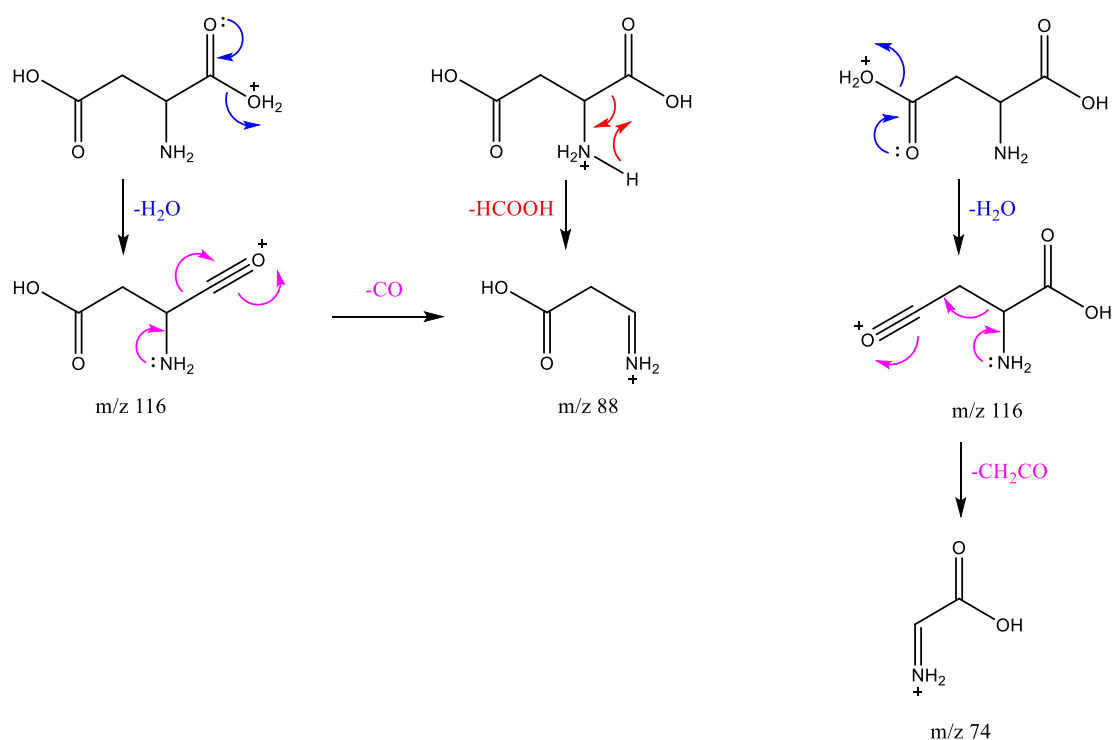
Scheme 2 A general representation of the fragmentation pathway for the loss of formic acid (HCOOH) in positive ion mode, R is the side chain.

Scheme 2 is a general representation of the fragmentation mechanism for the loss of formic acid (HCOOH). It is non-discriminating, especially with the aliphatic amino acids. The side chain is represented by 'R'. The nitrogen of the amino group is the protonation site and withdraws electron density to promote a bond breakage between the chiral carbon and carboxylic carbon. The products include a neutral formic acid molecule and the iminium ion, and these are the characteristic product ion combinations of amino acids in positive ion mode CID.

ii. Acidic Group +ve CID MS/MS

Aspartic and glutamic acids share the same neutral losses of HCOOH and H₂O. The suggested mechanisms are presented in Schemes 3 and 4. For both acidic amino acids, protonation can either occur on the nitrogen (which leads to a loss of HCOOH - mass 46) or on the OH group of the carboxylic acid (which leads to a simultaneous loss of CO -mass 28). Overall, there is a total loss of 46.

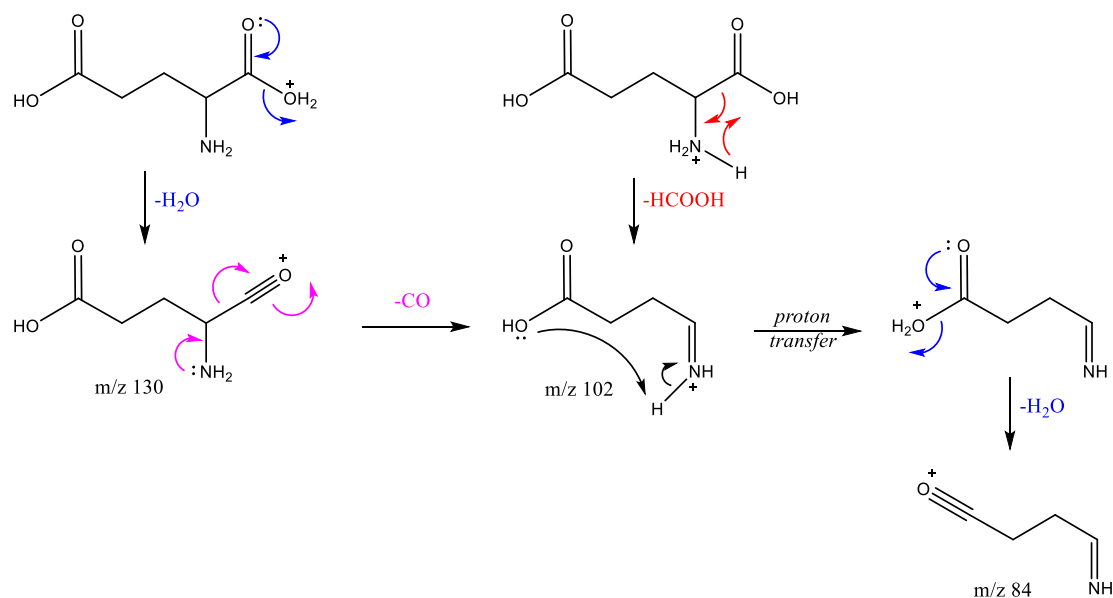
One difference exists in the mechanisms of aspartic acid (Asp) and glutamic acid (Glu). Asp gives a product ion 74 (see right-hand side of Scheme 3). It is initiated by protonation of the OH group of the side chain carboxylic acid. A water loss occurs, followed by a loss of ketene (mass 42), resulting in the stabilised immonium ion.



Scheme 3 Proposed mechanism for aspartic acid in positive ion mode. Product ions are m/z 116, m/z 88, and m/z 74. The colour scheme is consistent with the previous energy breakdown graphs.

Glutamic acid does not have ketene loss in its fragmentation pathways. Instead, it has a second H₂O loss after an initial 46 loss from the precursor ion. In Scheme 3, the proton transfers from the nitrogen to the oxygen of the OH group of the carboxylic side chain (see ion m/z 102). Product ion 102 then loses

H₂O to product ion m/z 84. A CO loss is expected to follow, but due to the limitation of lower mass product ions detection, it was not seen in the MS/MS spectra of glutamic acid.



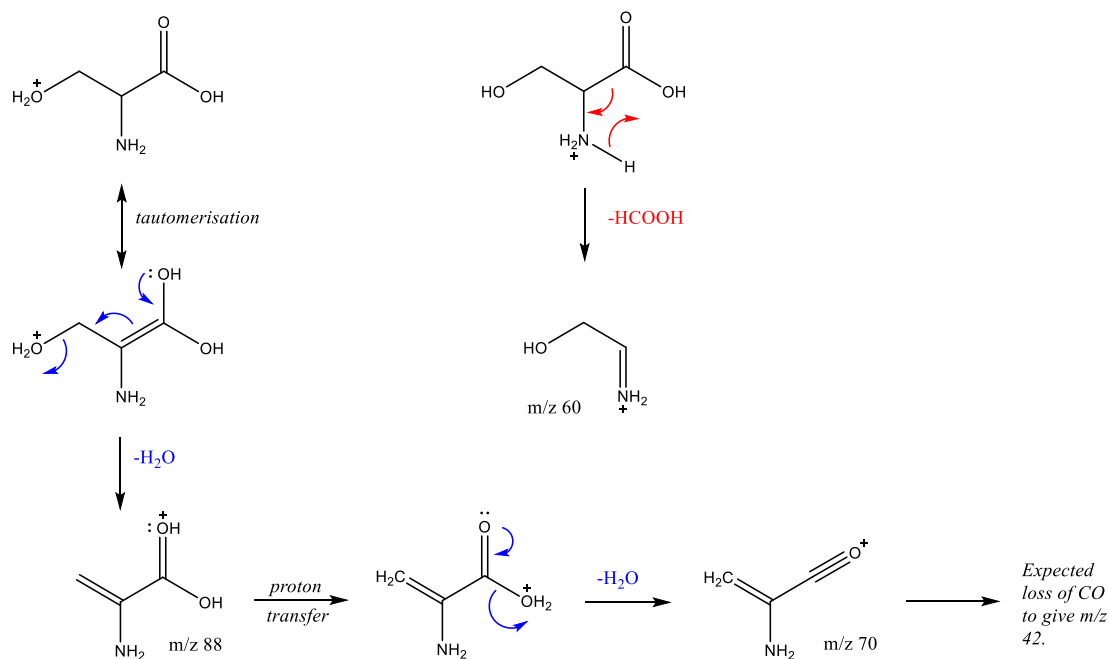
Scheme 4 Proposed mechanism for glutamic acid in positive ion mode. Product ions are m/z 130, m/z 102, and m/z 84.

iii. Hydroxylic Group +ve CID MS/MS

Serine and threonine have the same neutral losses in their tandem mass spectra. They both lose HCOOH or have two consecutive losses of water. Hence, only the mechanism for serine is provided (Scheme 5). Amongst the possible protonation sites, the hydroxyl group of the side chain of serine and the nitrogen on the amine are the most likely. Therefore, there are two possible routes resulting from either of these protonation sites, each leading to a different fragmentation pathway.

The right side of Scheme 5 is the classic loss of the HCOOH route. The left side shows initial protonation of the alcohol side chain, followed by a tautomerisation, which leads to the initial loss of H₂O. This is followed by a proton transfer to the OH group in the carboxylic acid. The positive charge can now shift to the carboxylic alcohol, resulting in the second loss of H₂O. It is important to note that the structure of the product ion at m/z 70 has a CO ready to fragment. It is, therefore, reasonable to assume that this ion will go on further fragmentation to form an ion at m/z 42. Unfortunately, this is outside of the mass range of the Orbitrap, and it is impossible to detect it. However, this route has been

confirmed by theoretical calculations using Gaussian 09. The relevant data will be presented in a later section.

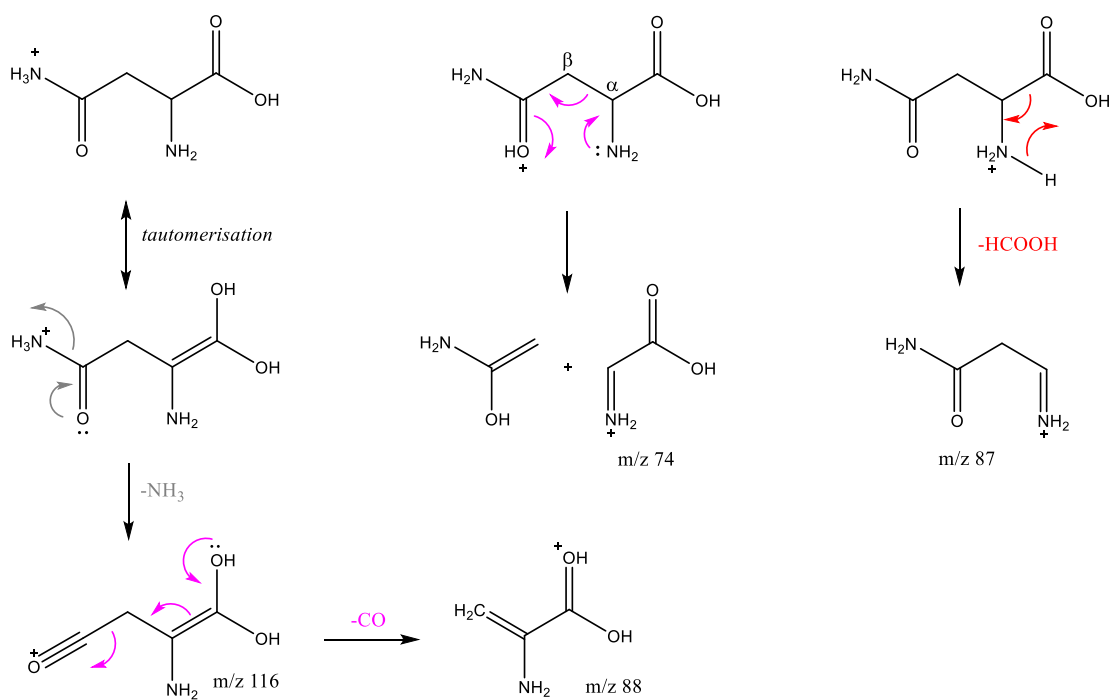


Scheme 5 Proposed mechanism for serine in positive ion mode. Product ions are m/z 88, m/z 70, and m/z 60.

iv. Amidic Group +ve CID MS/MS

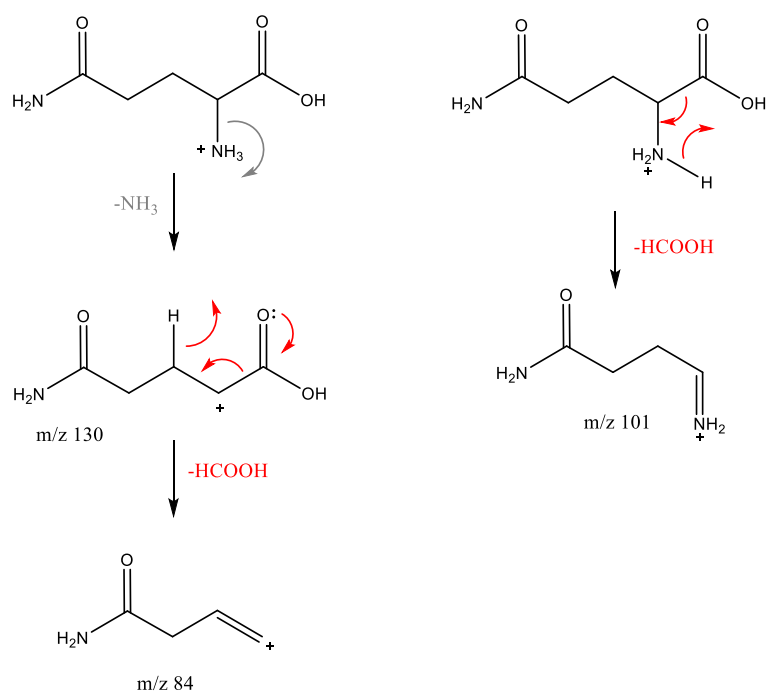
Asparagine and glutamine have a few differences in their fragmentation pathways. The mechanisms presented here are examined later by computational methods.

There is more than one possible protonation site with the amidic amino acids. Gaussian modelling reveals that the most energetically favourable site is usually the amine nitrogen. Because of various possible starting structures, three fragmentation pathways have been drawn for asparagine (Scheme 6). The one on the left starts with the amide nitrogen being the protonation site. It is followed by a tautomerisation, which leads to a loss of NH₃. This then goes on to fragment by loss of CO to give the final ion at m/z 88. The mechanism in the middle starts with the protonation site on the oxygen of the carboxamide, and it leads to a bond breakage between the α and β carbon, resulting in ion m/z 74 and a neutral loss of 59. Lastly, the mechanism on the right is the classic loss of HCOOH.



Scheme 6 Proposed mechanism for asparagine in positive ion mode. There are multiple possible protonation sites. Product ions are m/z 116, m/z 88, m/z 87, and m/z 74.

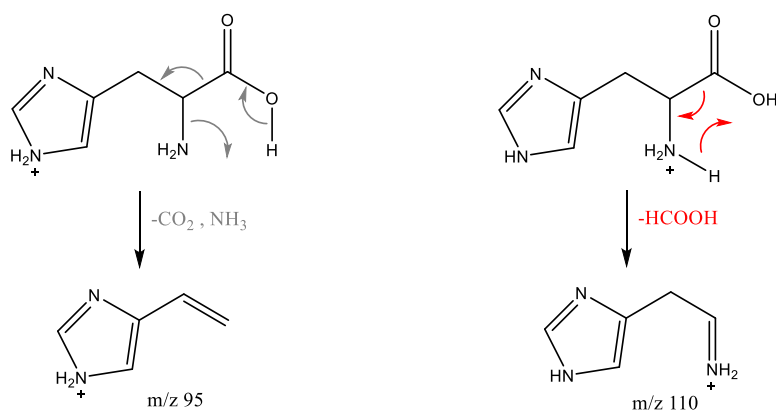
Glutamine has three product ions (Scheme 7). The mechanism provided here includes the classic loss of NH₃ via a simple bond stretch. It is then followed by a loss of HCOOH to produce ion m/z 84. Loss of HCOOH can also result from the protonated precursor ion, shown in the mechanism on the right. In comparison, the fragmentation pathways for glutamine are simpler. The absence of one carbon in asparagine in its side chain has facilitated the production of one additional fragmentation pathway. This demonstrates the side chain's effect on amidic amino acids' fragmentation pathways.



Scheme 7 Proposed mechanism for glutamine in positive ion mode. Product ions are m/z 130, m/z 101, and m/z 84.

v. Basic Group +ve CID MS/MS

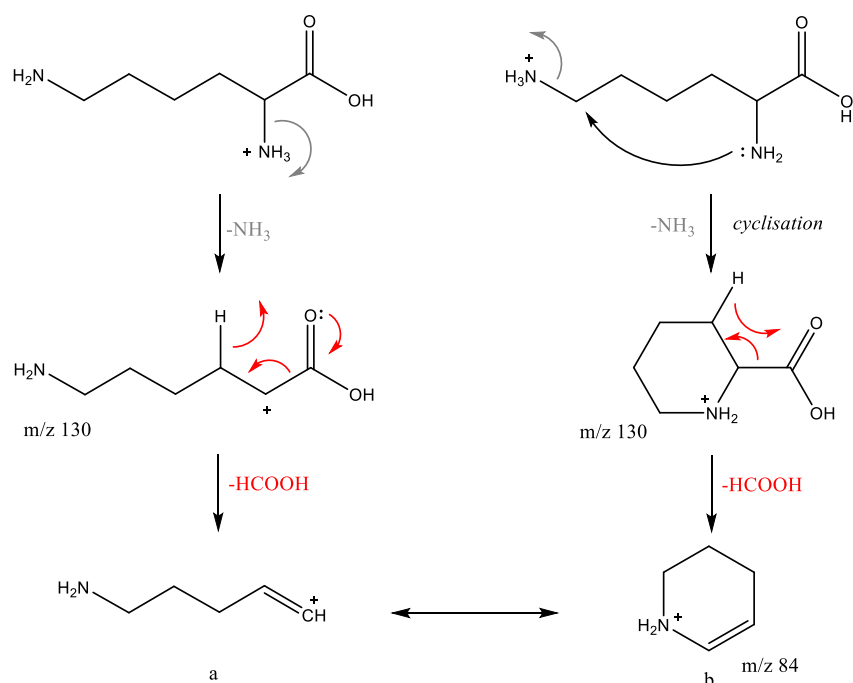
Among the basic group of amino acids, histidine has the simplest fragmentation pathway; therefore, it is discussed first (Scheme 8). The route on the left is a charge remote fragmentation. Protonation occurs on the nitrogen of the imidazole side chain instead of the amine. The positive charge is not involved in the actual fragmentation. Both carbon dioxide and ammonia are lost simultaneously, with the resulting product ion maintaining the positive charge on the same nitrogen. The route on the right is proposed to represent the classic HCOOH loss.



Scheme 8 Proposed mechanism for histidine in positive ion mode. Product ions are m/z 110 and m/z 95.

Different from histidine, lysine does not have an initial HCOOH loss in its MS/MS spectrum, although it does have an H₂O loss, which is assumed to be the result of protonation on the OH of the carboxylic acid group. Therefore, the product ion m/z 129 is not shown in Scheme 9. With two amines, it would be expected that lysine has an abundant loss of NH₃. Both amines can be protonated. Theoretical calculations revealed that the two protonated structures have very similar energies, with a difference of 11 KJ/mol in energy. Protonation on the first amine closest to the carboxylic acid is slightly more favourable due to carbonyl oxygen stabilising the positive charge.

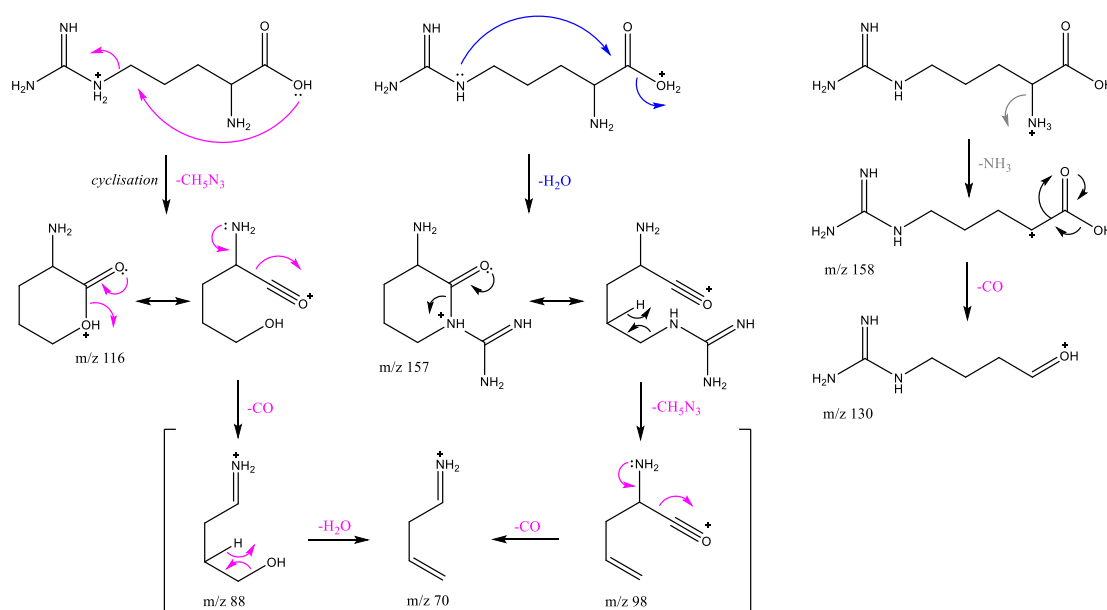
Loss of HCOOH only occurs after the NH₃ loss (Scheme 9). Two routes have been proposed. The mechanism on the left starts with a simple bond stretch to lose NH₃. It then goes on to lose HCOOH to form product ion m/z 84. The mechanism on the right shows protonation on the second amine. Cyclisation occurs when ammonia is lost from the side chain. Product ion m/z 130 can then lose a HCOOH giving ion b at m/z 84, which is product ion a's cyclic resonance structure. It is assumed that the mechanism on the left is energetically more favourable, as the cyclisation step on the right could require more energy.



Scheme 9 Proposed mechanism for lysine in positive ion mode. Product ions are m/z 130 and m/z 84.

Arginine has additional product ions compared to both lysine and histidine. Again, as with lysine, there is no HCOOH loss, probably due to the presence of the guanidine group on the side chain. However, a distinctive loss of 59 (corresponding to guanidine) in its MS/MS spectrum results in a product ion at m/z 116 (Scheme 10). The formation of a six-membered ring is one of the driving forces for the loss of the guanidine group. In addition, the loss of NH_3 can be followed by a loss of 28, and an OH migrates onto the positively charged carbon, resulting in product ion m/z 130, which has not been observed with the other basic amino acids. Product ion m/z 157 is caused by a loss of H_2O , and, as with other amino acids, it is proposed that the carboxylic acid OH group is the protonation site. However, in this case, a cyclisation mechanism is suggested to help stabilise the charge. Possible evidence to support this is that there is no further CO loss after the first water loss. A simple loss of CO from m/z 157 would result in an ion at m/z 129, but this is not observed.

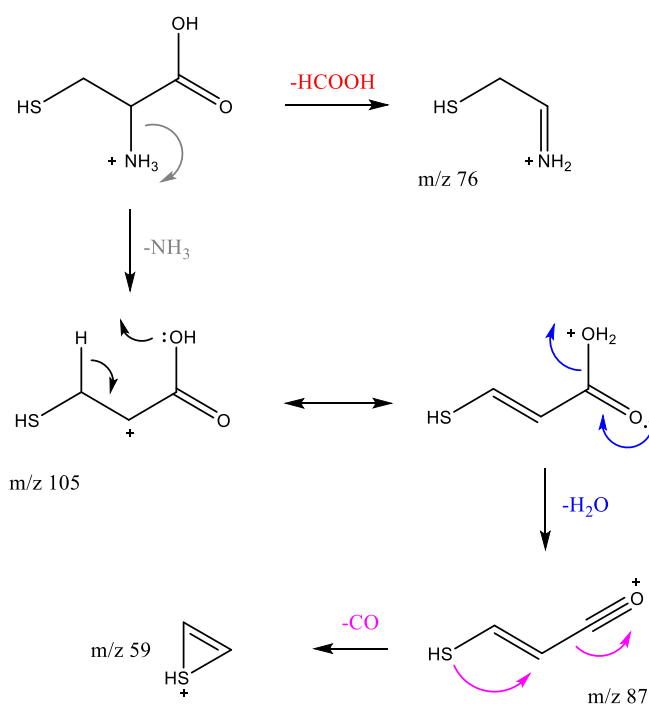
In addition, product ion m/z 116 is believed to move its positive charge on the oxygen, leading to a CO loss to produce ion m/z 88. An H_2O loss can then be followed by a hydride shift to produce ion m/z 70. It is also suggested that ion m/z 157 can have a resonance structure in which the ring opens and loses its guanidine group. This product ion, m/z 98, uses the lone pair on the nitrogen to create a $\text{C}=\text{N}$ double bond, pushing off a CO and producing the same product ion m/z 70. These ions are drawn in brackets, as their peak intensities are low in the MS/MS spectra. It provides an opportunity to examine them further in MS^3 in future studies.



Scheme 10 Proposed mechanism for arginine in positive ion mode. Product ions are m/z 158, m/z 157, m/z 130 and m/z 116. Product ions in brackets m/z 98, m/z 88 and m/z 70 are observed in low intensities.

vi. Sulfur Containing Group +ve CID MS/MS

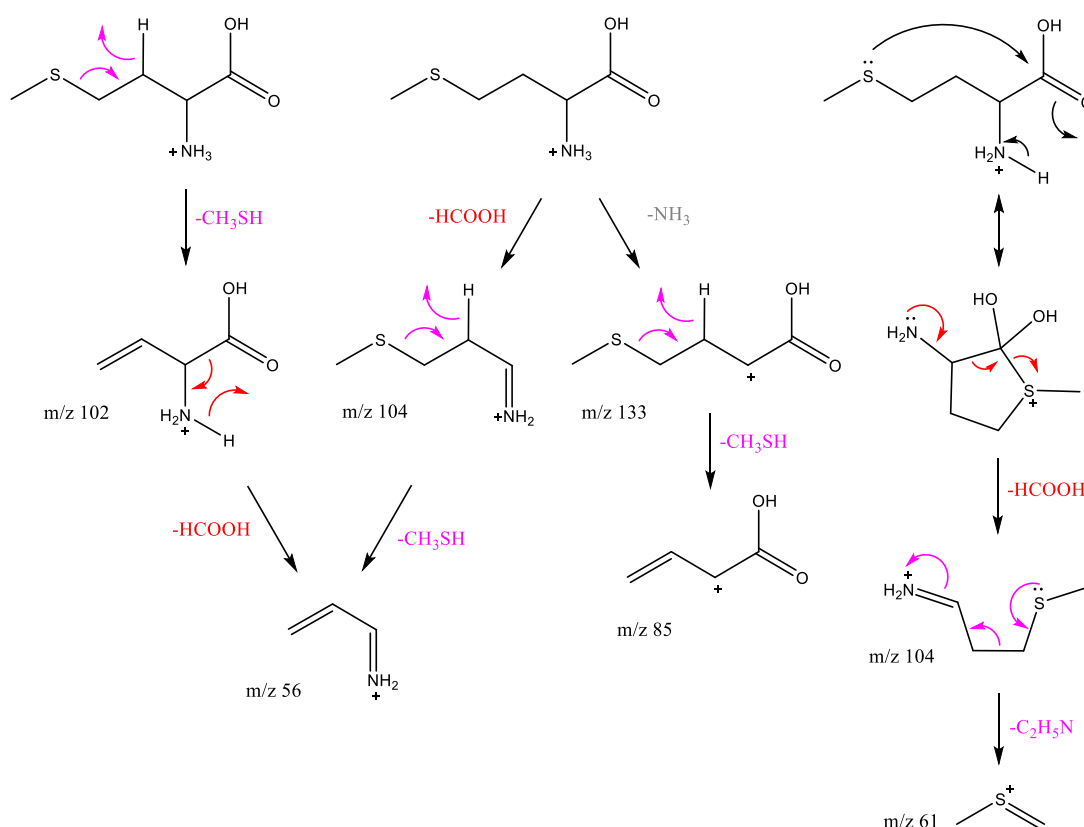
Cysteine and methionine both have sulfur containing side chains, but the former has a thiol, and the latter has an S-methyl thioether. Therefore, their MS/MS spectra and fragmentation mechanisms vary slightly. Cysteine has the classic loss of HCOOH to give product ion m/z 76. The other route is the loss of NH_3 to give ion m/z 105. A proton transfer occurs to move the positive charge onto the hydroxyl oxygen, promoting a second loss of H_2O . The lone pair on the sulfur acts like a nucleophile and attacks the double-bonded carbon, which facilitates a further loss of CO to produce ion m/z 59 (Scheme 11). Ion m/z 59 is proposed to adopt the cyclic structure of a thiirene, with a positive charge on the sulfur, to provide stability.



Scheme 11 Proposed mechanisms for cysteine in positive ion mode. Product ions are m/z 105, m/z 87, m/z 76 and m/z 59.

The presence of the S-methyl thioether group in methionine makes a neutral loss of CH_3SH inevitable (product ion m/z 102, Scheme 12 top left). The first route suggested is for this loss, followed by a further loss of HCOOH to produce product ion m/z 56. It can also be produced by first losing HCOOH (product ion m/z 104) and then CH_3SH . In addition, an initial loss of NH_3 is also observed in the MS/MS spectrum, leading to a loss of CH_3SH to produce ion m/z 85 (shown in the middle). The fourth route is when the lone pair on the sulfur attacks the carboxylic carbon. A five-membered ring is formed. The lone pair on

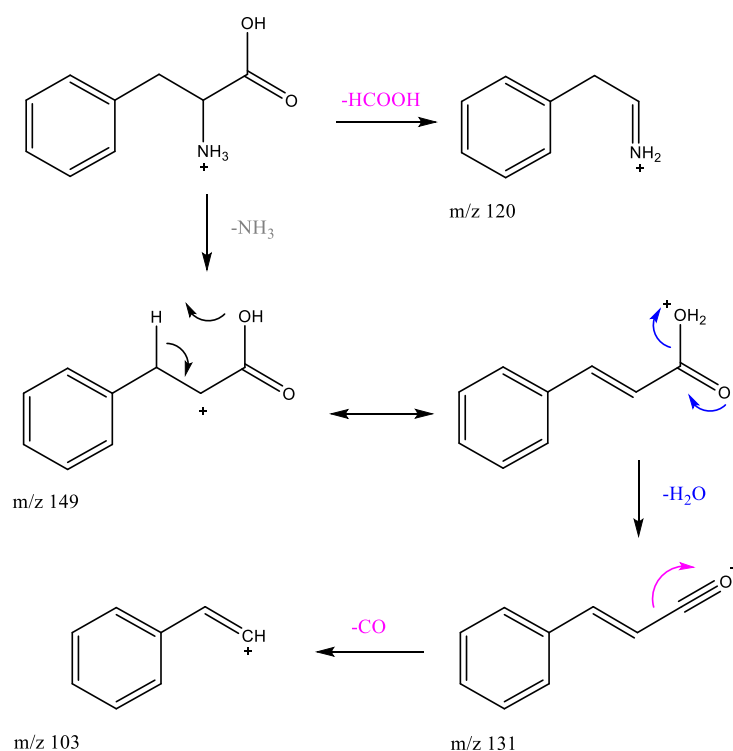
the nitrogen, combined with the electron withdrawing action of the positively charged sulfur, promotes a loss of HCOOH to give ion m/z 104. This can further lose C_2H_5N to produce ion m/z 61. Overall, both the MS/MS spectra and energy breakdown graph of methionine have revealed that a loss of HCOOH is very favourable. Having two fragmentation pathways leading to the same product ion m/z 104 is very likely to be the reason.



Scheme 12 Proposed mechanism for methionine in positive ion mode. Product ions are m/z 133, m/z 104, m/z 102, m/z 85, m/z 61 and m/z 56.

vii. Aromatic Group +ve CID MS/MS

Phenylalanine has a highly abundant loss of HCOOH and comparatively lower abundance for the other losses listed in Scheme 13. It is proposed that after the initial loss of NH_3 , product ion m/z 149 continues to fragment by loss of H_2O via a proton transfer to the OH group. The lone pair on the oxygen then donates its electrons to form a triple bond (product ion m/z 131, second structure), leading to a valid route for a CO loss. This gives the final product ion m/z 103. Compared to the more direct route for the loss of HCOOH, this route is less favourable, as demonstrated in its CID energy breakdown graph in the previous section.

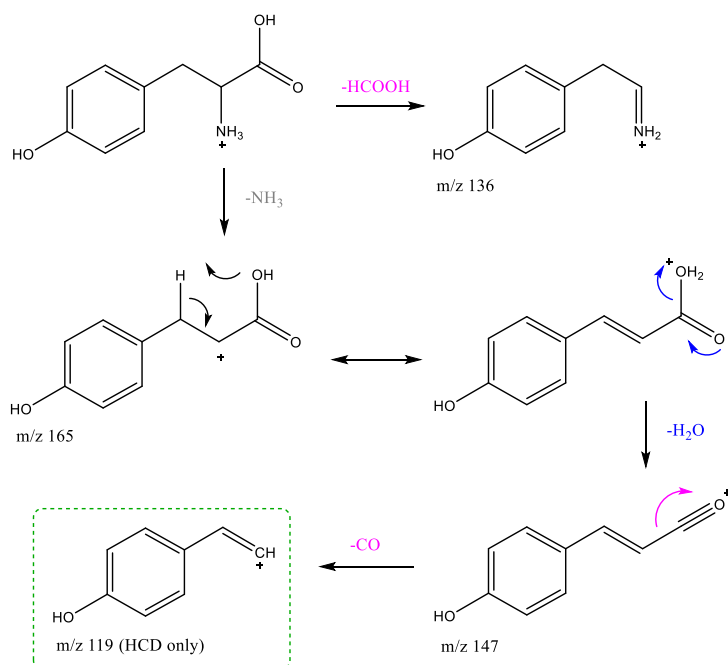


Scheme 13 Proposed mechanism for phenylalanine in positive ion mode. Product ions are m/z 149, m/z 131, m/z 120 and m/z 103.

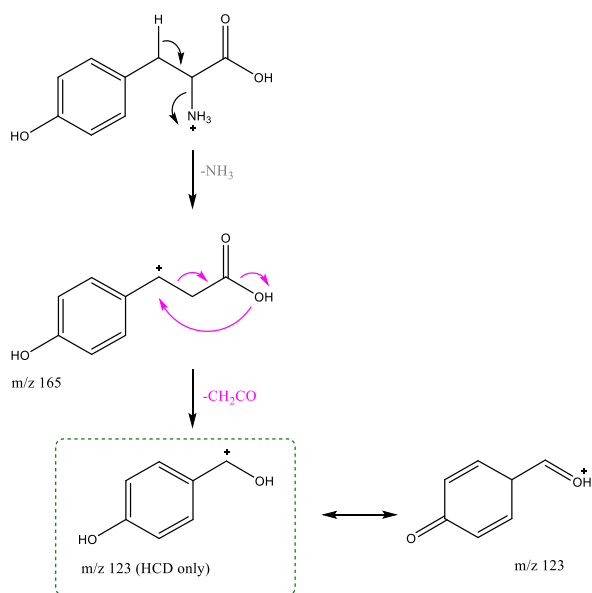
Tyrosine has a similar fragmentation route to phenylalanine. However, the breakdown graph shows distinctive differences in the peak intensity for the loss of HCOOH versus the loss of NH₃. For tyrosine, the loss of NH₃ is more intense than HCOOH. The reason for this is shown in Scheme 14, where there is an additional possible mechanism for the loss of NH₃. Instead of a simple bond stretch leading to the loss of NH₃, it is proposed that it can also be lost by a 1,2-hydride shift, resulting in a carbocation with the positive charge right next to the carbon closest to the phenol ring. The various resonance structures of this cation reveal that it is further stabilised by transferring the positive charge onto the oxygen (see Scheme 14).

The original route (same as for phenylalanine) is shown in Scheme 14, where product ion m/z 147 results from water loss from m/z 165. Further fragmentation produces ions at m/z 119 by a CO loss (this ion is only observed in HCD). These two possible routes for the loss of NH₃ explain the high abundance of product ion m/z 165 compared to m/z 136 (loss of HCOOH). In addition, as mentioned earlier in this chapter, HCD of the aromatic amino acids results in more fragmentation. In the MS/MS spectrum of tyrosine, two more product ions have been observed: ions at m/z 123 and m/z 119.

Tryptophan only gives a peak for the loss of NH_3 in CID, therefore, the mechanism for this classic route will not be shown here. However, tryptophan produces more interesting product ions in HCD, which will be discussed in detail in the next section.



(a)



(b)

Scheme 14 (a) Proposed mechanism for tyrosine in positive ion mode. Product ions are m/z 165, m/z 147, m/z 136, and m/z 119. (b) The additional mechanism proposed for tyrosine for loss of NH_3 . Product ions are m/z 165 and m/z 123.

viii. Aromatic Group +ve HCD MS/MS

The previous section shows energy breakdown graphs for aromatic amino acids in HCD. They reveal extensive fragmentation, especially in the case of tryptophan. In this section, mechanisms for these additional fragmentations are presented and discussed. Phenylalanine gives the same fragmentation in CID and HCD and will not be discussed further. The focus of this part will be on tyrosine and tryptophan.

In HCD, tyrosine has two extra product ions at m/z 123 and m/z 119. The mechanism for ion m/z 119 has already been shown in Scheme 14(a) as the product of a CO loss from ion 147. However, in HCD, the peak intensity of m/z 119 has increased considerably with higher collision energy. This shows that in HCD, multiple neutral losses are more accessible than in CID.

Scheme 14 shows the mechanism for the generation of product ion 123. It starts with a 1,2-hydride shift and loss of NH_3 to produce ion m/z 165. Then, the OH group migrates onto the carbon with a positive charge, promoting simultaneous bond dissociation, leading to a neutral loss of ethenone ($\text{H}_2\text{C}=\text{C}=\text{O}$) and product ion m/z 123 generations. This mechanism is more complex, requiring more bond dissociations and atom migrations, which probably explains why it is only observable in HCD. The migration of the hydroxyl ion from the carboxylic group is also supported in the literature.¹⁹⁹ Their study stated that for protonated tyrosine and tryptophan, the elimination of NH_3 has the lowest energy threshold, whereas for protonated phenylalanine, the elimination of HCOOH is lower. This correlates to our observations in energy breakdown graphs for tyrosine and tryptophan.

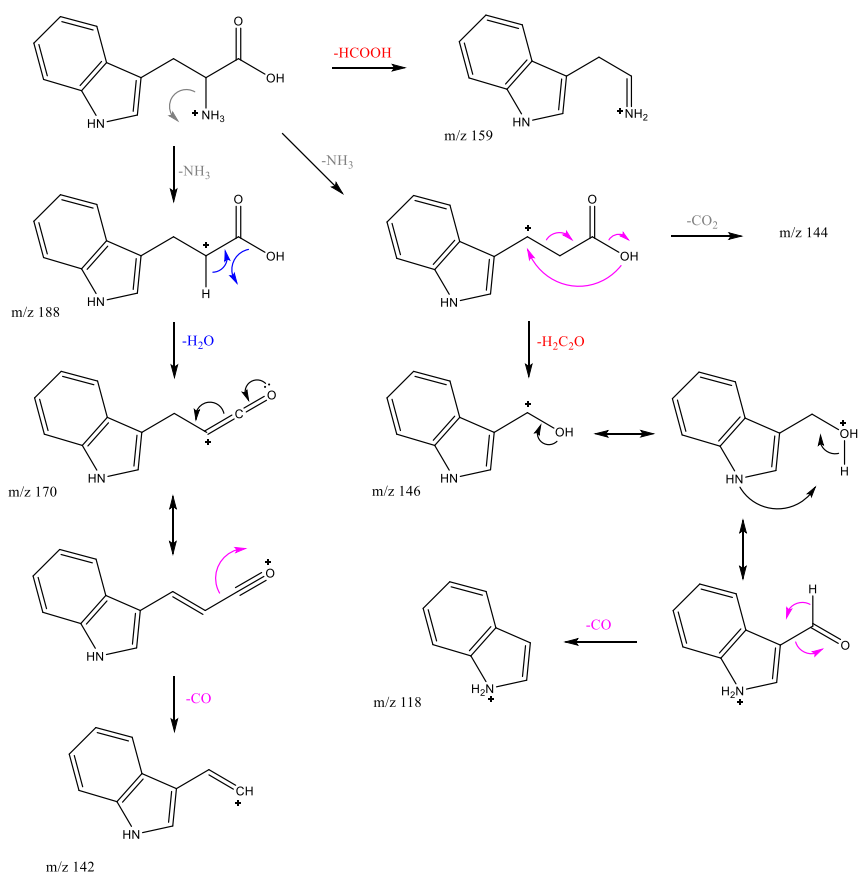
As mentioned in the previous section, the HCD-MS/MS of tryptophan reveals more complex and extensive fragmentation. To confirm the precursor ion of these fragmentations, an MS^3 analysis has also been performed. The following section discusses the mechanism for the ions observed in the MS^3 spectrum and HCD-MS/MS.

Scheme 15 is the proposed mechanism for the fragmentation of tryptophan observed in CID and HCD, as well as MS^3 spectra. Interestingly, it shares the fragmentation pathways with phenylalanine and tyrosine. Firstly, it has the classic loss of HCOOH , then two routes for the loss of NH_3 . The mechanism on the left starts with a simple C-N bond stretch, as seen in phenylalanine and tyrosine. This is followed

by water loss and a loss of CO. The other possible route is initiated by a 1,2-hydride shift and a loss of NH_3 , followed by two possible pathways leading to either a CO_2 loss or a loss of 42 (the latter has also been seen with tyrosine). The resulting product ion from the 42 loss can be in resonance with a more stabilised structure, allowing further CO loss.

The complexity of the fragmentation pathways of tryptophan in

Scheme 15 could be a valid reason why these product ions are not observed in CID-MS/MS. Interestingly, even with an indole side chain, tryptophan still shares a considerable amount of its fragmentation with phenylalanine and tyrosine. However, there are differences resulting from the side chain. The increased stability given by the indole side chain makes tryptophan capable of having more product ions in HCD. From this part of the study, it is important to appreciate the advantages of also using HCD to complement the results obtained with CID, and it should be applied more widely in related areas in the future when additional fragmentation routes may be required.

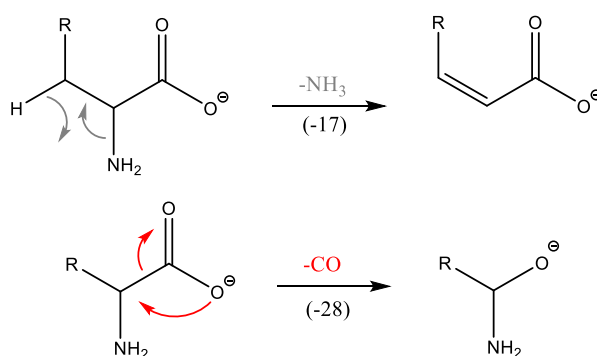


Scheme 15 Proposed mechanism for tryptophan in positive ion mode. Product ions are m/z 188, m/z 170, m/z 159, m/z 146, m/z 142 and m/z 118.

2.3.4.2. Negative Ion Mode

The mechanisms for the 20 amino acids in negative ion CID-MS/MS are investigated in this section. Due to the mass limitations, there are no spectra for alanine and glycine, and these two are excluded from the discussion. The most common neutral loss is NH_3 ; the second is H_2O , followed by CO . As a result, duplicate mechanisms for each amino acid will not be proposed. The extent of variation in the fragmentation pathways within the same side chain functionality group is generally smaller than that observed in positive mode CID. Therefore, their mechanisms are less complicated than those elucidated in the positive ion mode.

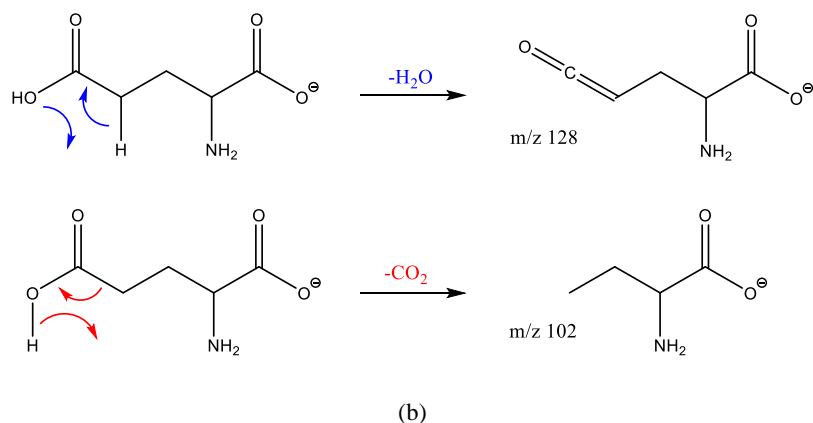
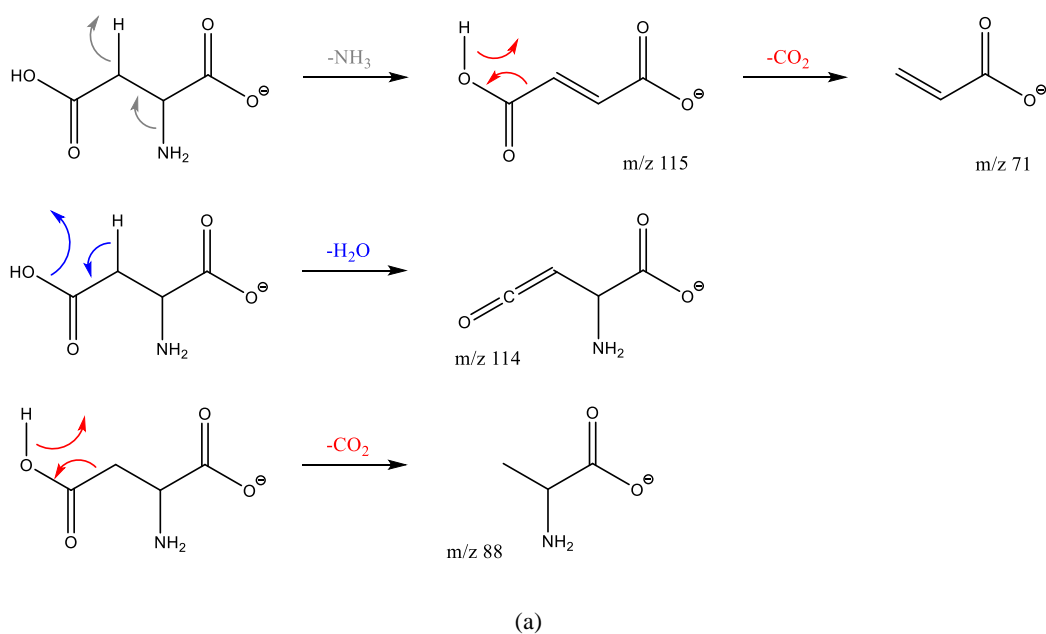
Scheme 16 represents two suggested fragmentation pathways in negative ion mode for losses of NH_3 and CO . The top route is for the loss of NH_3 , which results from a charge remote fragmentation. The negative charge on the oxygen is not involved and remains on the same atom throughout the mechanism. The bottom route is for the loss of CO . In this case, the O^- group migrates onto the chiral carbon, leading to a neutral loss of CO . These two routes are readily observed in negative ion MS/MS spectra of all the target amino acids. Therefore, the mechanisms presented here generally use 'R' as the side group. These routes will not be discussed further.



Scheme 16 A general representation of the fragmentation pathways for the loss of NH_3 (top) and CO (bottom) in negative ion mode, R is the side chain.

i. Acidic Group -ve CID MS/MS

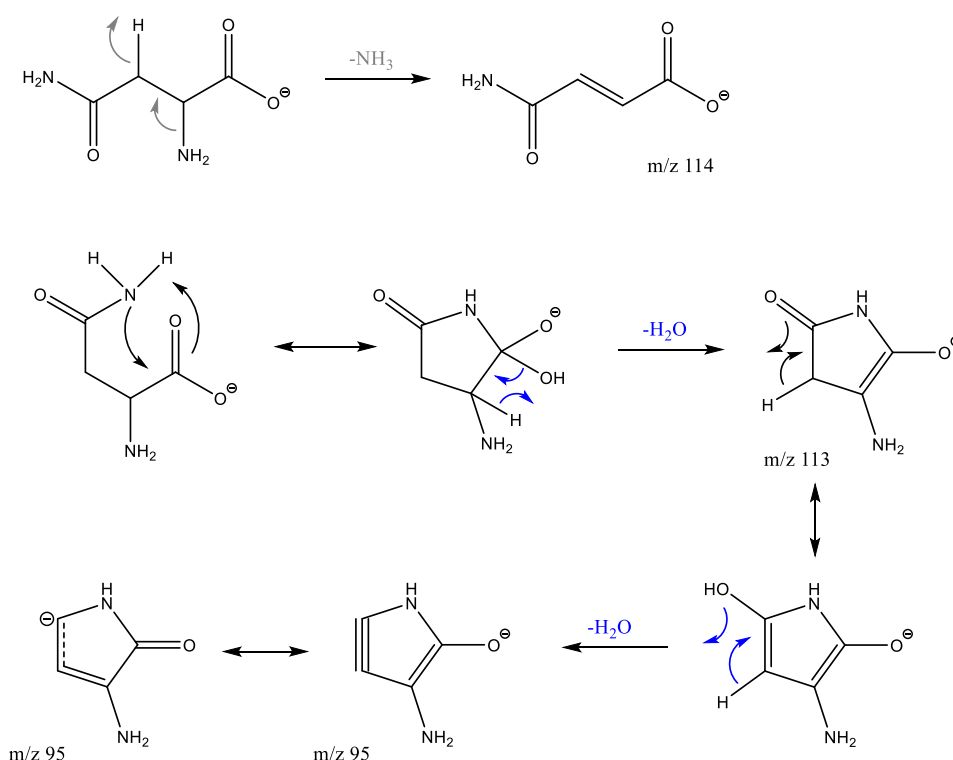
Aspartic and glutamic acid share common neutral losses of H₂O and CO₂, as shown in Scheme 17. However, glutamic acid does not lose NH₃ in its negative ion mode MS/MS spectrum, unlike aspartic acid. A possible reason for this could be that compared to a neutral loss of NH₃, and it is more favourable for glutamic acid to lose a water molecule. By examining Scheme 17 (a) of aspartic acid and (b) glutamic acid, it is obvious that changes in their side chains influence their fragmentation pathways in negative ion mode as well as in the positive ion mode.



Scheme 17 (a) Proposed mechanism for aspartic acid in negative ion mode. Product ions are m/z 115, m/z 114, m/z 88 and m/z 71. (b) Proposed mechanism for glutamic acid in negative ion mode. Product ions are m/z 128 and m/z 102.

ii. Amidic Group -ve CID MS/MS

Asparagine and glutamine have the same neutral losses in their MS/MS spectra in negative ion mode. Both amino acids have losses of NH_3 followed by two consecutive losses of H_2O . The proposed mechanism for asparagine is given in Scheme 18. Glutamine is assumed to follow the same fragmentation pathway and is therefore not shown. It is interesting to note that although they have different fragmentation pathways in positive ion CID-MS/MS, they have the same neutral losses and fragmentation pathway here in negative ion mode.



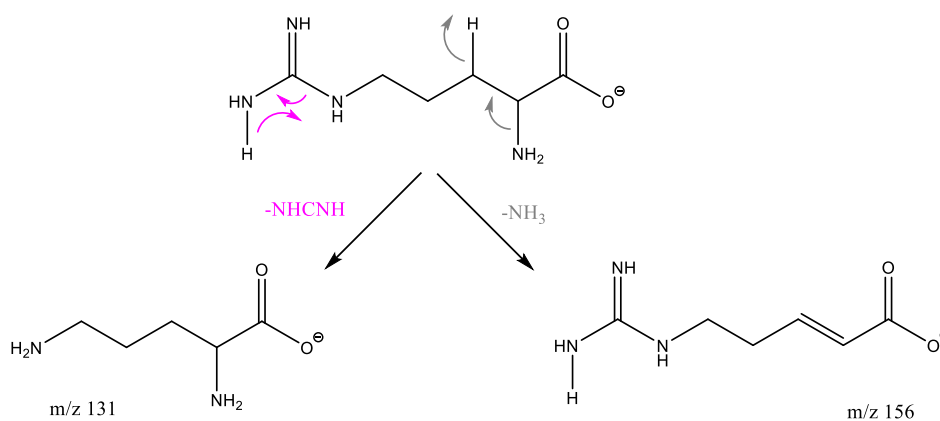
Scheme 18 Proposed mechanism for asparagine in negative ion mode. Product ions are m/z 114, m/z 113 and m/z 95.

Scheme 18 shows a cyclisation step with the amidic nitrogen of the side chain attacking the carbon of the carboxylic acid. The first water loss is followed by a resonance structure transformation, promoting the second water loss. The final structure of product ion at m/z 95 has been theoretically optimised by Gaussian 09.

iii. Basic Group -ve CID MS/MS

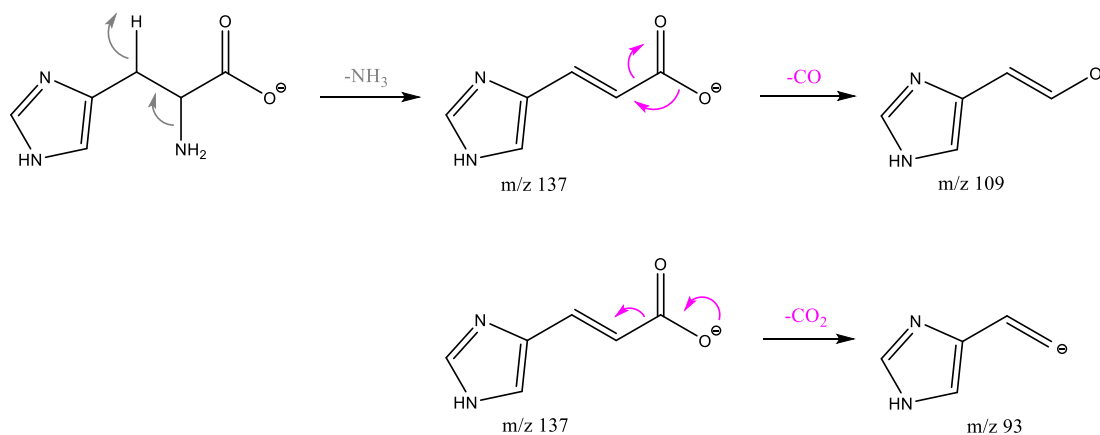
Although arginine and histidine belong to the basic group, they have very different side chains resulting in their negative ion CID-MS/MS being very different. Their mechanisms are therefore drawn separately to highlight the differences. Generally, the side chain variation in negative ion mode MS/MS has little effect on the fragmentation pathways. Arginine and histidine demonstrate examples of exceptions to that observation.

Scheme 19 is the proposed mechanism for arginine. The route on the left is a loss of 42, which is unique to arginine due to the presence of its guanidino side chain. On the left is the loss of NH_3 by charge remote fragmentation.



Scheme 19 Proposed mechanism for arginine in negative ion mode. Product ions are m/z 156 and m/z 131.

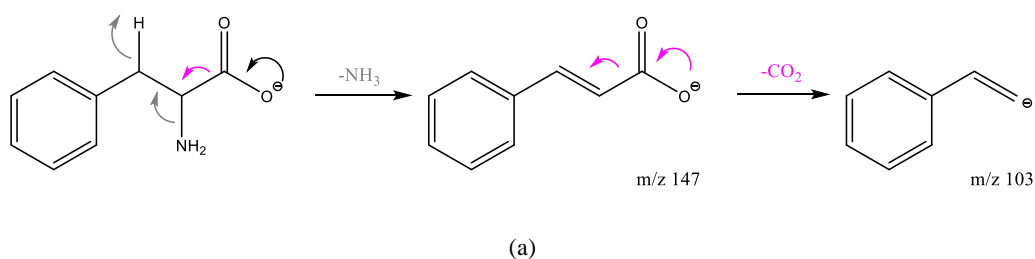
Histidine has a completely different fragmentation pathway (Scheme 20). First, it lost NH_3 , then product ion 137 loses CO by migrating the negatively charged oxygen onto the double bonded carbon, resulting in ion 109. In addition, product ion 137 can also sustain a CO_2 loss, producing ion 93. The extra neutral losses are likely to be supported by the delocalising effect from the imidazole side chain. Compared to arginine, histidine has two more secondary losses. It is also worth mentioning that in positive ion CID, arginine has more product ion peaks than histidine. In contrast, histidine has more product ion peaks than arginine in the negative ion mode.

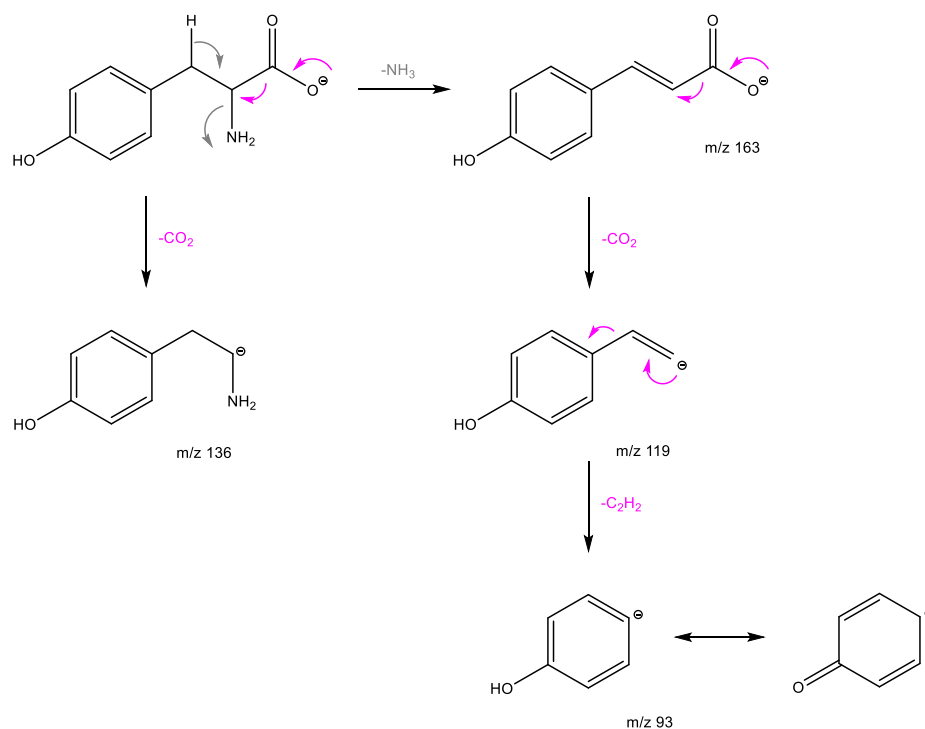


Scheme 20 Proposed mechanism for histidine in negative ion mode. Product ions are m/z 137, m/z 109 and m/z 93.

iv. Aromatic Group -ve CID MS/MS

In negative ion CID-MS/MS, phenylalanine has the least product ions. Tryptophan and tyrosine have the same fragmentation pathways, with two additional product ions compared to phenylalanine. Compared with the data obtained in positive ion CID, phenylalanine has fewer product ions in negative ion (four neutral losses in positive ion mode compared to two neutral losses in negative ion CID). On the other hand, there are more fragmentation peaks for both tryptophan and tyrosine in negative ion mode compared to positive ion mode. However, when the mass spectra obtained in positive ion HCD-MS/MS are considered, HCD still gives the greatest number of product ions for these two aromatic amino acids.





(b)

Scheme 21 (a) Proposed mechanism for phenylalanine in negative ion mode. Product ions are m/z 147 and m/z 103. (b) Proposed mechanism for tyrosine in negative ion mode. Product ions are m/z 163, m/z 136, m/z 119 and m/z 93.

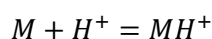
Scheme 21 illustrates the proposed mechanisms for phenylalanine and tyrosine. Since tryptophan has the same neutral losses as tyrosine, it is not included here. Phenylalanine has an initial loss of 17, followed by a loss of 44 (Scheme 21). Tyrosine has two initial neutral losses from the precursor ion (NH_3 and CO_2), giving product ions m/z 163 and 136 (Scheme 21). A secondary loss of CO_2 is observed to give product ion m/z 119. A further unique loss of acetylene (HCCH , mass 26) is observed with tyrosine and tryptophan. It is proposed that the reason for this could be that when compared to phenylalanine (phenyl group on the side chain), tyrosine (phenol side chain) and tryptophan (indole side chain) have more electronegative atoms (oxygen and nitrogen, respectively) in their side chains, resulting in increased stabilisation of the final anion m/z 93 resulting from the loss of 26. A more stabilised resonance structure has been proposed and examined by computational structural optimisation.

2.3.5. Computational Studies

Computational analysis with Gaussian 09 has been performed alongside the experimental data to reinforce and understand the fragmentation observed. Firstly, to gain insight into the most likely site(s) of protonation, the proton affinity (PA) and gas phase basicity (GPB) of all basic sites of each of the amino acids have been calculated. Secondly, the fragmentation pathways have been modelled for certain amino acids to examine the energy costs and reaction profile diagrams have been sketched where appropriate. Thirdly, the stabilities and energies required for the neutral losses of HCOOH and NH₃ have been compared to answer why specific fragmentation is seen in some MS/MS spectra. In addition, theoretical data obtained from Gaussian, such as optimised energies and bond lengths (generally bonds which undergo fragmentation), have also been compared to peak intensities of related neutral losses. Lastly, transition states for some fragmentation pathways of a few interesting amino acids have been simulated. These have been combined into a short movie to assist data visualisation. Overall, Gaussian has been proven to be an extremely powerful tool to understand the ions observed and the pathways followed in tandem mass spectrometry studies of amino acids.

2.3.5.1. Proton Affinity

The proton affinity (PA) of a molecule is defined as the energy released when a proton is added in the gas phase to produce a cation (see Equation 23). It is computed as the energy difference between the system of interest and the same molecule with one additional proton, for example, as in Equation 24. Here M represents any neutral amino acid, and H⁺ is a proton. MH⁺ is, therefore, the protonated amino acid.



Equation 23

$$E(M) - E(MH^+)$$

Equation 24

Gaussian 09 has been used to compute the PAs of all the possible basic sites of every amino acid to investigate the most likely sites of protonation. The information acquired can then be applied to help

elucidate the fragmentation pathways of these amino acids. Table 5 Calculated proton affinities for all 20 amino acids in kcal/mol. References have been provided in two columns, experimental (ref 200) and evaluated (ref 201). It shows the value of PA of the most favourable protonation sites of all 20 amino acids. All PAs and additional references have been provided to compare theoretical data. Lysine and methionine have been labelled red as the results are more than 10 kcal/mol, different from both reference values. Results labelled in blue are theoretical data having more than 10 kcal/mol differences from one of the reference values. In general, there is a very good agreement in the results obtained in this research compared to the two published references.^{200,201} This provides great confidence in using these results for further energy studies and transition state simulations. Anomalies marked in red have been further re-calculated using the B3LYP basis set, but there are still disagreements with the references. The reasons for this are not currently clear. However, the values calculated here should still be treated with great confidence as they can be rationalised when compared to others in the same side chain functionality groups. Overall, the most favourable protonation site is on the nitrogen of the amine group.

Name	Proton Affinity (kcal/mol)	Experimental ^a	Evaluated ^b
Alanine	225.5	222.1±2.9	214.2
Arginine	247.7	>243.2	247.8
Asparagine	222.3	227.2±2.2	223.6
AsparticAcid	220.6	220.2±2.1	216.4
Cysteine	214.4	214.6±2.7	214.0
GlutamicAcid	228.6	240.6±1.9	226.4
Glutamine	228.7	228.5±2.2	225.0
Glycine	210.7	211.7±3.2	210.5
Histidine	235.2	228.5±2.2	234.5
Isoleucine	220.8	225.0±3.1	218.6
Leucine	218.9	214.0-216.5	217.4
Lysine	220.9	242.6±3.4	235.6
Methionine	215.3	228.5±2.2	224.1
Phenylalanine	220.4	216.5-218.1	219.9
Proline	225.6	218.8-222.1	222.1
Serine	217.4	220.2±2.1	215.2
Threonine	221.7	228.5±2.2	219.5
Tryptophan	227.6	231.0±3.3	226.9
Tyrosine	221.3	227.2±2.2	220.9
Valine	217.9	222.1±2.9	216.5

Table 5 Calculated proton affinities for all 20 amino acids in kcal/mol. References have been provided in two columns, experimental (ref²⁰⁰) and evaluated (ref²⁰¹).

2.3.5.2. Gas Phase Basicity

As Gaussian calculations have investigated proton affinity, the gas phase basicity (GPB) of the 20 amino acids has also been examined and summarised. Gas phase basicity is related to proton affinity, and it is the negative of the Gibbs energy change, ΔG° , in the gas phase associated with Equation 25. Data are obtained from the thermochemistry section in the Gaussian output files. It is computed as:

$$G(M) - G(MH^+)$$

Equation 25

M is the neutral amino acid, MH^+ hence is the protonated amino acid. Table 6 gives GPB values of the 20 amino acids. Unlike PA, GPB includes entropic terms in its definition, providing more insight into the thermodynamics of amino acid protonation in the gas phase. The same references in the previous section have been used to compare the data.^{200,201} Most data agree with the published experimental and evaluated results. Anomalies have been highlighted in red. Because their PAs are already in disagreement with the references, it is expected that their GPB values would also show inconsistencies. The evaluated GPB of alanine is much smaller than both results obtained in this study and the experimental value from reference a. Thus, it is highlighted in blue. From Table 5 and Table 6, it has been shown again that Gaussian can be used as a powerful analytical tool in assisting the understanding of mass spectrometry studies. Results obtained here provide a solid foundation for further energy related studies of fragmentations in the gas phase.

Name	Gas Phase Basicity (kcal/mol)	Experimental ^a	Evaluated ^b
Alanine	217.7	212.3±2.5	206.4
Arginine	237.7	>233.8	237.0
Asparagine	212.3	217.4±2.6	212.8
AsparticAcid	210.6	210.3±2.6	208.6
Cysteine	204.4	205.4±2.8	206.2
GlutamicAcid	218.6	223.7±1.9	215.6
Glutamine	218.7	219.5±2.6	214.2
Glycine	202.9	202.4±3.2	202.7
Histidine	225.2	219.5±2.6	223.7
Isoleucine	213.0	214.8±3.1	210.8
Leucine	211.1	214.8±3.1	209.6
Lysine	210.9	226.0±3.4	221.8
Methionine	205.3	219.5±2.6	213.3
Phenylalanine	212.6	217.4±2.6	212.1
Proline	217.8	221.9±2.9	214.3
Serine	209.6	210.3±2.6	207.6
Threonine	211.7	219.5±2.6	211.7
Tryptophan	219.8	221.9±2.9	216.1
Tyrosine	213.5	217.4±2.6	213.1
Valine	210.1	212.3±2.5	208.7

Table 6 Gas phase basicity of the 20 amino acids. References are the same as in Table 5. Anomalies have been highlighted in blue and red.

2.3.5.3. Formation Energy of Carbocations and Peak Intensity

Previous MS results have shown that the loss of HCOOH is non-discriminating for the 20 proteinogenic α -amino acids. Loss of NH₃ is also prominent among the neutral losses. The relationship between the relative formation energy of these product ions and their peak intensities has been investigated in this section. Gaussian 09 is used to calculate the relative formation energy of the two product ions for 16 α -amino acids from the loss of HCOOH and the loss of NH₃. Alanine, glycine, proline and valine have been excluded from the study as their masses are too small to be analysed on the Orbitrap. Table 7 lists all 16 α -amino acids studied, a comparison between the peak intensities of loss of HCOOH vs loss of NH₃, and a comparison of their relative formation energies

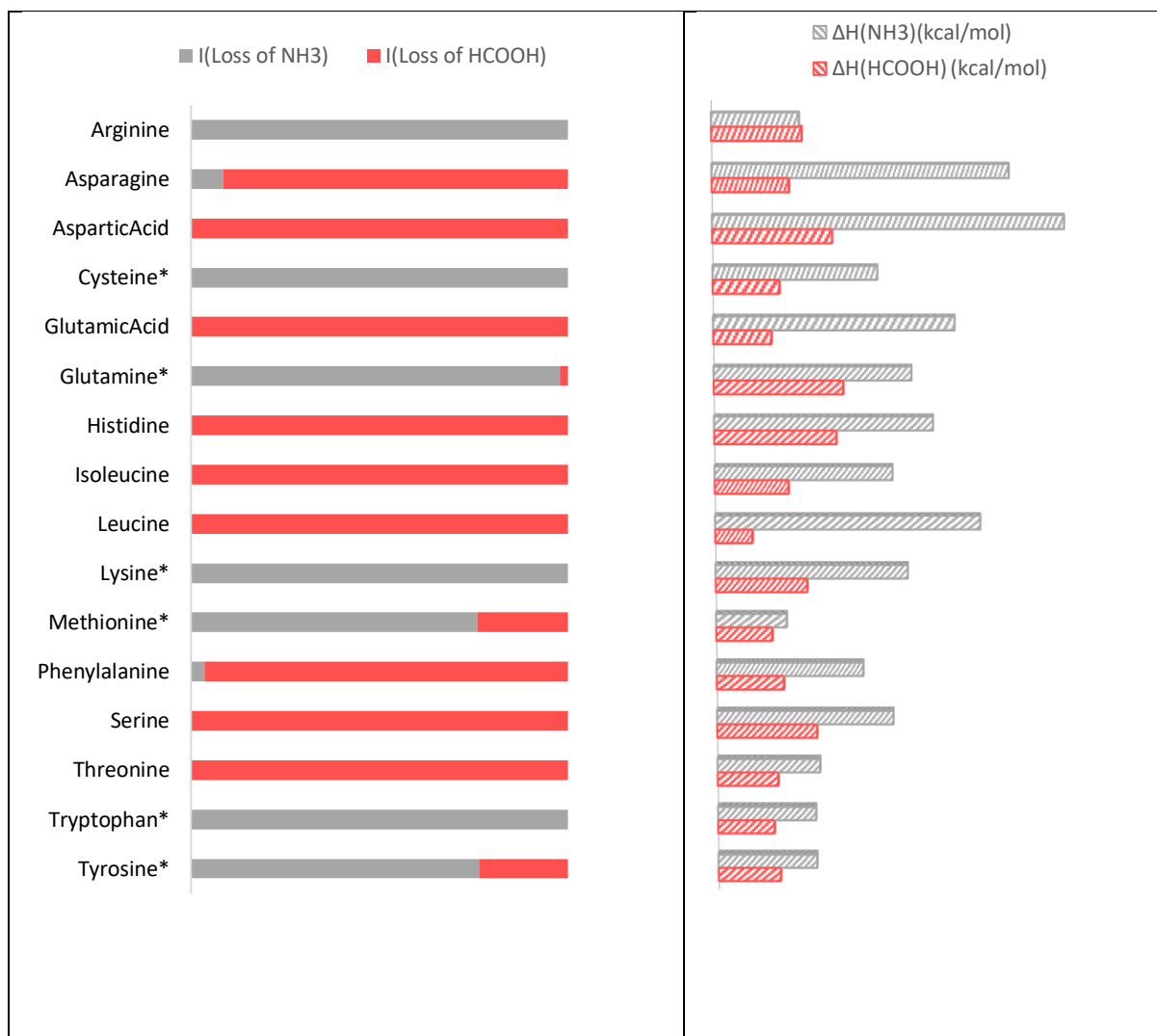


Table 7 Comparison of relative formation energies and peak intensities of loss of NH_3 and loss of HCOOH of the 16 amino acids studied. Anomalies are marked by *. I is the peak intensity. $\Delta H(\text{NH}_3)$ is the relative formation energy of $[\text{M}+\text{H}-\text{NH}_3]^+$. $\Delta H(\text{HCOOH})$ is the relative formation energy of $[\text{M}+\text{H}-\text{HCOOH}]^+$.

It can be observed from Table 7 that the higher the observed peak intensity of that carbocation, the lower the relative formation energy of a carbocation. For example, it costs less energy for arginine to lose NH_3 than HCOOH in its fragmentation pathway, therefore the peak intensities of product ion $[\text{M}+\text{H}-\text{NH}_3]^+$ is expected to be higher in its tandem mass spectrum, loss of NH_3 is more favourable than the loss of HCOOH . This trend remains true with most of the α -amino acids. However, some anomalies are marked by *.

Among these anomalies, cysteine and methionine both prefer the loss of NH_3 , despite the lower formation energies for the loss of HCOOH carbocations. A possible reason for this is the contribution of sulfur on the side chain. In the fragmentation pathway of cysteine, loss of NH_3 leads to further product

ions facilitated by the sulfur. As for methionine, the loss of NH_3 promotes an additional loss of methanethiol.

In the case of glutamine, the initial loss of NH_3 is followed by a secondary loss of HCOOH , which is the more favourable route. It suggests that $[\text{Gln}+\text{H}-\text{NH}_3]^+$ is more stable than $[\text{Gln}+\text{H}-\text{HCOOH}]^+$.

Although it costs more energy for lysine to lose an NH_3 on the α -carbon, its side chain lysyl provides an alternative route for the loss of NH_3 . Lastly, for aromatic amino acids tyrosine and tryptophan, higher peak intensities for loss of NH_3 can be rationalised by additional mechanisms shown previously in sections vii and viii from **Part 2.3.4.1**.

2.3.5.4. C-C Bond Length and Peak Intensity

The computational study collected three bond lengths of 19 amino acids from the Gaussian 09 geometry optimisation calculation. These bond lengths are represented in Figure 32, where a is the bond length of the C-C bond, b is the bond length of the C-X bond, and c is the C-N bond. It is assumed that a correlation may be found by observing the relationship between the normalised bond length and peak intensity of neutral loss HCOOH .

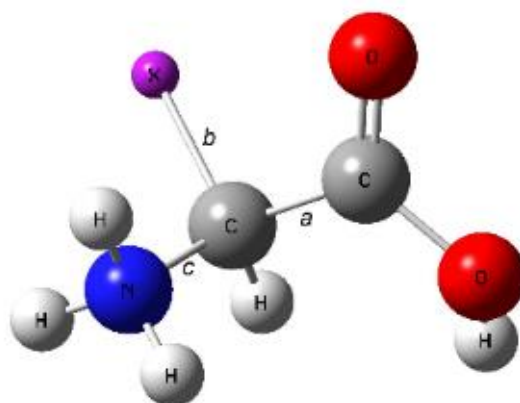


Figure 32 Illustration of bond lengths (a , b and c) collected in any amino acid. Atoms have been labelled. The side group is represented by X . Bond lengths $a(\text{C-C})$, $b(\text{C-X})$ and $c(\text{C-N})$ are recorded and used in the calculations. Grey, carbon; white, hydrogen; blue, nitrogen; red, oxygen; and purple, side group.

In this part of the study, $a_{(C-C)}$ has been normalised by Equation 26, represented by R .

$$R = \frac{a_{(C-C)}}{a_{(C-C)} + b_{(C-X)} + c_{(C-N)}}$$

Equation 26

A scatter chart has been plotted based on R obtained from 19 α -amino acids and their peak intensity of loss of HCOOH. In addition to the 16 α -amino acids used in the previous study, three more have been added. These are H-Lys(Alloc)-OH, H-Tyr(tBu)-OH and L-allo-Isoleucine. The first two are derivatives of lysine and tyrosine, and the latter is a stereoisomer of isoleucine. Their structures are shown in Figure 33.

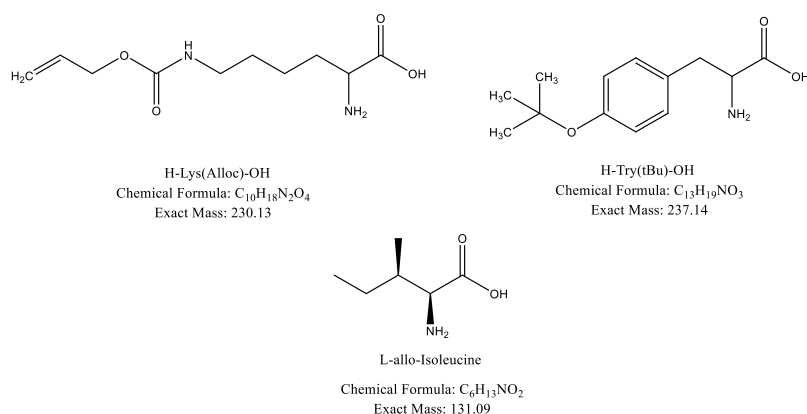


Figure 33 Structure, chemical formula and exact mass of H-Lys(Alloc)-OH, H-Tyr(tBu)-OH and L-allo-Isoleucine.

All calculations were performed by DFT using the B3PW91 6-311++G(d,p) basis set. It can be observed from the chart that there are clusters which are formed mostly by amino acids from the same side chain functionality group (Figure 34). For example, tryptophan, H-Tyr(tBu)-OH and tyrosine are close together, as they are all aromatic amino acids. The same applies to lysine, H-Lys(Alloc)-OH and arginine due to their basic side chains. Aliphatic amino acids isoleucine and L-allo-Isoleucine are found in the same cluster, and cysteine and methionine are also found to position close by. This shows the direct influence of side chain functionality on the fragmentation pathways. In addition, it can be observed that by adding an allyloxy carbonyl group on the amine of lysine, the peak intensity of the loss of HCOOH increased nearly three times. This raised interest in analysing more varieties of amino acids to produce a more detailed investigation of the impact of side chain functionality on their fragmentation pathways.

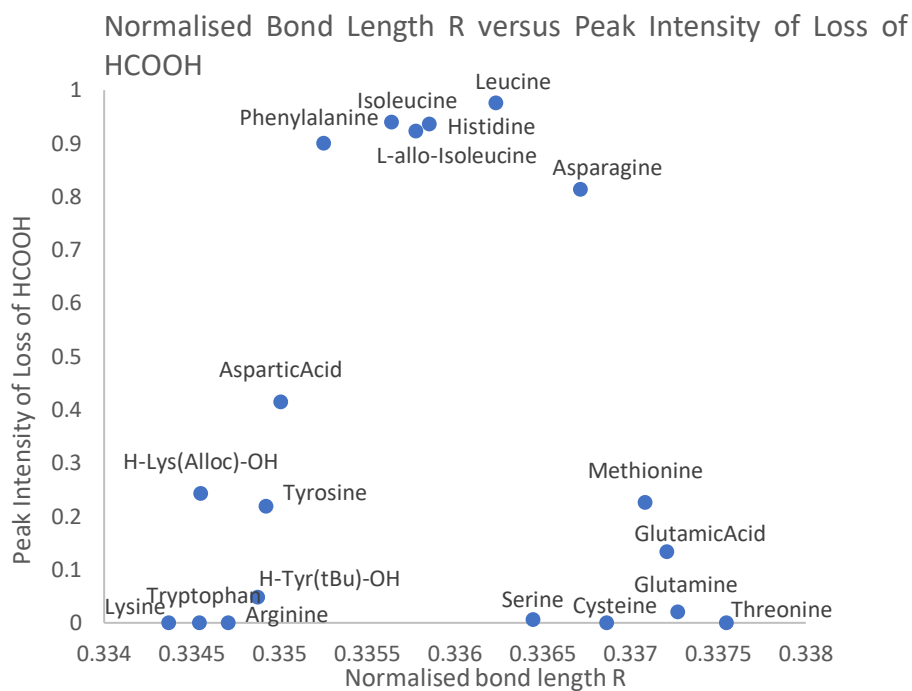


Figure 34 A scatter chart of normalised bond length R versus peak intensity of loss of HCOOH. Calculated by DFT at B3PW91 6-311++G(d,p). All 19 α -amino acids are labelled.

2.3.5.5. Reaction Pathway and Geometry Optimisation

The energy breakdown graphs for the fragmentation of the protonated amino acids give insight into the energy cost to induce various fragmentations and the effect that increasing the collision energy has on the intensity of these fragmentation peaks. However, it does not provide enough detail on exactly how much energy it takes to cause every neutral loss, nor does it identify the energy changes occurring during the fragmentation. To further understand the individual fragmentation pathways, Gaussian calculations are necessary. In this part of the study, geometry optimisation by Gaussian 09 has been used to locate possible transition structures. Reaction path calculations have been requested by the 'IRC' (intrinsic reaction coordinate) keyword to verify that the optimised transition structure connects the desired reactants and products. Once a reaction pathway has been established, the reaction profile for the fragmentation of a protonated amino acid can be constructed. However, it has to be noted that although the intrinsic reaction coordinate provides a convenient description of the reaction progress, real molecules have more than infinitesimal kinetic energy and will not follow the intrinsic reaction path.²⁰² It is important not to attribute them too much chemical or physical significance.

Among the 20 amino acids, the reaction profile of serine has been created based on its fragmentation pathway. The energy cost for every step has been calculated. Additionally, the transition structures have been computed and confirmed by IRC. A short video has been created to visualise the fragmentation pathway of serine for the loss of the second H₂O and CO. From the discussion in the previous section, and serine has competing fragmentation pathways. One comprises two losses of H₂O and a further loss of CO; the other is the loss of HCOOH. Its energy breakdown graph suggests that the first route is much more dominant than the latter. Close examination of both reaction profiles reveals that compared to a loss of HCOOH, loss of an H₂O group costs much less energy.

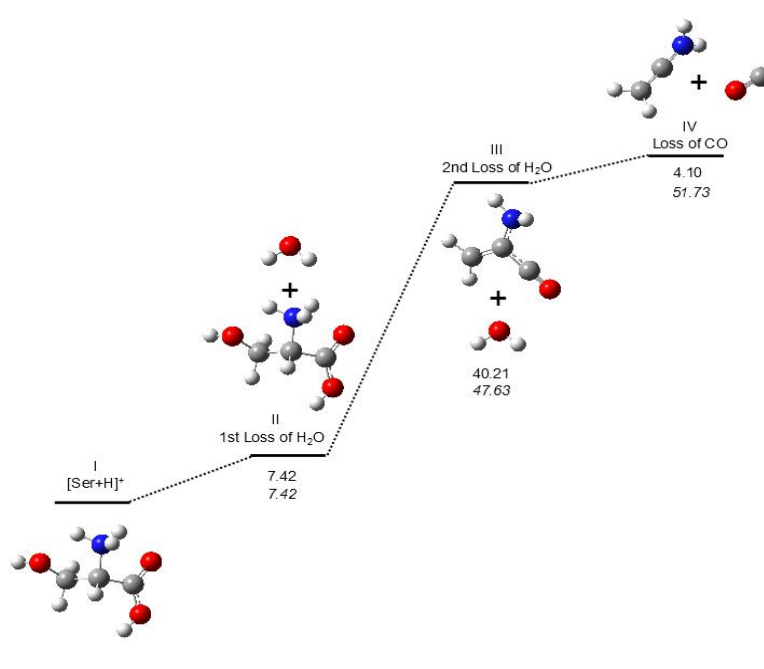


Figure 35 Reaction profile of protonated serine for loss of the first and second H₂O (structure II and structure III) and CO (structure IV), calculated at B3PW91 6-311++G(d,p). Upper number ΔH to previous structure; lower number in italic, ΔH(total), to protonate serine in KJ/mol. Colours of atoms: grey, carbon; white, hydrogen; red, oxygen and blue, nitrogen.

In Figure 35, protonated serine (structure I) loses the first H₂O (structure II). This initial step costs only 7.42 kcal/mol. It then loses a second H₂O, costing 40.21 kcal/mol for this stage. Finally, a loss of CO follows. The overall fragmentation pathway requires 51.73 kcal/mol.

In comparison, Figure 36 shows that it costs 21.43 kcal/mol for protonated serine to lose HCOOH. This is almost three times more than fragmentating an H₂O. Therefore, the loss of HCOOH is not

energetically favourable. Hence loss of H₂O is more intense in the energy breakdown graph of protonated serine.

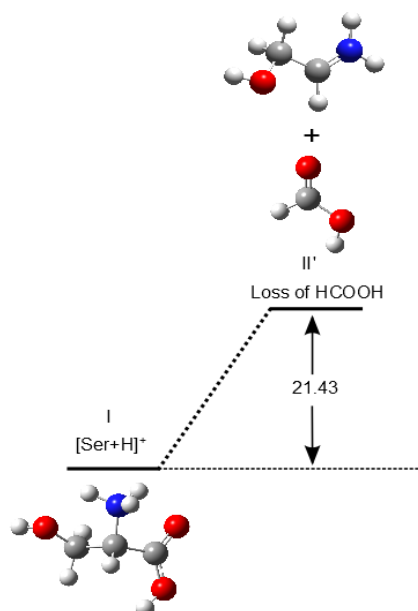


Figure 36 Reaction profile of protonated serine for loss of HCOOH, calculated at B3PW91 6-311++G(d,p). Number, ΔH , in kcal/mol. Colours of atoms: grey, carbon; white, hydrogen; red, oxygen and blue, nitrogen.

Furthermore, the lack of loss of NH₃ in the MS/MS spectra of protonated serine can also be explained by its energy costs. Compared to a very favourable loss of H₂O, which only costs 7.42 kcal/mol, or a less favourable (but still viable) loss of HCOOH which requires 21.43 kcal/mol, a loss of NH₃ needs 37.83 kcal/mol to occur. Loss of NH₃ costs the most energy and hence becomes the least energetically favourable route. It is, therefore, not seen in the neutral losses of protonated serine.

As well as using Gaussian calculations to visualise fragmentation pathways, it can also be used in determining the optimised geometry of product ions. For example, in the negative ion mode mass spectra of deprotonated asparagine, product ion m/z 95 is proposed to have a negative charge on oxygen (Scheme 18). However, Gaussian geometry optimisation revealed that the negative charge is more likely to be shared between C1 and C2. The bond length of C1 and C2 is 141 pm, which is in between the length of a C-C and C=C bond. Figure 37 is the result of Gaussian 09 geometry optimisation.

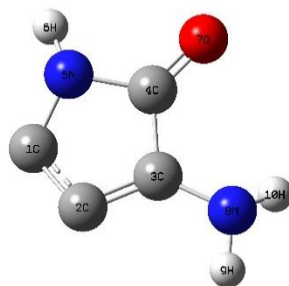
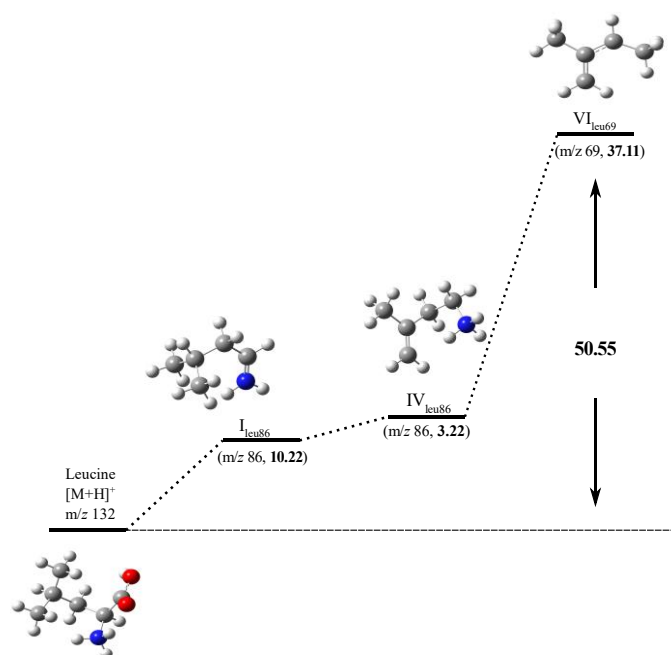


Figure 37 Optimised structure of product ion m/z 95 of deprotonated asparagine in negative ion mode. Atoms are labelled. Grey, carbon; white, hydrogen; blue, nitrogen; and red, oxygen.

2.3.5.6. Computational Data on Leucine and Isoleucine

Gaussian modelling has been used to rationalise the lack of m/z 69 loss for leucine. The energy costs for each fragmentation pathway to form product ion m/z 69 were modelled and shown in Figure 38. Fragmentation pathways from their proposed mechanisms were constructed and labelled. In comparison, Leu costs more to initiate the loss of HCOOH in the first step. Although the energy difference between structure I and IV is only 3.22 kcal/mol, the overall energy costs to form structure VI are higher than that of Ile. Ile has more stabilised structures which could be structurally optimised in its suggested fragmentation pathway. There is a noticeable increase in energy from structure I to III of Ile. However, this is offset by a 17.27 kcal/mol decrease from structure III to IV. The total energy costs to form product ion m/z 69 for Ile are lower when compared to Leu, suggesting a more energetically favourable fragmentation route. This part of the data supported the observation of a lack of product ion m/z 69 for leucine and formed a part of the paper published in 2020.¹⁶⁵

(a) Leucine



(b) Isoleucine

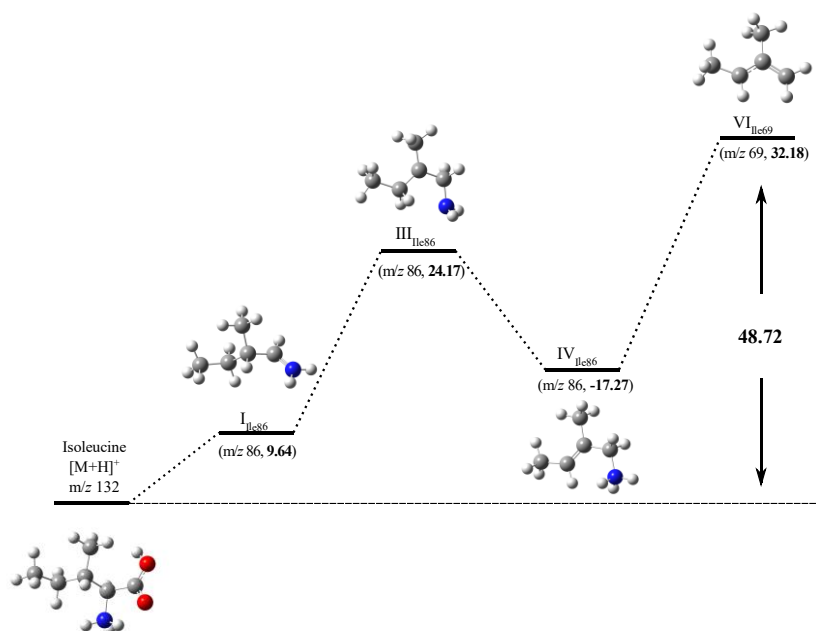


Figure 38 Potential energy diagrams for the fragmentation of (a) Leu and (b) Ile. Numbers in brackets are mass to charge ratio and relative formation energy (kcal/mol) relative to the previous structure. The total energy (kcal/mol) costs to form structure VI is marked in bold. Carbon: grey; Nitrogen: blue; Hydrogen: white.

2.4. Conclusion

In this part of the project, a systematic analysis has been performed to study the fragmentation pathways of α -amino acids using tandem mass spectrometry. It has been discovered that side chain functionalities of these 20 α -amino acids play a significant role in influencing their fragmentation pathways. In the positive ion mode, loss of HCOOH is predominant and non-discriminating. This is followed by the loss of NH₃ and H₂O. Amino acids in the same side chain functionality group generally have the same or similar neutral losses. However, variations have been found within the same group, usually explained by the additional carbon atom or slight structural change in their side chains. In addition, structural isomers leucine and isoleucine show differences in their MS. It has been investigated in detail and explained by several approaches. HCD has also been employed as an additional fragmentation mode for the aromatic amino acids phenylalanine, tyrosine and tryptophan, generating more product ions. These additional ions allow for a more detailed analysis and comparison of their mechanisms.

MS/MS spectra of these 20 α -amino acids in the negative ion mode have also been recorded. Unlike in the positive ion mode, loss of NH₃ is the most abundant in the negative ion mode. The second most frequent loss is the loss of H₂O, followed by losses of CO and CO₂. Side chain functionalities still have influences on their respective fragmentation pathways. However, there is less uniformity within the same amino acid side chain functionality group.

Energy breakdown graphs have been plotted for the fragmentations of protonated amino acids. Energy breakdown graphs correlate the peak intensities of the different product ions to the collision energy used in the tandem mass spectrometry experiment. By analysing these graphs, it is apparent that the loss of HCOOH generally happens at very low collision energy. In addition, the effect of increasing the collision energy on different neutral losses can be observed from these graphs. This part of the research also noticed that the energy breakdown graphs obtained from HCD have a very distinguishable shape compared to those obtained from CID.

To understand the fragmentation pathways of these α -amino acids, mechanisms have been proposed for most of the side chain functionality groups in both positive and negative ion modes. It has also been found that amino acids within the functionality group usually share the same mechanism. Although differences are present in cases where more neutral losses are observed and/or several competing routes are possible.

Computational analysis has been utilised in this study to investigate the fragmentation pathways of the 20 α -amino acids. Calculations performed by Gaussian 09 have provided insight into the most likely sites of protonation, proton affinity and gas phase basicity. Moreover, the relative formation energies of the carbocations and bond lengths have been calculated to compare with the peak intensities. This explains why certain α -amino acids prefer the loss of HCOOH or NH₃. Lastly, IRC calculations in Gaussian 09 have made locating transition states and following reaction paths achievable. The reaction profile of serine has been created, as well as a simulation of its fragmentation pathway. This is a great way of visualising fragmentations observed in mass spectra of certain amino acids and can be further applied to a larger variety of natural products. It has been shown that Gaussian is a very powerful analytical tool and its ability to complement studies performed by tandem mass spectrometry.

In addition to the systematic analysis of the 20 α -amino acids, a more detailed study on the fragmentations of leucine, isoleucine and L-allo-Isoleucine has also been included in this report. Several approaches have been used in rationalising the different fragmentation pathways observed in their mass spectra. It could be a future interest of this research to analyse more amino acid structural isomers by IMS-MS/MS.

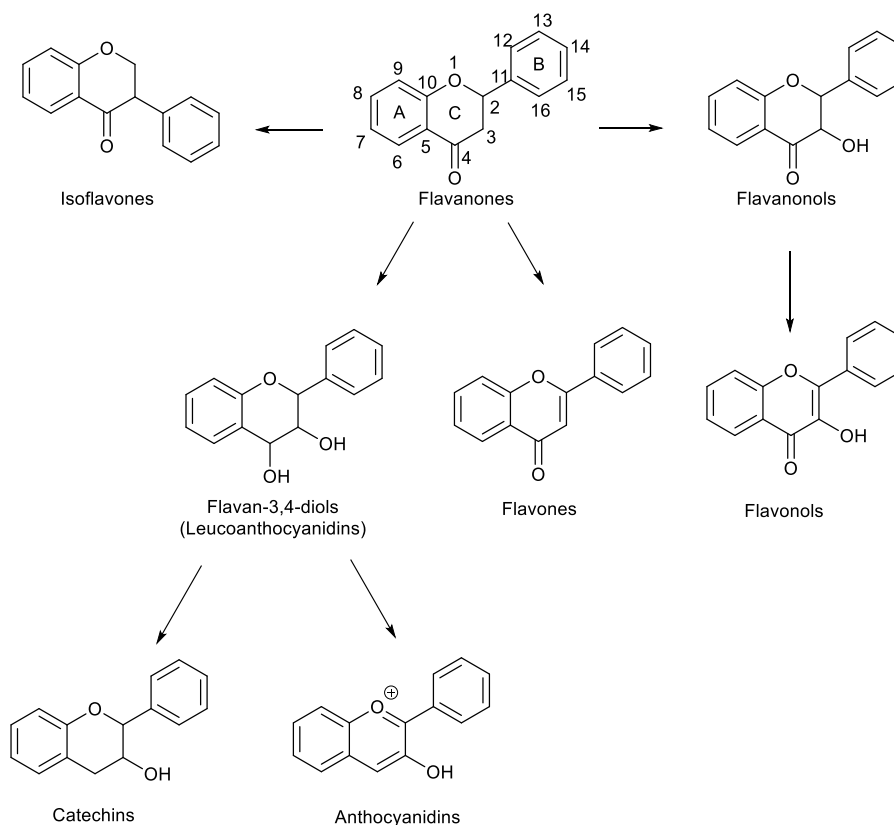
To conclude, MS combined with computational modelling has proven to be a significantly valuable tool in studying α -amino acids. It will be of interest to expand this research method to perform studies of other small molecules and natural products. The knowledge and skills obtained from this research can provide a necessary platform for the rest of this research. The results have proven that the methodologies used are applicable to this class of complex/related small natural products. The next chapter will study a class of compounds with considerable isomeric structures and more competing fragmentation pathways – the flavonoids.

3. Flavonoids

3.1. Flavonoids Overview

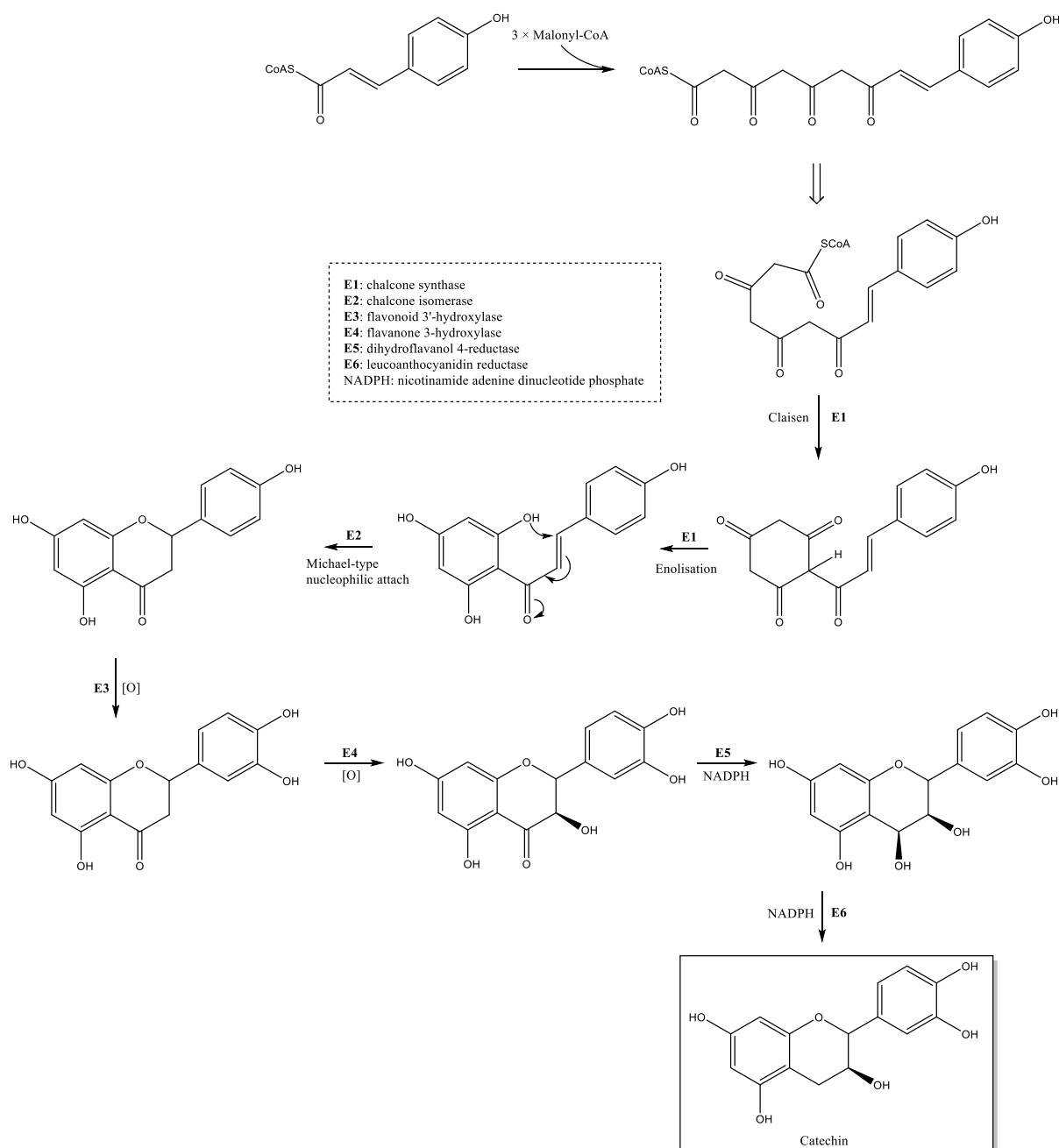
Flavonoids are a class of polyphenolic secondary metabolites mainly found in plants. They are present in fruits, nuts and vegetables and have become a part of our daily dietary intake. Flavonoids often contribute to the colour of most flowers. Vegetables use them for their growth and defence against disease, infection, provide UV protection and assist in pollination.²⁰³ Due to their unique structures, flavonoids have been found to have many important biological activities such as anti-oxidative, anti-inflammatory, anti-mutagenic and anti-carcinogenic. Their extensive applications include nutraceutical, pharmaceutical, medical, and cosmetic uses.

Flavonoids are widely distributed in the edible parts of food plants as flavonoid glycosides. However, some are found in their free form (aglycone), which can exhibit phenolic properties.²⁰⁴ Due to the increasing awareness of the health benefits and pharmaceutical applications of flavonoids, studies of flavonoids using mass spectrometry have been rising dramatically. In the later section of this chapter, a few highly cited studies are summarised, and their results are discussed. A skeleton structure of flavonoids is shown in Scheme 22, along with their variation classes - the degree of oxidation on ring C determines which subclass the flavonoid belongs to. During this project, flavonoids from all listed classes were investigated. A full list of these flavonoids is in Appendix: 2. List of Analytes.



Scheme 22 Different classes of flavonoids. Figure recreated from the literature.²⁰⁴

Scheme 23 below shows an illustration of the biosynthesis of flavonoids.²⁰⁵ The biosynthesis follows the phenylpropanoid metabolic pathway, which begins with phenylalanine in the production of 4-hydroxycinnamoyl-CoA. This can then cooperate with malonyl-CoA to give chalcones, the backbones of flavonoids. It is followed by the chalcone's ring closure to give the flavone's typical three-ring structure. If a series of enzymatic modifications are carried on, many products, including flavanones and flavonols, can then be synthesised.



Scheme 23 Reproduction of the mechanism for the biosynthesis of catechin

A few review articles in the literature cover an overview of the chemistry and biological activities of flavonoids.^{203, 206} It has been stated that the configuration, number of hydroxyl groups and substitution of functional groups on the structures of individual flavonoids affect their bioavailability, metabolism and biological activities. The first flavonoid to be identified was rutin. It was initially extracted from oranges and was referred to as vitamin P. However, this was soon correctly replaced by its flavonoid class name.²⁰⁷ The specific subclasses of flavonoids differ by the oxidation level and substitution pattern on the C-ring. Within the same class, flavonoids vary by substitution on the A- and B-rings (see Scheme 22). Among all classes of flavonoids, flavonols are the most abundant in foods.²⁰⁶ Preparation and

processing of food will decrease the flavonoid content; however, the extent of reduction depends on the methods used.²⁰⁸

The isoflavonoid glabridin from *Glycyrrhiza glabra* was found to inhibit low-density lipoprotein oxidation by scavenging free radicals.²⁰⁹ In some studies, the human metabolism of flavonoids and the mechanism for their antioxidant activity have been studied.²¹⁰ The hepatoprotective activity of several flavonoids has also been investigated.²¹¹ In addition, antibacterial, anticancer, anti-inflammatory, and antiviral functions are also well known for flavonoids. Finally, a paper has also stated methods for microbial production of flavonoids using metabolic engineering and synthetic biology.²¹²

A second review paper focused on the current research trends and developments on flavonoids. This includes the human metabolic pathways, functions and applications, as well as potential usage as pharmaceutical drugs.²⁰³ Flavonoids are found in all parts of plants,²¹³ Figure 39²⁰³ gives a clear representation of flavonoid classes, subclasses and their natural sources. Current research and trends have been summarised in detail in the literatures.^{214,215} This includes studies on the applications of flavonoids as anti-cholinesterase, anti-inflammatory, steroid-genesis modulators, xanthine oxidase modulators, countering antibiotic resistance, and disease-combating, amongst others. In addition, the mechanisms of flavonoids have been described based on their capacity to behave as antioxidants. Finally, many of the functions and applications of flavonoids are listed in Figure 40, which is simplified from a more detailed version of the review paper.²⁰³

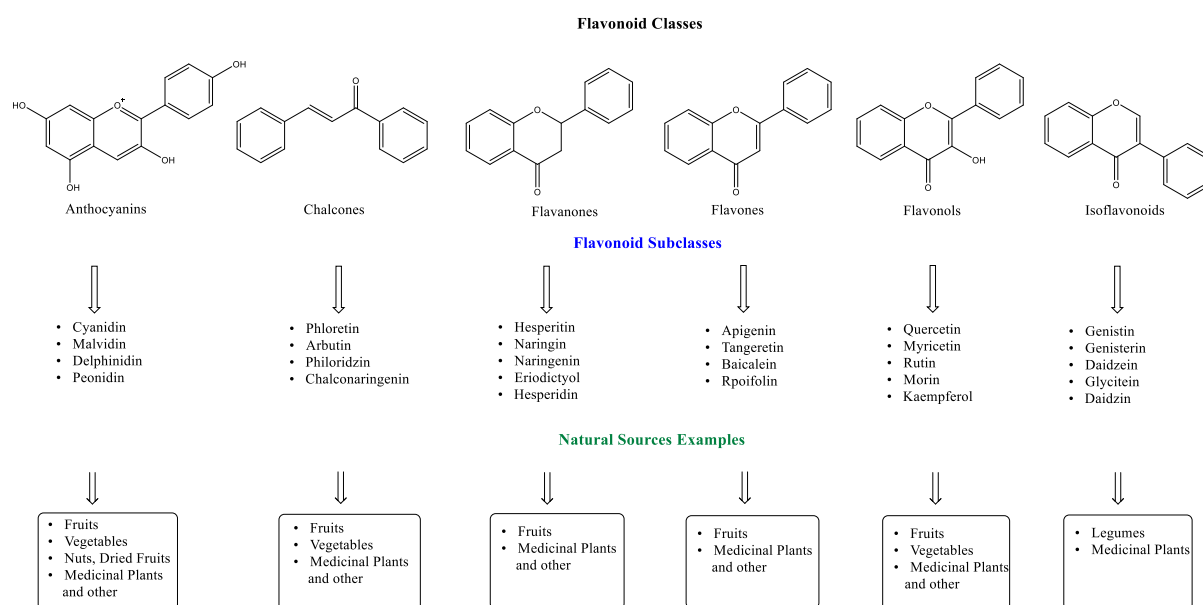


Figure 39 An illustration of flavonoid classes, subclasses, and natural sources examples.

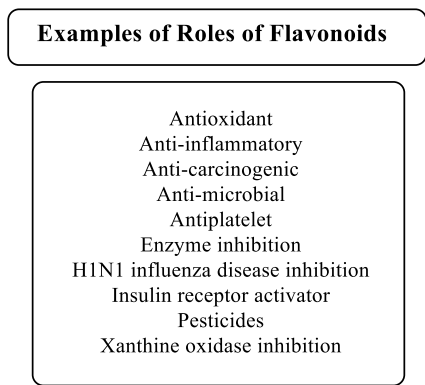


Figure 40 Examples of the roles of flavonoids in various bioactivities, human health, and agriculture.

Despite the antioxidant activities reported widely in the literature of flavonoids, one study has reported that morin and naringenin could induce a concentration-dependent peroxidation of nuclear membrane lipids concurrent with DNA strand breaks in rat liver nuclei.²¹⁶ The reported reactions were believed to be promoted by metal ions such as iron or copper. The published paper demonstrated the possible dual role of morin and naringenin in mutagenesis and carcinogenesis. This further proves that there are more scopes for research into the properties of flavonoids.

3.1.1. Studies of Flavonoids by Mass Spectrometry

Flavonoids remain a popular topic in scientific research, especially using techniques such as mass spectrometry (MS). The available literature on flavonoid analysis by various MS techniques is overwhelming. Here a few are selected chronologically to demonstrate some notable results.

Methane CI-MS was used in the study of flavonoids in the early 1970s.²¹⁴ A range of flavonoids was studied, including flavones, flavonols, flavanones and flavanols. The flavones and flavonols were shown to produce no significant fragmentations during CI-MS. However, the flavanones and flavanols gave characteristic fragmentations believed to be useful in structural elucidation analysis. Later in the 1970s, the CI-MS study of flavonoids was reinvestigated.²¹⁷ Eleven flavonoids were analysed using methane as the reagent gas and six using isobutane. When isobutane was employed, the production of $[M+CH_3]^+$ adduct ions by methylation was favourable for the flavonoids with hydroxy groups. Loss of propene, C_3H_6 , was observed from the $[M+C_2H_5]^+$ when methane was used with flavonoids having methoxy groups.

A study published in 2000 reviewed the existing applications of MS for the identification and structural studies of flavonoid glycosides.²¹⁸ It included several techniques such as EI and CI and the desorption

ionisation methods (FAB and LSIMS). In EI, it was reported that molecular ions were observed, but product ions dominated the spectra. Ions are found to have charge retention on the aglycone. However, there was no fragmentation due to retro Diels-Adler (RDA) reactions. This review paper stated that CI did not give satisfactory results in either positive or negative ion modes. As for the desorption ionisation techniques, ions observed are typically $[M+H]^+$, $[M+Na]^+$, $[M+K]^+$ and $[M-H]^-$. Cleavage of the sugar moiety has been observed extensively, and the relative intensities of the sample ions in the spectra depend on the particular ionisation technique and structure of the analytes.

The later study implemented ESI coupled with HPLC-MS/MS to investigate 14 flavonoids and their monohydroxy- to pentahydroxy-substitutes.²¹⁹ Spectra of deprotonated flavonoids were reported in the negative ion mode, in addition to the protonated ions $[M+H]^+$ of naringenin and galangin. Five general fragmentation processes were proposed (Figure 41). The identification of genistein from wood pulp was also demonstrated in this study.

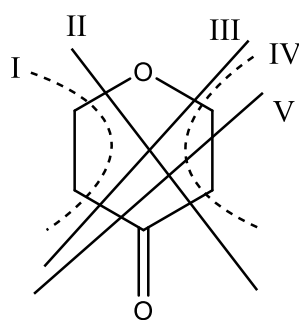


Figure 41 A schematic showing the proposed representation of the five general cross ring fragmentation processes observed in the C ring during the CID-MS/MS of flavonoids.²¹⁹

In 2007, HPLC-MS/MS was used to study 12 flavonoids in positive ion mode.²²⁰ Specific fragmentations were found following dehydration and loss of CO from protonated molecular ions. In addition, cleavage on the C rings could be used to identify the different subclasses and the substitution patterns of the A and B- rings can also be determined. Unknown flavonoids were also studied using the characteristic fragmentation patterns discovered in this study. HPLC-MS was also used to investigate isoflavones and their conjugates in soy products with positive ion APCI.²²¹ In this study, it was stated that the spectra from APCI provided the most sensitivity and structurally useful information. The paper also demonstrated that extraction at higher temperatures caused changes in isoflavone composition and should be avoided. The study also showed that soy products such as soymilk, tofu and soy molasses contained the isoflavone β -glucosides as the major flavonoids.

Another study used LC-ESI-MS/MS to characterise flavonol glycosides in tomato extracts and human plasma.²²² The study examined representative aglycones and glycosides of flavonols, flavones and

isoflavones to investigate any relationship between their structures and their mass spectra. An ESI-HPLC-MS-DAD method was developed to achieve direct observation of flavonols in tomato puree extract. The study also reported on the successful detection of intact flavonol glycosides in the plasma of healthy volunteers. This provided evidence of the physiological absorption of flavonoid glycosides after consuming vegetables.

Besides studies performed in the positive ion mode, fragmentations of flavonoids have also been explored in the negative ion mode. The earliest study used EI-MS in negative ion mode to analyse flavones, flavonols and their permethyl ethers.²²³ It described a new fragmentation pathway of 6-methoxyflavones and the formation of $[M-H]^-$ and $[M-H_2O-H]^-$ ions. This study demonstrated the suitability of EI to obtain the structural information of flavones and flavonols. It showed that the presence of the loss of methoxy groups on the 3, 5, 6, and 8 carbon and substitution on ring C could be deduced from their spectra. Three decades later, 11 naturally occurring flavonoid aglycones were studied using ESI MS in the negative ion mode.²²⁴ Several fragmentation pathways were proposed (Figure 42) with supporting evidence from five other synthetic flavonoid aglycones. Some highly specific negative product ions were discovered based on the Retro Diels-Alder (RDA) process. Unusual CO, CO₂ and C₃O₂ losses appeared to be observed only in the negative ion mode. Moreover, one unique lactone-type structure was found in the spectra instead of a ketene, and it was proposed to be a diagnostic RDA ion in the negative ion mode.

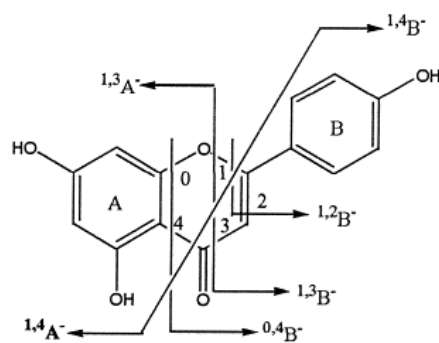


Figure 42 A schematic showing the nomenclature adopted for the various retrocyclisation cleavages observed in the negative ion mode.²²⁴

A detailed review paper was published in 2004 covering the application of MS/MS for the structural analysis of flavonoids.²²⁵ It stated that APCI and ESI were the favoured methods for ionising flavonoids. Negative ion mode gives the highest sensitivity but limited fragmentation, which allows for assessing the molecular masses of various flavonoids when the concentration is low. Apart from protonated or deprotonated molecular ions, adducts with solvent, acid molecules, and molecular complexes have been found in the mass spectra. Positive ion mode gives more structural information, which can provide

sufficient information for identifying known compounds when combined with negative ion mode. Fragmentation pathways of flavonoid aglycones, O-glycosides, C-glycosides and acylated glycosides have all been considered in this review paper.

An additional study has been performed in the negative ion mode by chip-based nanospray MS/MS.²¹⁵ Distinctive low-mass ions were discovered from the deprotonated precursor ion. It was possible to identify the flavonoid class by examining their product ions' patterns. The methodology employed in this study was applied to a green tea extract, and expected flavonoids were discovered. Quercetin was found to be present at about 2% of the most intense ion in the spectrum.

Besides investigations on pure flavonoids, research has been carried out on the effect of transition metal chelation of iron and copper with flavonoids by ESI MS/MS.²²⁶ In this study, the preferred stoichiometry of metal:flavonoid as 1:2 is reported. The preferred metal-binding sites are the 5-hydroxyl and 4-oxo groups for flavones and naringenin. Redox reactions were observed during the experiment through the change of the oxidation state of the metal, alongside the oxidation of the flavonoids by loss of hydrogen.

Further studies have been carried out on the chelation of metals with flavonoids to enhance detection by ESI-MS.²²⁷ The complex ion $[M^{II}(\text{flavonoid-H})\text{byp}]^+$ was formed using a neutral auxiliary ligand, 2,2'-bipyridine. It was reported that the intensity was more than double that of protonated flavonoids and was 1.5 orders of magnitude more than that of deprotonated flavonoids. During the study, six divalent transition metals were used. Cu^{2+} was found to give the most straightforward spectra, followed by Co^{2+} . Eleven flavonoids were investigated, including common ones such as hesperidin, rutin, naringin, genistin and catechin.

Finally, a study of flavonoids in 2003 reported on radical cleavages observed by CID MS/MS.²²⁸ This agrees with our results obtained from this research. In the paper, it was reported that depending on the structures of the flavonoid glycosides, both homolytic and heterolytic cleavage of the O-glycosidic bond could occur. Both deprotonated radical aglycone $(Y_0\text{-H})^-$ and aglycone $(Y_0)^-$ product ions have been observed. The relative abundances of the radical aglycone compared to the aglycone fragments from flavonol-3-O-glycosides increase with the number of hydroxyl substituents in the B ring. The order established for the 3-O-glycoside in the paper follows kaempferol < quercetin < myricetin. It was also proposed that the position and nature of the flavonol glycosides' sugar substitution affect the fragmentation of the radical aglycones.

3.1.2. Aims and Objectives

Over 5000 naturally occurring flavonoids have been discovered in plants. Based on their chemical structures, they are divided into different classes. Six flavonoid classes are selected and studied in this part of the research. All of the flavonoids studied are commercially available. They have been chosen due to their structural similarity and range of isomeric properties.

MS/MS has been used to study flavonoids for over two decades.²⁰³ This part of the research aims to analyse the spectra of 30 flavonoids in the positive and negative ion mode and identify any diagnostic product ions using the same methodology developed in the study of the α -amino acids. These diagnostics product ions can provide important structural information on the protonated molecular ions. Hence, they could be used as fingerprints for identifying unknown flavonoids.

Energy breakdown graphs are then sketched to demonstrate the relationship between collision energy and peak intensity of selected product ions. Breakdown graphs demonstrate the differences between structural isomers by displaying distinctive features for each flavonoid when compared. Combined with proposed fragmentation pathways, the data shows that breakdown graphs can effectively differentiate flavonoid structural isomers. Proton affinities of these flavonoids are calculated to confirm the most energetically favourable protonation site, and this helps to add validity to the proposed mechanisms.

3.2. Experimental

3.2.1. Flavonoids Samples

30 flavonoids from six classes were analysed. All were obtained from Sigma Aldrich (purity \geq 90%). The flavonoids were dissolved in 1:1 methanol:water (1 mg/mL) to make up a standard solution.

List of flavonoids used in this experiment: Quercetin, Morin, Rhamnetin, Isorhamnetin, Kaempferide, Diosmetin, Kaempferol, Fisetin, 5-Hydroxyflavone, Galangin, Baicalein, Apigenin, Luteolin, Scutellarein, Chrysin, Daidzein, 6-Hydroxyflavanone, 7-Hydroxyflavonone, Gallocatechin, Epigallocatechin, Catechin, Epicatechin, Aromadectin, Eriodictyol, Genistein, Sakuranetin, Naringenin, Genkwanin, Hispidulin, Wogonin.

3.2.2. Instrumentation

Positive and negative ESI-MS analyses were performed on an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific) using the heated ESI source. The analyte solutions were diluted to 0.1 mg/mL in methanol:water solvent prior to analysis and were delivered using the autosampler module of an RS3000 UHPLC system (Dionex) at a 10 μ L/min flow rate. Tandem mass spectra were recorded in CID-MS/MS mode on isolated precursor ions (50-500 m/z window) with collision energies ramped from 0 to 50 eV. Energies were increased by 1 eV every 12 seconds. The resulting runs were 10 minutes per sample. A full scan (50 – 500 m/z) was obtained then the selected precursor ion was automatically selected in data dependent mode, isolated and underwent CID-MS/MS using dry N₂ as the collision gas. For all experiments, the acquisition time was 0.25 mins per scan, with a flow rate of 10 μ L/min, mass analyser FTMS mode at a resolution of 240,000, 2 microscans were summed with maximum ion accumulation time of 200 ms.

3.2.3. Computational

Gaussian 09¹²⁸ software was used to perform all optimisation and frequency calculations. All structures were optimised to give a stationary point (minima) on the potential energy surface. Frequency calculations were performed to obtain zero-point energy (a correction to the electronic energy of the molecule, taking in the effects of molecular vibrations, which persist even at 0 K). Density functional theory (DFT) was used, with the hybrid exchange correlation functional and basis set at B3PW91 6-311++G(d,p). This basis set was selected based on a good result from the previous amino-acid research.

3.3. Results and Discussion

3.3.1. Overview

This chapter discusses the MS/MS spectra of selected flavonoids with reference to common neutral losses. Based on these neutral losses, the diagnostic product ions are then summarised. Example adaptation of energy breakdown graphs are demonstrated, and selected structural isomers from various flavonoid classes are investigated. Furthermore, computational modelling data is presented to show the proton affinity of selected flavonoids.

Based on previous literature²⁰⁴, a numbering system is also used in this study. Figure 43 shows the basic structure of a flavonoid. Ring A is marked in red and ring B in blue. The atom counting starts from the heterocyclic oxygen and ends on the ring B benzene carbon. This numbering system is used throughout the analyses of flavonoids.

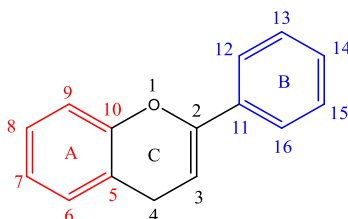


Figure 43 The basic structure and labelling system for the flavonoids.

3.3.2. Common Losses

Positive ion mode ESI-MS/MS spectra were collected for all 30 flavonoids, and the neutral/radical losses from the precursor ions were tabulated (see Table 8). All flavonoid names are coloured to differentiate their classes. **Purple: Flavonol**; **Blue: Flavanone**; **Green: Isoflavone**; **Brown: Flavone**; **Red: Flavanol** and **Grey: Flavanolol**. Flavonoids are listed in descending order of their protonated molecular ion as the precursor ion $[M+H]^+$. The mass range is from m/z 317 to m/z 239. Common neutral losses are listed for comparison and mapped with colour intensity.

The first trend to notice is that flavonoids from the same group do not always have the same common losses, although there are degrees of similarity. Secondly, common losses are mostly dependent on the

location and number of -OH groups (loss of H₂O), -MeO groups (loss of CH₃ radical), and the presence of a C=O function in ring C (loss of CO, mass 28). Any losses higher than 100 mass units result from ring cleavage through retro Diels-Adler (RDA) processes on ring C. Further examination of Table 9 summarises the typical neutral losses observed. Due to the number of flavonoids studied, this chapter mainly focuses on presenting the results and discussion for selected flavonoids with the best established spectra.

Name	Precursor m/z	Common Losses												
Rhamnetin	317	15	18	28	30	46	56	74	138	150	166	178	180	
Isorhamnetin	317	15	32	43	60	164	178							
Gallocatechin	307	18	156	168										
Epigallocatechin	307	18	156	168										
Quercetin	303	18	28	45	46	56	74	102	108	138	132	154	166	182
Morin	303	18	28	42	46	56	70	74	138	150	154	148	176	
Kaempferide	301	15	28	32	42.9	56	60	136	140	148	162	151		
Diosmetin	301	15												
Hispidulin	301	15												
Catechin	291	18	126	140	144	152	168							
Epicatechin	291	18	126	140	144	152	168							
Aromadedin	289	18	46	94	136									
Eriodictyol	289	18	110	125	126	136								
Kaempferol	287	18	28	46	56	74	110	122	134					
Fisetin	287	18	28	46	56	74	102	138	150	166				
Luteolin	287	18	42	46	108	126	134	152						
Scutellarein	287	18	28	46	110	118	168							
Sakuranetin	287	119	120	140										
Genkwanin	285	15	30	43	118									
Wogonin	285	15												
Naringenin	273	18	28	36	42	46	60	66	72	84	102	110	120	126
		150	154	166	182									
Genistein	271	18	28	56	112	118	122	126						
Galangin	271	18	28	42	46	56	74	90	106	118	149			
Baicalein	271	18	18	18	28	46	102	148						
Apigenin	271	42	46	118	126	152								
Chrysin	255	42	46	68	102	108	126							
Daidzein	255	18	28	56	92	110	118							
6-Hydroxyflavanone	241	18	28	42	46	78	83	104	110					
7-Hydroxyflavanone	241	15	18	28	42	46	70	78	86	104	110			
5-Hydroxyflavone	239	44	62	88	102	106	150							

Table 8 List of the flavonoids studied, their protonated molecular precursor ion m/z, and common neutral losses observed. The colour heat map indicates increases in masses. The flavonoids are coloured by class: Purple: Flavonol; Blue: Flavanone; Green: Isoflavone; Brown: Flavone; Red: Flavanol, and Grey: Flavanolol.

Common Loss	Mass
CH ₃ ·	15
H ₂ O	18
CO	28
CH ₄ O	32
2H ₂ O	36
C ₂ H ₂ O	42
C ₂ O	44
H ₂ O+CO	46
2CO	56
CH ₄ O+CO	60
H ₂ O+2CO	74
3CO	84

Table 9 Identification of some of the most common neutral losses observed for positive ion ESI-MS/MS analysis of the 30 flavonoids studied in this project.

3.3.3. Diagnostic Product Ions

A nomenclature is added for clarity for the assignment of the product ions. For example, a ring cleavage labelled 'rCx,yAorB' represents the cleavage location on ring C between carbon-carbon positions x and y, leaving either ring A or B intact.

During this investigation, three product ions were found to be diagnostic for the structure of their precursor ions. These are m/z 167, m/z 153, and m/z 151. Among these, m/z 153 is the most commonly observed and results from cleavage on ring C between atoms 1 and 4, leaving ring A intact and is therefore designated as $\mathbf{rC}^{1,4}_A$. Product ion m/z 167 is also due to cleavage on ring C between atoms 1 and 4, leaving ring A intact and is therefore designated as $\mathbf{rC}^{1,4}_A$. The difference between these two ions is due to the presence of a methoxy group on ring A. Lastly, ion m/z 151 is formed from cleavage on ring C between carbons 3 and 10, therefore designated as rC3,10A. All three product ions are listed in Table 10, along with the flavonoids (and their classes) that produced them.

Product Ion	Name	Nomenclature	Class
<i>m/z</i> 153	Quercetin	$rC^{1,4}_A$	Flavonol
	Morin		
	Isorhamnetin		
	Kaempferide		
	Kaempferol		
	Galangin		
	Apigenin		Flavone
	Luteolin		
	Chrysin		
	Aromadedrin		Flavanolol
	Eriodictyol		Flavanone
	Naringenin		
Genistein	isoflavone		
<i>m/z</i> 167	Rhamnetin	$rC^{1,4}_A$	Flavonol
	Sakuranetin		Flavanone
	Genkwanin		Flavone
<i>m/z</i> 151	Gallocatechin	$rC^{3,10}_A$	Flavanols
	Epigallocatechin		
	Catechin		
	Epicatechin		

Table 10 The key diagnostic product ions, with the flavonoids (and their classes) where they were observed. See Table 8 for the colour code.

Table 11 is a demonstration of the diagnostic properties of the product ions. This allows the observation of the structural correlation between precursor and product ions. Quercetin, sakuranetin, and epigallocatechin are used as examples. All three product ions are coloured to highlight the exact structural resemblance to their respective precursor ions. Therefore, it can be concluded that the observation of *m/z* 153 demonstrates the presence of two OH groups on ring A (at C-6 and 8) and a C=O at C-4 in ring C. Observation of *m/z* 167 indicates a presence of an OH group on C-6, and a methoxy group on C-8, as well as a ketone on C-4 on ring C. Finally, product ion *m/z* 151, shows the presence of no ketone in ring C, i.e., C-4 is an unoxidized CH₂, but two OH groups on ring A on C-6 and C-8. The presence of these product ions gives a clear insight into the structures of their respective

flavonoid precursor ions. It will have a substantial application in the structural elucidation of unknown flavonoids in future work.

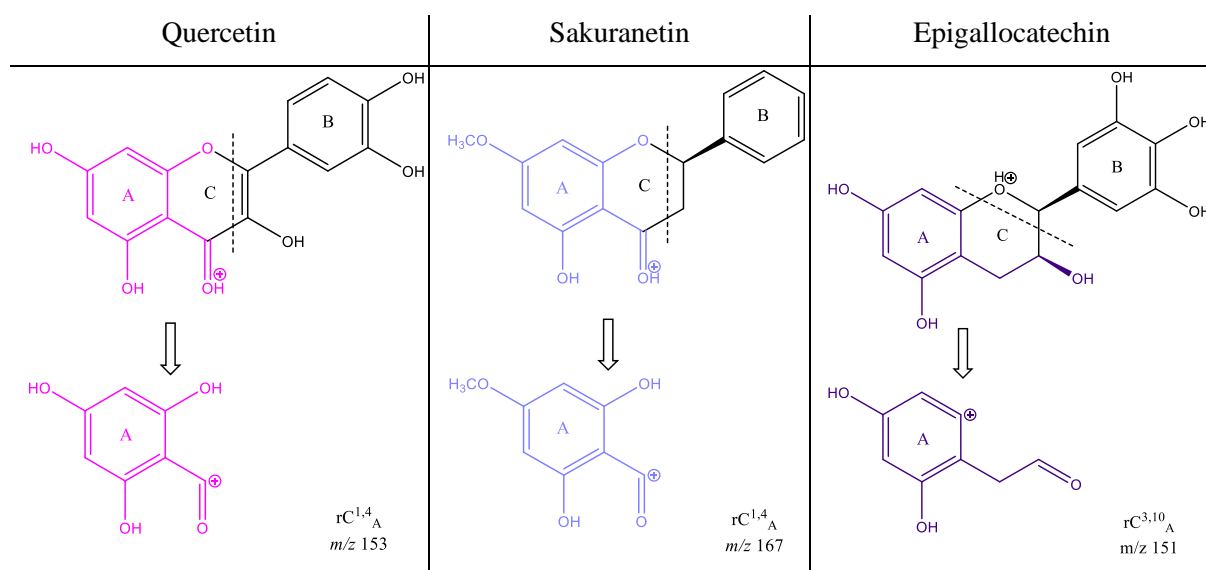


Table 11 The production of the diagnostic product ions m/z 153, m/z 167 and m/z 151 exemplified by quercetin, sakuranetin and epigallocatechin. These ions provide direct insight into the structure of the protonated precursor flavonoids.

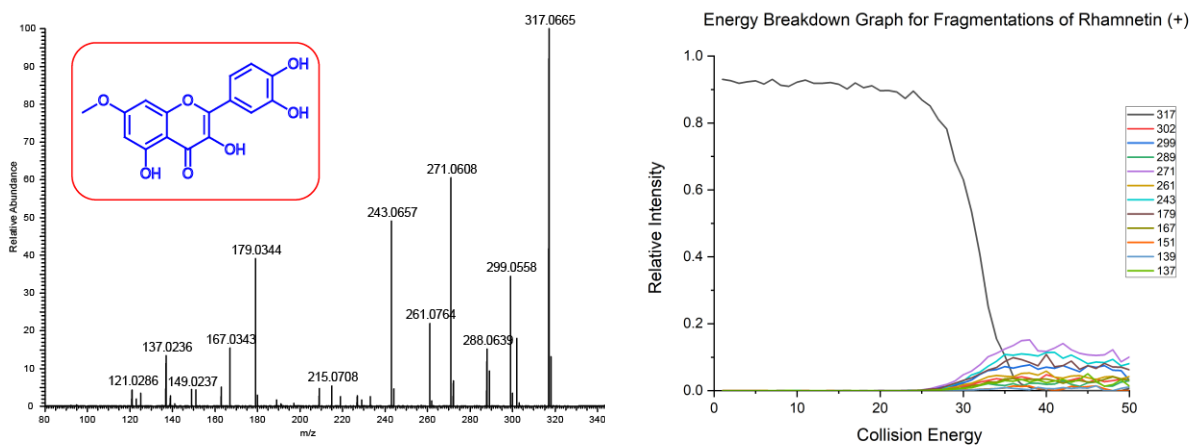
3.3.4. Energy Breakdown Graphs

Energy breakdown graphs are plots of collision energy versus relative intensity for the selected product ions. This is sometimes referred to as ‘energy resolved mass spectrometry’ or ‘2D mass spectrometry’ as they are produced by recording product ion spectra at every collision energy across a range of energies. These were successfully used to aid the interpretation of the fragmentation of the α -amino acids in Chapter 2 and are therefore employed here for the same reason. They also show the different energy thresholds for alternative fragmentation routes and showcase secondary losses during the fragmentation processes. Energy breakdown graphs have been used previously for various studies^{229,230,231}, as well as for the study of flavonoids.²³² During this project, energy breakdown graphs were plotted for all 30 flavonoids from their CID-MS/MS analysis. The following section selects three pairs of isomeric flavonoids for a detailed discussion.

i. Rhamnetin and Isorhamnetin +ve CID MS/MS

Rhamnetin and isorhamnetin are structural isomers differentiated by the location of the methoxy group (see Figure 44 for structures). Each of them has one methoxy group and four OH groups. Both rhamnetin and isorhamnetin glycosides have been discovered in cloves, and their pharmacological uses as potential therapy for type 2 diabetes have been investigated in the literature.²³³ Their energy breakdown graphs shown in Figure 44 indicate a degree of difference. Firstly, there are fewer product ions for isorhamnetin, and secondly, the difference in the intensity of product ion m/z 302. It is almost four times more intense for isorhamnetin compared to rhamnetin. These observations are both due to the intensity of the loss of the methyl radical. In the case of rhamnetin, the losses of H₂O and CO are in competition and a much-preferred route of fragmentation. Except for the common losses of the methyl radical and of a CO, rhamnetin and isorhamnetin have noticeably different fragmentation pathways, despite being positional isomers.

(a) Rhamnetin



(b) Isorhamnetin

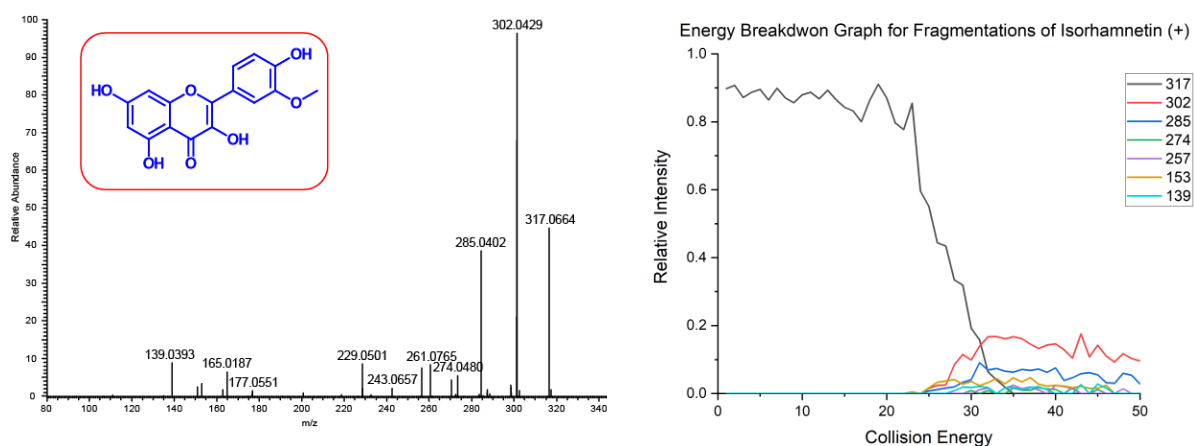
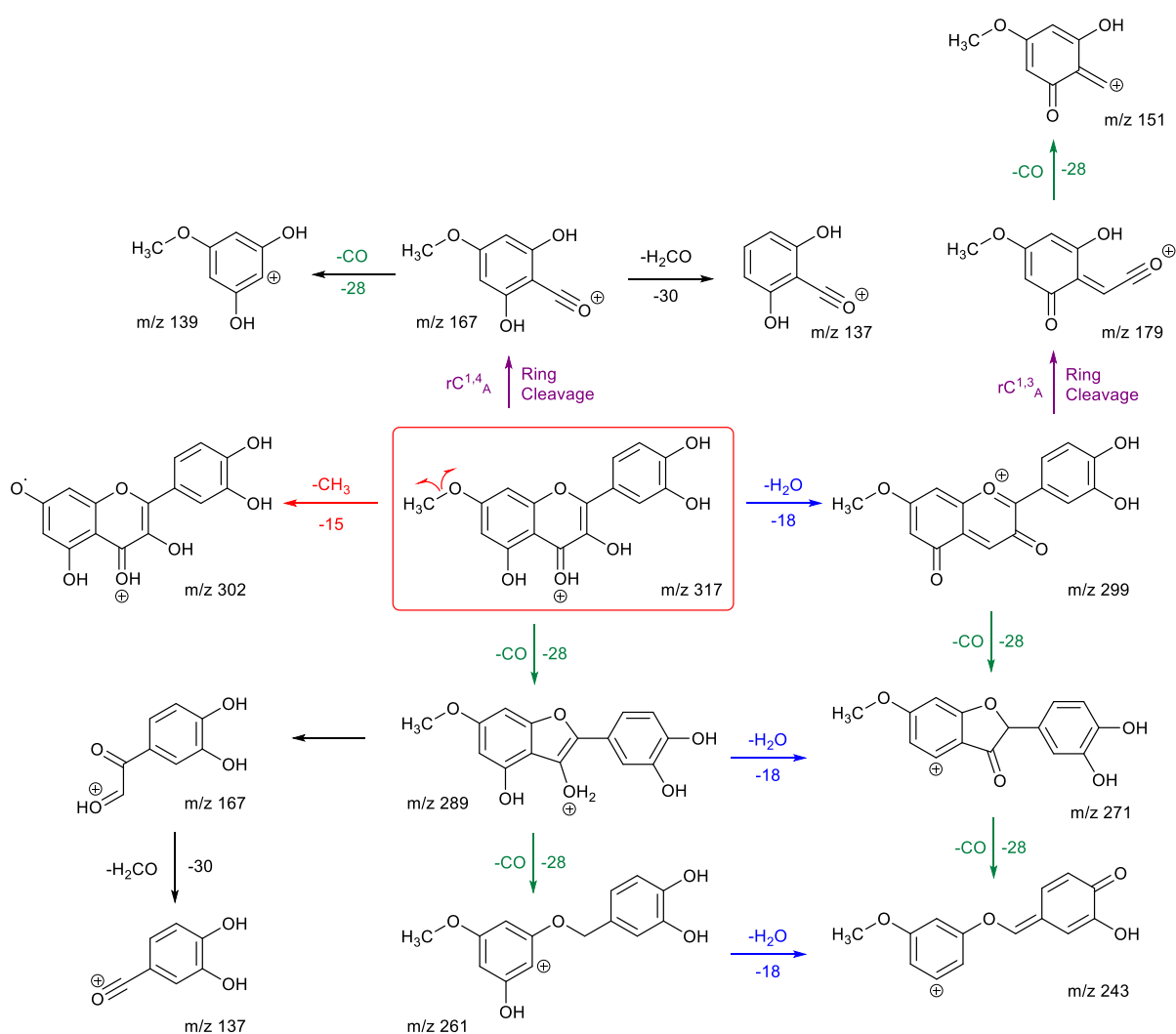


Figure 44 Positive ion mode CID MS/MS spectra, energy breakdown graphs and structures for (a) rhamnetin and (b) isorhamnetin.

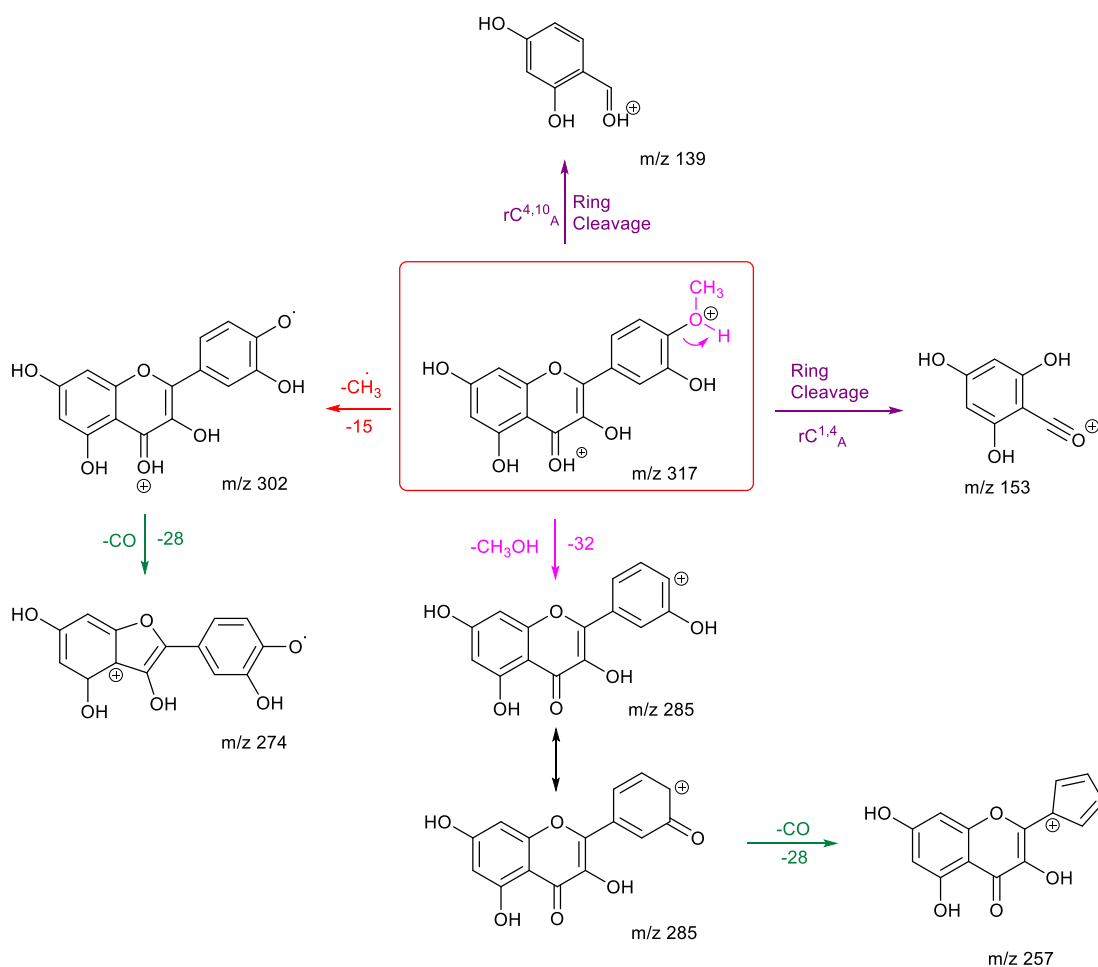
The fragmentation pathways and structures of the product ions are proposed in Scheme 24 and Scheme 25. The protonation site is on carbonyl oxygen of ring C unless stated otherwise. Scheme 24 shows multiple losses of CO from rhamnetin, either after an initial loss of H₂O or from the precursor ion. The competition between the loss of a methyl radical and other routes makes this route less favourable. The two ring cleavages are marked in purple. The first occurs after a water loss to produce *m/z* 179, which is the third most intense peak in the energy breakdown graph. The second ring cleavage results in *m/z* 167. However, this peak can also be produced through a CO loss and a C-O bond breakage. Both routes further lose COH₂ to produce *m/z* 137, although the structure on the bottom left is more likely for product ion *m/z* 137.

Isorhamnetin has fewer product ions than rhamnetin, with *m/z* 302 as the most prominent. This can be explained by the formation of a stabilised radical, the structure of which is shown in Scheme 25. Compared to the structure of ion *m/z* 302 for rhamnetin, *m/z* 302 for isorhamnetin is stabilised by the adjacent hydroxy group and conjugation through rings B and C. The production of *m/z* 274 further proves this, as product ion *m/z* 302 is stable enough to undergo secondary fragmentation by loss of CO. A ring closure is proposed for ion *m/z* 274 to explain why it does not fragment further. The second most intense peak is *m/z* 285, which results from the loss of methanol. Protonation is proposed on the methoxy oxygen to induce an ortho elimination. This specific loss for isorhamnetin is diagnostic for the presence of a methoxy group on ring B at the ortho position, which agrees with previous literature.²³² Kaempferide, which has a methoxy group at the same location, exhibits the same mechanism.²³⁴

The structural differences between rhamnetin and isorhamnetin are also shown in the location of ring cleavages. Isorhamnetin has two ring cleavages $rC^{1,4}_A$ and $rC^{4,10}_A$. The first occurs at ring C atoms 1 and 4 to produce ion m/z 153. As mentioned previously, this m/z 153 ion is a diagnostic ion for ring A with two OH groups. The second ring cleavage occurs at ring C atom locations 4 and 10, forming product ion m/z 139. This is very different to rhamnetin, and it shows that even though both flavonols have the same number of OH groups and a methoxy group, the location of these functional groups significantly affects the formation of the product ions. The relative stability of the different radical ions produced influences the entire spectra. The stability of the ring B product ions also controls which fragmentation pathways would be followed.



Scheme 24 Proposed fragmentation pathway for rhamnetin in positive ion mode. Product ions are m/z 302, 299, 289, 271, 261, 243, 179, 167, 151, 139 and 137.



Scheme 25 Proposed fragmentation pathway for isorhamnetin in positive ion mode. Product ions are m/z 302, 285, 274, 257, 153 and 139.

ii. Kaempferol and Fisetin +ve CID MS/MS

As with rhamnetin and isorhamnetin, both kaempferol and fisetin belong to the flavonol class. Kaempferol is found in many fruits and vegetables, including broccoli, tea and nuts.²³⁵ Reports have been published on its anticancer activity for various organs, including breast, ovarian, gastric, lung, pancreatic, and blood cancers.²³⁶⁻²³⁸ Its method of action includes interfering with the signalling pathways needed for the survival of cancer cells. In breast cancer, kaempferol causes cell death by an extracellular signal-regulated kinase-dependent mechanism.²³⁹ Fisetin is also present in different fruits and vegetables, it has been reported to act on multiple pathways in the reduction of the impact of ageing and diseases on the central nervous system, and it has been accepted as a nutritional dietary supplement in the US.²⁴⁰

Both kaempferol and fisetin have four OH groups in their structures. However, kaempferol has two OH groups on ring A, whereas fisetin has only one. The opposite is true for ring B – (see structures in Figure 45). Differences in their product ions can be viewed in their energy breakdown graphs. Kaempferol has a dominating ion at m/z 177, resulting from a ring cleavage. In comparison, fisetin favours the loss of mass 46 (losses of H_2O and CO) to produce the ion at m/z 241. Product ion m/z 177 occurs at a much lower collision energy than the others for kaempferol. The product ions of fisetin all occur at similar collision energies as can be seen by the energy breakdown graph.

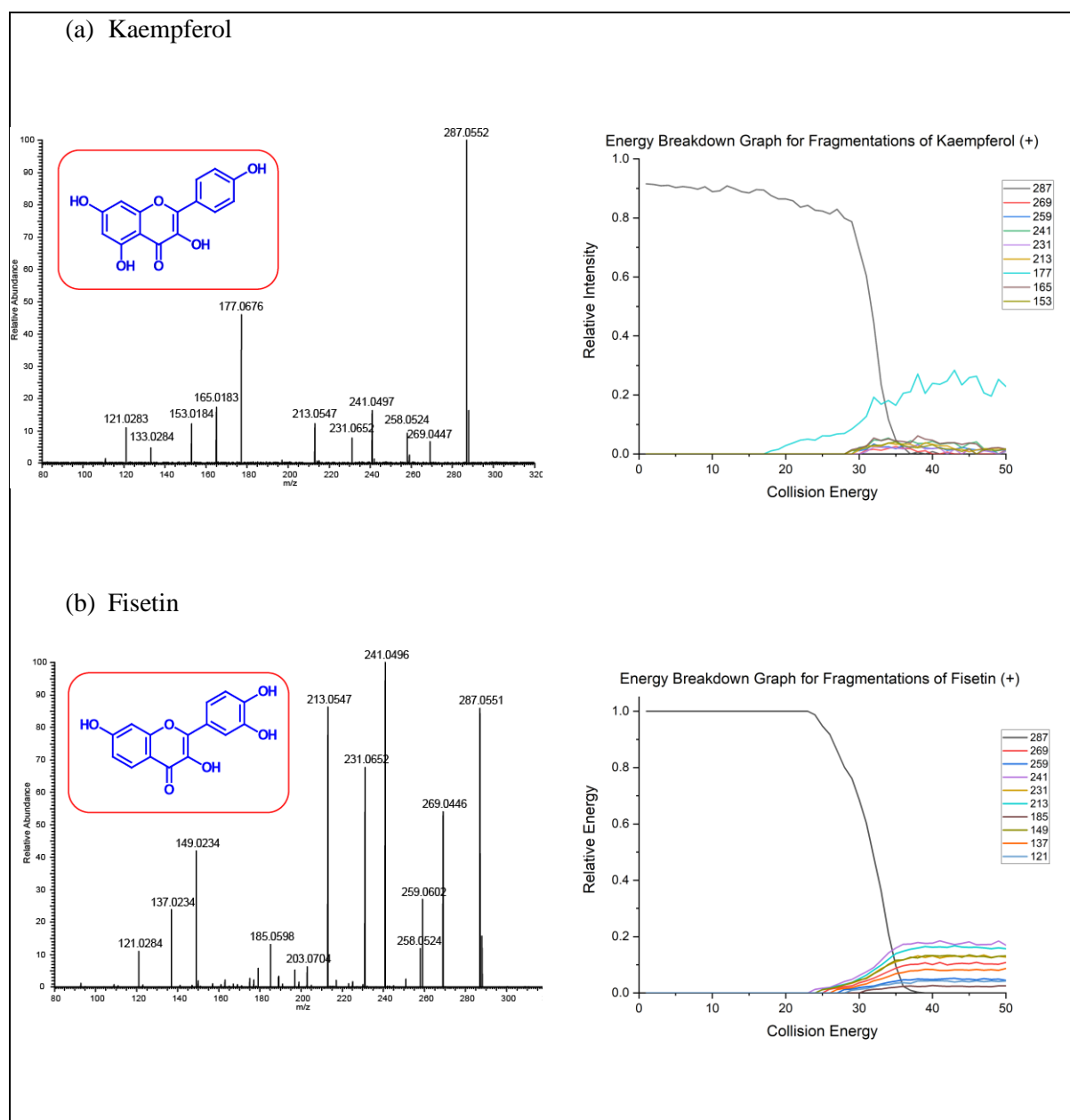
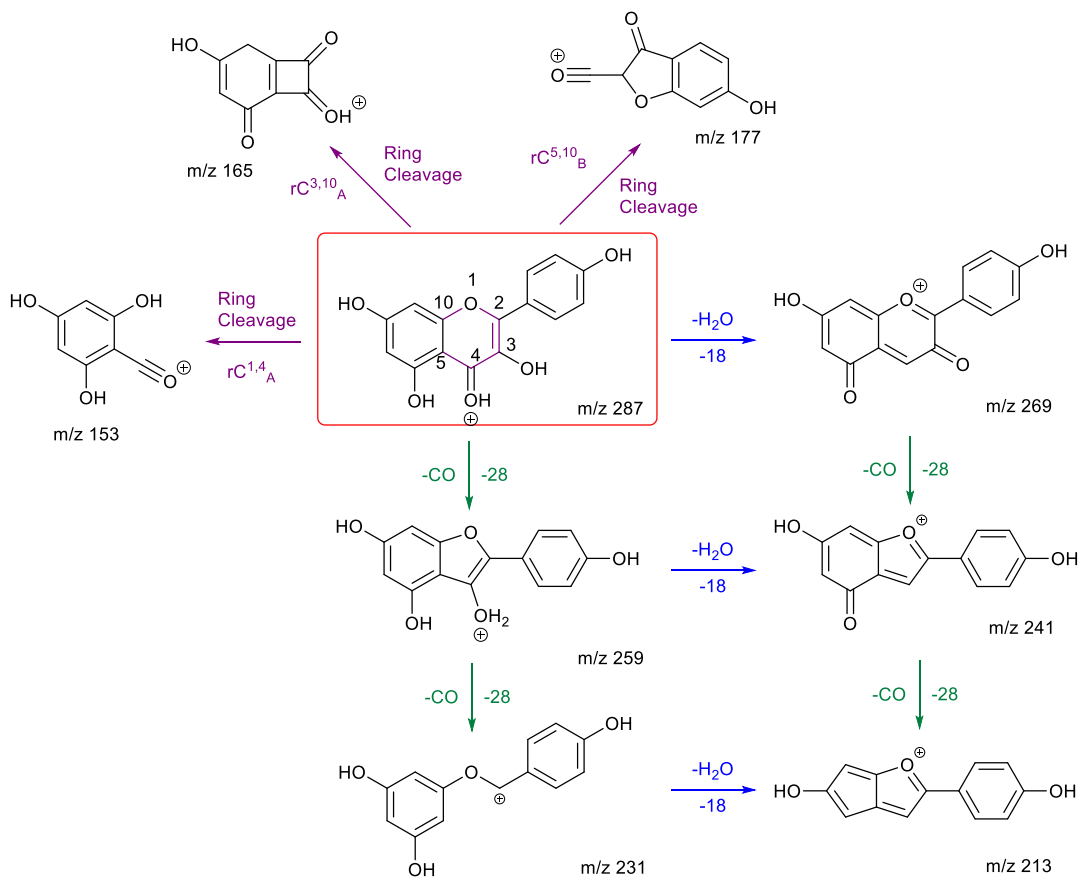


Figure 45 Positive ion mode CID-MS/MS spectra, energy breakdown graphs and structures for (a) kaempferol and (b) fisetin.

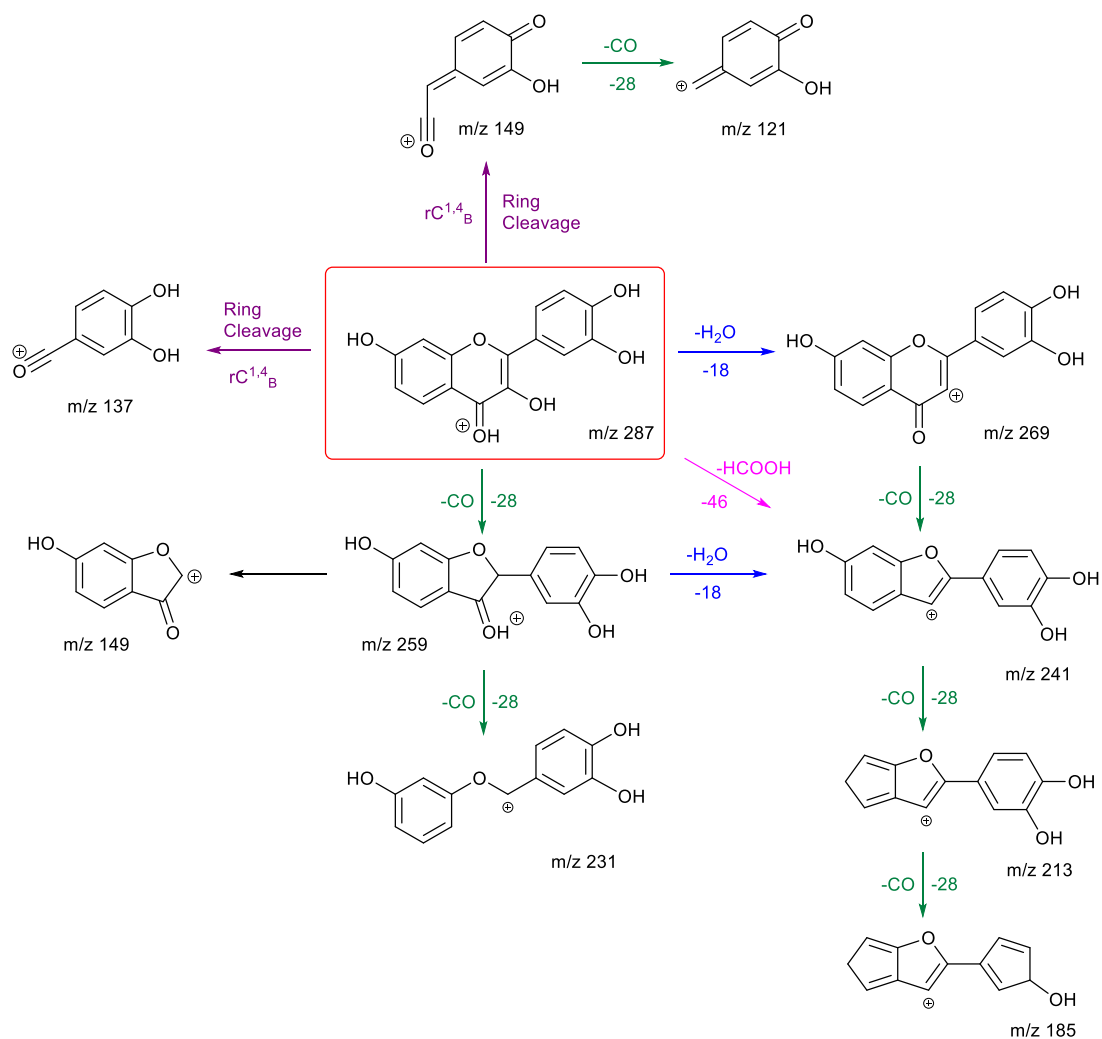
Fragmentation pathways for kaempferol and fisetin are proposed (Scheme 26 and Scheme 27) to gain insights into their product ion structures and rationalise their energy breakdown graphs. Ring C is labelled for clarity. Kaempferol has extensive ring cleavages, and the bonds broken are coloured purple. The high peak intensity of product ion m/z 177 could be due to a stabilised five-membered ring structure with a tertiary carbocation. The diagnostic product ion m/z 153 is seen as kaempferol has the same structure on ring A with two OH groups on atoms 6 and 8. Extensive H₂O and CO losses are also observed.

Fisetin follows the same number of water and CO losses to produce ions m/z 269, 259, 241, 231 and 213. From its energy breakdown graph, the ion at m/z 241 is the most intense and *occurs* at a relatively lower energy than the others. This suggests that it may be due to the loss of formic acid (HCOOH) in one step, as marked in lilac in Scheme 25, in addition to the first loss of H₂O then, followed by a CO. It then loses a CO to produce ion at m/z 213. This agrees with the observation that the ion m/z 213 is the second most intense peak (colour: aqua) in the energy breakdown graph. Compared to kaempferol, it has one less ring cleavage, and both of the product ions from these ring cleavages leave ring B intact. Product ion m/z 149 has two fragmentation pathways, with one occurring directly from the cleavage of ring C at atoms 1 and 4, and the other is by a carbon-carbon bond breakage from ion m/z 259. The first one is more likely according to its energy breakdown graph. Both routes can be followed by another loss of CO to give product ion m/z 121. Lastly, peak m/z 185 has the lowest intensity and requires the highest collision energy. This can be explained by a less favourable charge remote fragmentation and the need to break an already stabilised ring structure of m/z 213.

To summarise, kaempferol and fisetin differ in their locations, numbers, and remaining structures of the ring C cleavages. Product ions of these ring cleavages are key to their structural identification. Kaempferol has a diagnostic product ion m/z 153 which can distinguish itself from fisetin due to the location of the OH group.



Scheme 26 Proposed fragmentation pathway for kaempferol in positive ion mode. Product ions are m/z 269, 259, 241, 231, 213, 177, 165 and 153.



Scheme 27 Proposed fragmentation pathway for fisetin in positive ion mode. Product ions are m/z 269, 259, 241, 231, 213, 185, 149, 137 and 121.

iii. Chrysin and Daidzein +ve CID MS/MS

Chrysin belongs to the flavone class. It is often used as a fabric dye to yield bright yellow, pale yellow-orange, and chocolate brown shades on wool and silk.²⁴¹ Daidzein, present mainly in soy, is an isoflavone classed as a phytoestrogen because it has oestrogen-like properties. It has been reported to have weak oestrogenic and antiestrogenic effects and can provide some protection against osteoporosis.²⁴² Chrysin and daidzein are selected in this report as they are structural isomers but belong to different flavonoid classes. Both of their protonated molecular ions occur at m/z of 255. Their energy breakdown graphs are presented in Figure 46.

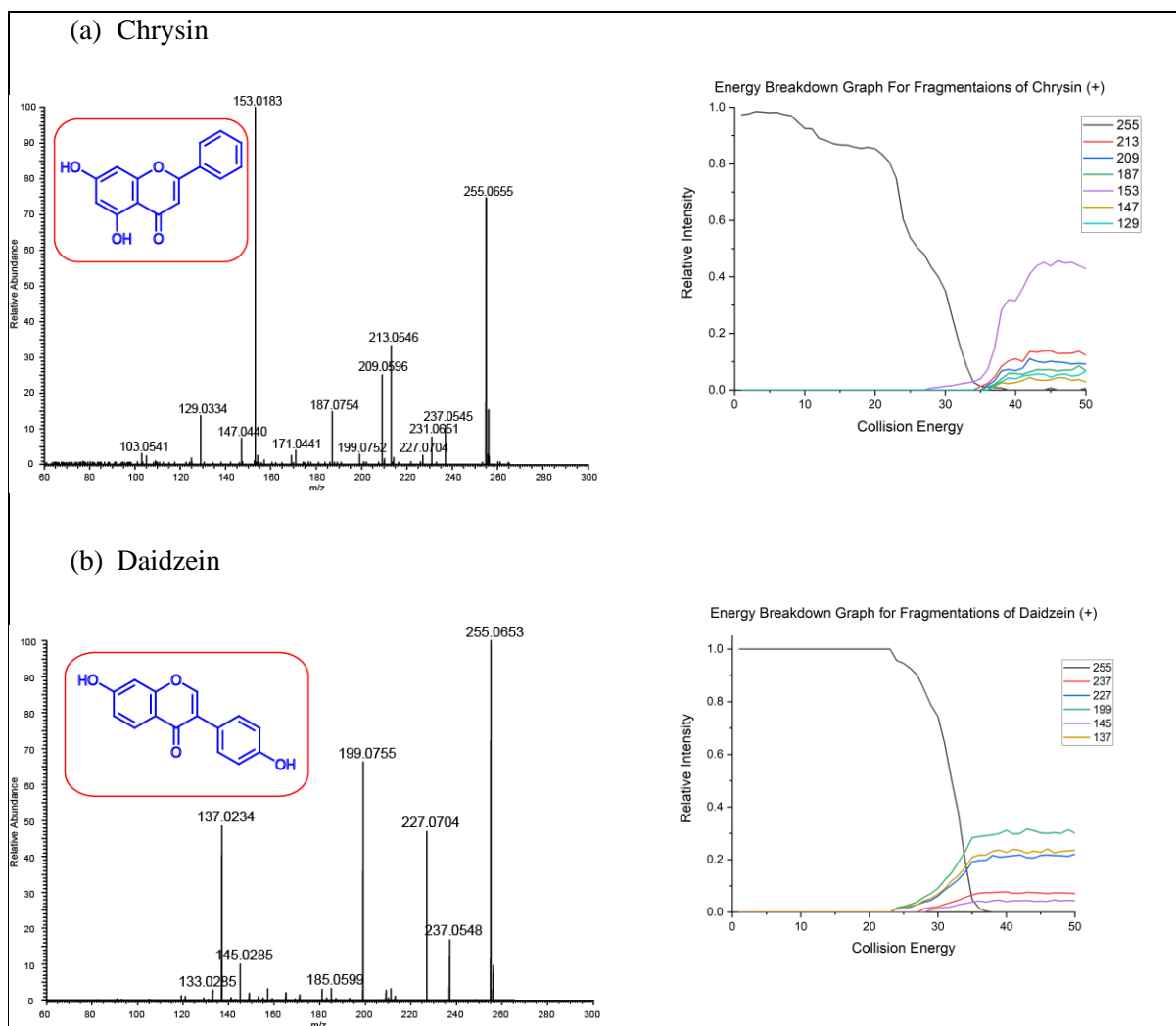


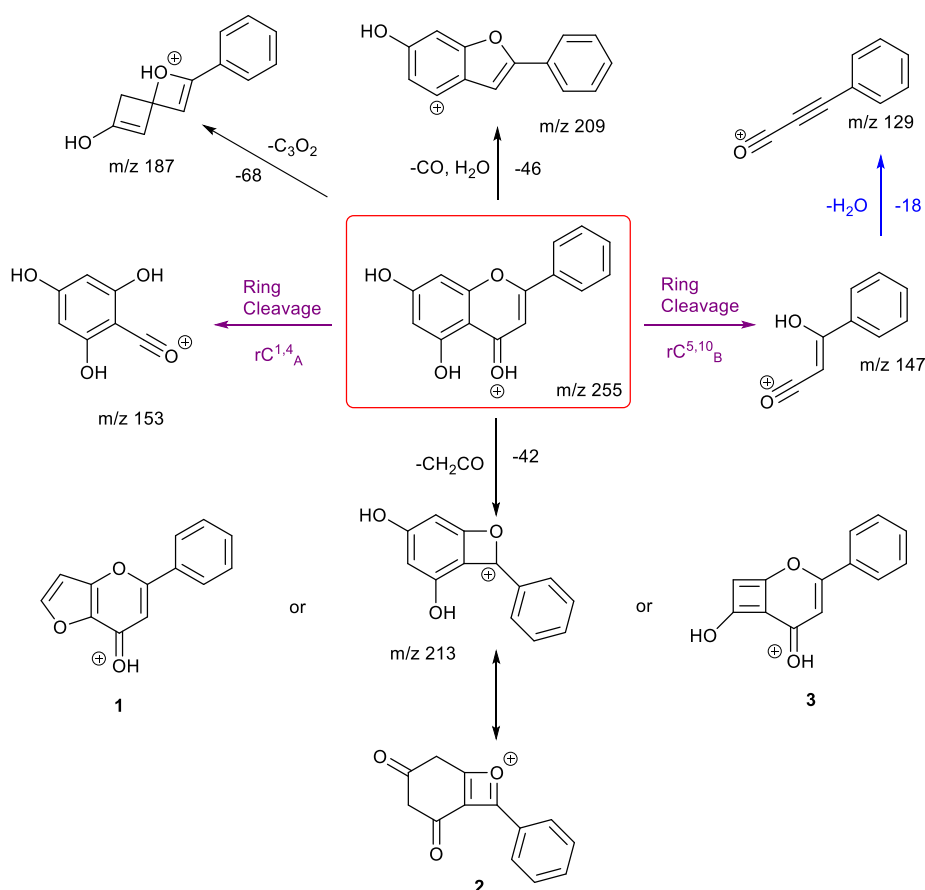
Figure 46 Mass spectra, energy breakdown graphs and structures for (a) chrysin and (b) daidzein in positive ion mode.

Figure 46 gives very distinctive energy breakdown graphs of chrysin and daidzein. The most intense peak for chrysin is the product ion m/z 153, the diagnostic peak from a specific structure on ring A. This peak also first appears at lower collision energy compared to others. In contrast, daidzein has three product ions at the same collision energy. These are m/z 199, 137 and 227 (in order of their peak intensities). As shown by Gaussian calculations in this research, which will be discussed in the next section, the ketone oxygen on ring C has the highest proton affinity, followed by the hydroxy oxygen on ring B. As a result, product ion formation at m/z 227 is much more favoured when compared to m/z 237 for daidzein.

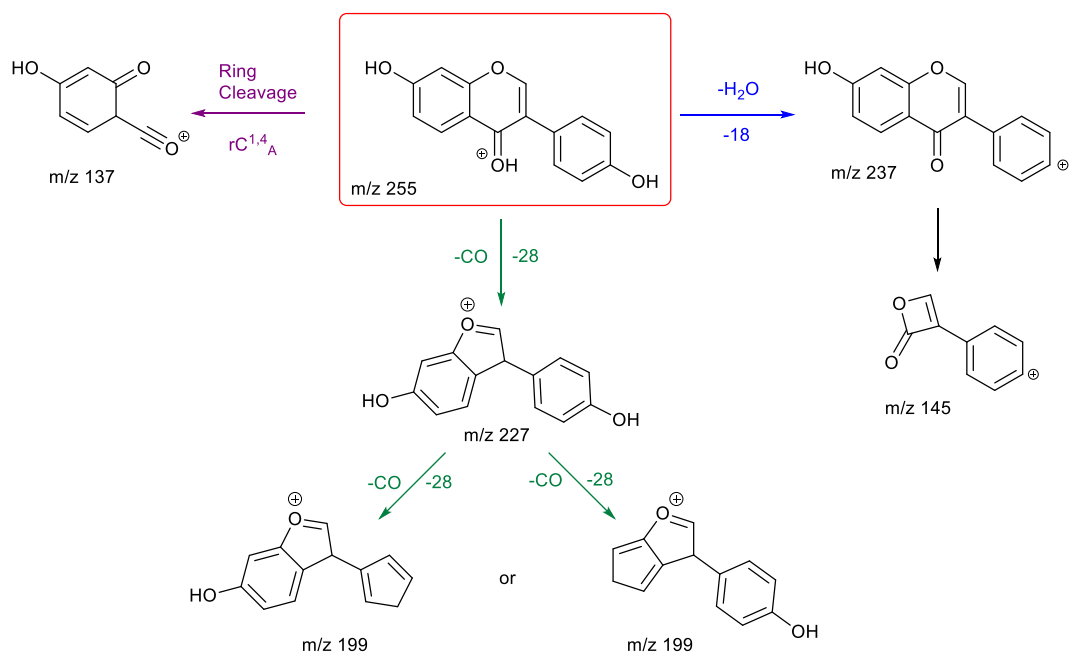
Scheme 28 and Scheme 29 are proposed to show the fragmentation pathways of chrysin and daidzein. Chrysin has two ring cleavages, including the diagnostic product ion for ring structure A. Also, it has a loss of formic acid to produce ion m/z 209. There are also extensive ring contractions for chrysin which agrees with previous literature.²⁴³ The ion m/z 187 is generated from a very unusual loss of carbon suboxide (C_3O_2), first observed in 2001 for flavones and later in a few other studies both in positive and

negative ion mode.²⁴³⁻²⁴⁵ Although no mechanisms have been proposed. This could be an opportunity for future work to investigate a viable mechanism in combination with computational modelling. Also, a loss of CH₂CO gives three possible structures for product ion *m/z* 213. Structure **3** has been shown previously in the literature.²⁴³ However, this research believes structure **1** is more stabilised. This has been shown by Gaussian modelling and could be explained by the aromatic rings A and C.

Daidzein has only one ring cleavage from the protonated precursor ion, resulting in product ion *m/z* 137. The most intense peak at *m/z* 199 is produced from the loss of a CO from *m/z* 227. Two possible structures are proposed for product ion *m/z* 199. Both involve a ring contraction to form cyclopentadiene in their structure, depending on which CO is lost during fragmentation. Product ion *m/z* 145 has the lowest intensity, possibly due to a less favourable charge remote ring cleavage from ion *m/z* 237. In summary, ring cleavage on ring C is very energetically favourable for both chrysin and daidzein. Chrysin has extensive ring contractions, resulting in the unusual neutral loss of C₃O₂. Daidzein also can undergo ring contraction to lose a CO.



Scheme 28 Proposed fragmentation pathway for chrysin in positive ion mode. Product ions are *m/z* 213, 209, 187, 153, 147 and 129.



Scheme 29 Proposed fragmentation pathway for daidzein in positive ion mode. Product ions are m/z 237, 227, 199, 145 and 137.

3.3.5. Common Losses Negative Ion Mode

MS/MS data for the 30 flavonoids in negative ion mode were also investigated, and neutral/radical losses of their deprotonated precursor ions were summarised in Table 12. Flavonoids are grouped and marked by different colours based on their classes. **Purple:** Flavonol; **Blue:** Flavanone; **Green:** Isoflavone; **Brown:** Flavone; **Red:** Flavanol and Grey: Flavanolol. The table is populated with their deprotonated molecular ion mass $[M-H]^-$, the mass range is from m/z 237 to 315. Common neutral losses are also listed for comparison and mapped with colour intensity.

Name	Deprotonated Precursor Ion m/z	Common Losses														
		28	44	62	72	84	108	122	150	180	194					
Quercetin	301	28	44	62	72	84	108	122	150	180	194					
Morin	301	18	28	44	62	72	114	126	138	150	152	154	176	194		
Rhamnetin	315	15	16	28	44	122	150									
Isorhamnetin	315	15	16	30	44	72	150	164								
Kaempferide	299	15	44													
Kaempferol	285	28	30	42	44	46	56	71	72	98	116	134	178			
Fisetin	285	28	122	150												
Galangin	269	28	42	44	46	56	68	72	86	100	116	126				
Genkwanin	283	15														
Hispidulin	299	15														
Wogonin	283	15														
Diosmetin	299	15														
5-Hydroxyflavone	237	28	42	54	66											
Baicalein	269	18	28	44	46	72	74	100								
Apigenin	269	42	44	68	72	86	88	118	120							
Luteolin	285	18	29	42	44	62	68	72	78	86	110	134	152			
Scutellarein	285	18	28	44	46	56	62	68	72	74	86	100	108	120	148	
Chrysin	253	42	44	68	72	88	102	110	146							
Gallocatechin	305	18	44	58	62	84	86	102	114	126	138	140	166	168	180	
Epigallocatechin	305	18	44	58	62	84	86	102	114	126	138	140	166	168	180	
Catechin	289	44	84	86	110											
Epicatechin	289	44	84	86	110											
Aromadetrin	287	28	44													
Daidzein	253	18	28	42	44	56	72	93	118	120						
Genistein	269	28	42	44	56	68	72	86	88	100	110	136				
Eriodictyol	287	136														
6-Hydroxyflavanone	239	18	27	28	42	44	78	104								
7-Hydroxyflavanone	239	42	91	104	130											
Sakuranetin	285	94	120	166	192											
Naringenin	271	94	120	146	152	164	178									

Table 12 List of the flavonoids studied, their deprotonated molecular precursor ion m/z, and common neutral losses observed. The flavonoids are coloured by class: Purple: Flavonol; Blue: Flavanone; Green: Isoflavone; Brown: Flavone; Red: Flavanol; and Grey: Flavanolol.

The first trend is consistent with the observations in the positive ion mode. Flavonoids from the same classes do not always exhibit the same common losses. They only have the same common losses if they are structural isomers from the same flavonoid class. For example, in the flavone group, Genkwanin and Diosmetin only produce a loss of 15, corresponding to a CH₃ radical loss. Gallocatechin and epigallocatechin, catechin and epicatechin, these two pairs of structural isomers are all flavanols. Hence they have the same neutral losses.

Secondly, compared to the positive ion mode CID MS/MS data, there are more lower mass neutral losses. The most common ones are summarised in Table 13. Like the positive ion mode, common losses in the negative ion mode also depend on the hydroxy and methoxy group locations. Other than the predominantly occurring loss of H₂O, loss of CO is frequently observed. Any losses higher than 100 mass units will likely indicate ring cleavage on ring C.

Common Losses	
Formulae	Mass
CH ₃ [·]	15
CH ₄	16
H ₂ O	18
CO	28
C ₂ H ₂ O	42
C ₂ O	44
H ₂ O+CO	46
2CO	56
CO ₂ +H ₂ O	62
H ₂ O+2CO	74
3CO	84

Table 13 Identification of some of the most common neutral losses observed for negative ion mode ESI-MS/MS analysis of the 30 flavonoids studied.

The negative ion mode MS/MS results have one major difference from the positive ion mode spectra. This is the loss of 16, which corresponds to a CH₄ methane loss. This has been previously published in the work from this research group, and now more data from this study is strong evidence to support this observation.²⁴⁶ Flavonols with neutral loss of CH₄ are highlighted in the box from Table 12. These are discussed in the next two sections.

3.3.6. Loss of CH₄

During the investigation of negative ion mode MS/MS spectra of flavonols, it was discovered that rhamnetin and isorhamnetin both have a loss of 16 corresponding to the loss of methane and confirmed by accurate mass measurements. This observation of CH₄ elimination is very interesting and has been published before with heterocyclic aromatic amines²⁴⁷ and the flavanone Hesperetin but not with any other flavonoids. Hence, this is discussed in more detail here.

Out of all the flavonoids with a MeO group, only rhamnetin and isorhamnetin exhibits a loss of CH₄ in their MS/MS spectra, whereas the other flavonol: kaempferide, flavones: genkwanin, hispidulin, wogonin and diosmetin do not. However, all the flavonoids mentioned share a common loss of CH₃[·], a methyl radical. This could be because the unique CH₄ loss requires a hydroxyl group on the meta carbon in ring C, which the flavone groups are lacking. An OH group on the neighbouring carbon is also essential to facilitate this mechanism. This explains why although kaempferide has a hydroxyl group on the meta carbon in ring C, and a MeO group on ring B, the lack of an OH group on the same ring prevents a loss of 16.

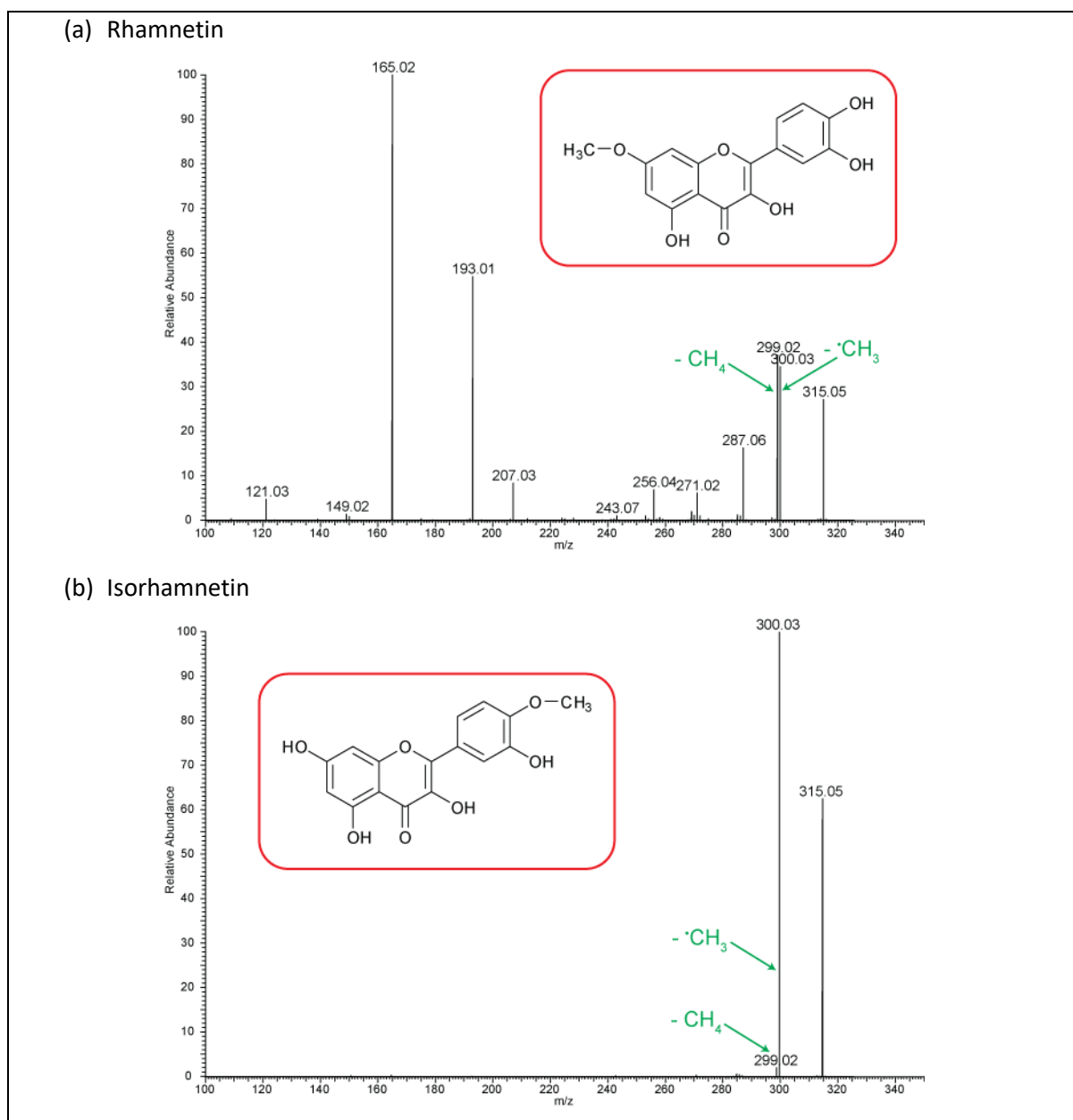
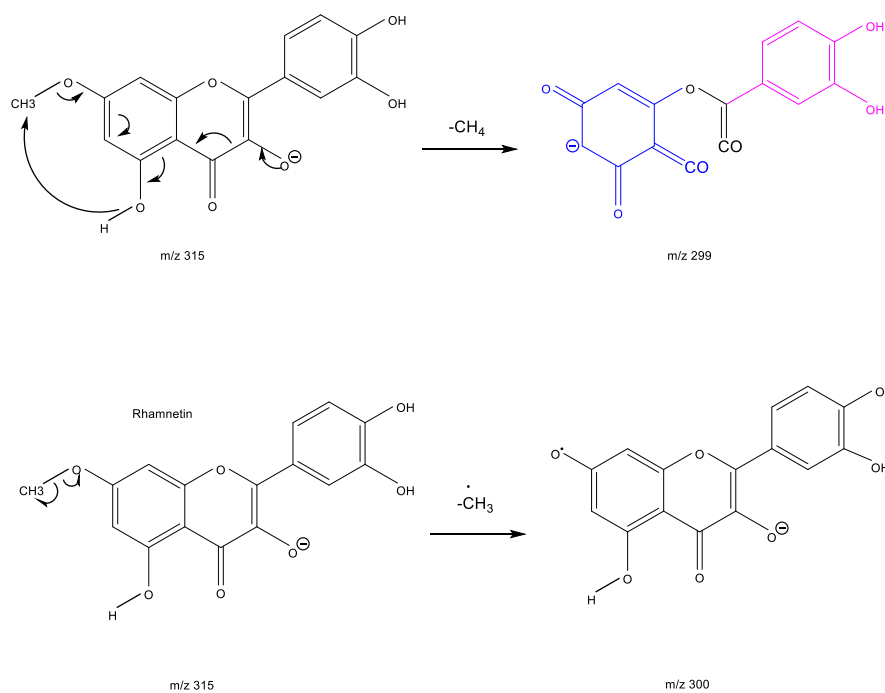
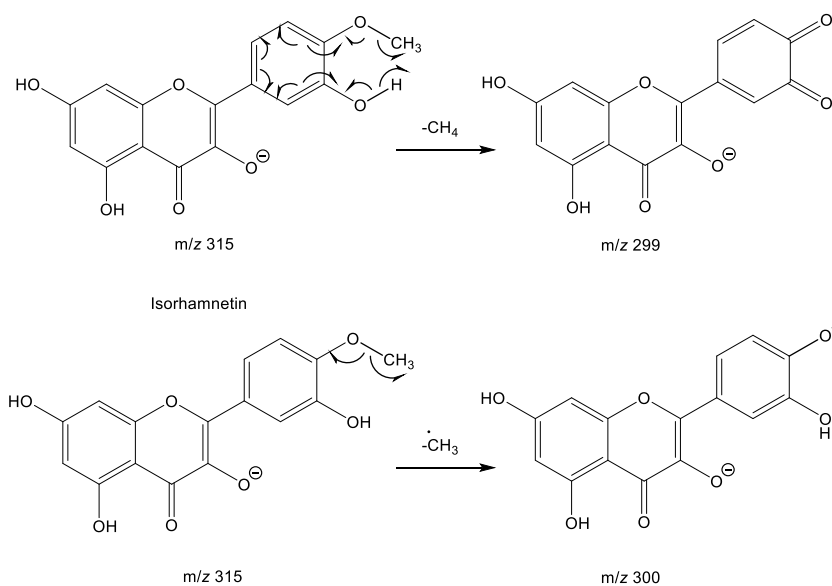


Figure 47 MS/MS spectrum of (a) rhamnetin and (b) isorhamnetin in negative ion mode. Peaks of m/z 300 and m/z 299 are highlighted, m/z 300 loss of CH_3 , m/z 299 loss of CH_4 .

Figure 47 shows MS/MS spectra for rhamnetin and isorhamnetin at collision energy 30eV and 27eV in the negative ion mode. Although rhamnetin and isorhamnetin are structural isomers and belong to the same flavonol group, their MS/MS spectra are distinctively different. They both have losses of CH_3 (m/z 300) and CH_4 (m/z 299). Compared to isorhamnetin, m/z 299 of rhamnetin is at a much higher intensity and is more intense than m/z 300. In contrast, for isorhamnetin, CH_3 loss is a much more favourable route.



Scheme 30 Proposed fragmentation pathway for rhamnetin in negative ion mode for the loss of CH₄ (top) and loss of CH₃ (bottom).



Scheme 31 Proposed fragmentation pathway for isorhamnetin in negative ion mode for the loss of CH₄ (top) and CH₃ (bottom).

Scheme 30 and 31 are the proposed fragmentation pathways for rhamnetin and isorhamnetin for generating product ions *m/z* 299 and 300 in the negative ion mode. In Scheme 30, the negative charge is proposed to be on the meta oxygen in ring C, facilitating a hydride shift to eliminate CH₄ on ring A. This also promotes a ring opening between C3 and C4 on ring C resulting in the structure demonstrated

in Scheme 31 for ion m/z 299. Fragmentation of this m/z 299 product ion further generates m/z 193 and m/z 165. For isorhamnetin, due to the different positions of the MeO and OH group, an alternative mechanism is proposed, as published in the literature for hesperetin.²⁴⁶ Unlike the 1,5 hydride shift, this mechanism is less energetically favourable and could be a reason for the lower intensity of peak m/z 299 for isorhamnetin. The charge remote fragmentation and the lack of ring opening could also be contributing factors to this lower intensity in Figure 47.

3.3.7. Competition of $\text{CH}_3\cdot$ and CH_4 Losses

Examining the energy breakdown graphs for rhamnetin and isorhamnetin in the negative ion mode brings our attention to their contrasting results. Both graphs were recorded with CID energy from 0-50eV for deprotonated molecular ion peak m/z 315.

Figure 48 shows distinctive differences in the product ions for these two flavonols, despite the similarity in their structures. Rhamnetin produces more intense peaks for product ions, especially in the lower mass range (m/z 165). In contrast, the loss of CH_3 dominates the isorhamnetin spectrum. In the breakdown graph for rhamnetin, peak m/z 299 started to occur at CID energy 15eV and m/z 300 at just above 20eV. As the collision energy increases, m/z 299 peak intensity also increases. This supports our theory that the loss of CH_4 , in the case of rhamnetin, is more energetically favourable than the loss of CH_3 .

Figure 48 (b) shows very low intensity for the loss of CH_4 for isorhamnetin and other product ions except m/z 300. This observation could be supported by a different mechanism for the loss of CH_4 , which was discussed in the previous section.

The striking differences in the mass spectra between these two structurally isomeric flavonols in the negative ion mode build an excellent foundation for isomerisation differentiation. The unique loss of CH_4 could be used for the structure elucidation of unknown flavonoids in future work.

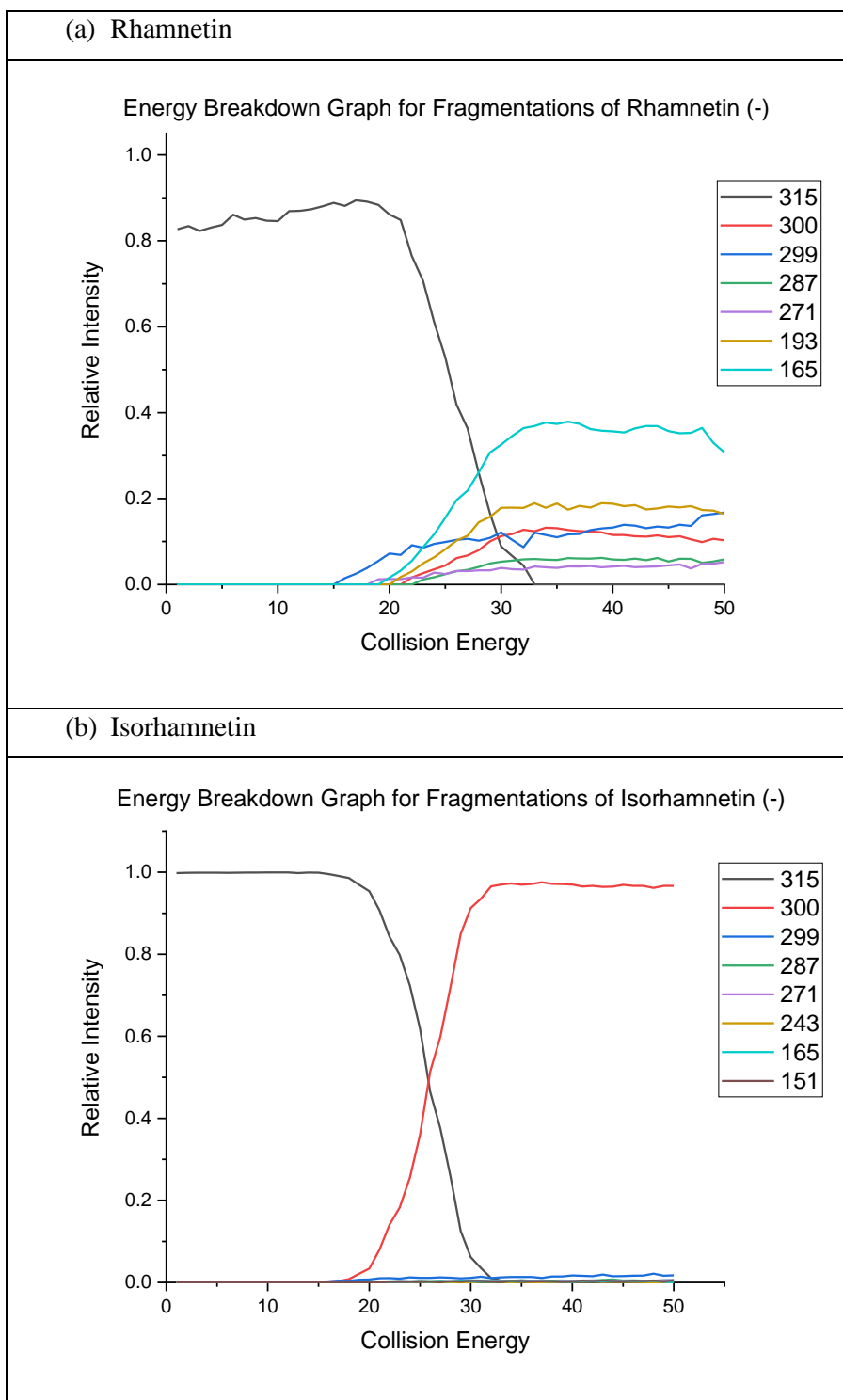


Figure 48 Energy breakdown graphs for (a) Rhamnetin and (b) Isorhamnetin in negative ion mode.

3.3.8. Proton Affinity

Proton affinity is the energy released by the addition of a proton to a system. It can be calculated using the energy difference between the target system and the protonated structure of the same target system. The calculation of proton affinity is useful because flavonoids have multiple possible protonation sites; it indicates which is the most energetically favourable protonation site and allows us to consider how targeted molecule fragments and why. Proton affinities for the selected flavonoids have been calculated.

Results from Gaussian 09 revealed that for the flavonoids calculated, the ketone oxygen is the most energetically favourable protonation site for most of the listed flavonoids (Table 14). This explains the extensive observation of ring C cleavages in their MS/MS spectra. However, it is essential to remember that this is theoretical modelling; in reality, protonation is still possible on other sites, especially other OH groups. Proton scrambling can also occur in CID-MS/MS due to the extra energies involved.

Name	Class	Preferred Protonation Site	PA(kcal/mol)
Kaempferol	Flavonols	Ketone oxygen	228.12
Fisetin	Flavonols	Ketone oxygen	223.15
Chrysin	Flavones	Ketone oxygen	232.68
Daidzein	Isoflavones	Ketone oxygen	228.61
5-Hydroxyflavone	Flavones	Hydroxyl ring A	223.44
6-Hydroxyflavanone	Flavanones	Ketone oxygen	216.14
7-Hydroxyflavanone	Flavanones	Ketone oxygen	215.71
Apigenin	Flavones	Ketone oxygen	223.13
Quercetin	Flavonols	Ketone oxygen or Hydroxy ring C	217.11

Table 14 Calculated proton affinity for selected flavonoids in kcal/mol.

3.4. Conclusion

This chapter investigated the MS/MS spectra for 30 flavonoids in positive ion mode. Common losses are summarised and compared. From these, three diagnostic product ions are identified, these are m/z 153 – indicative of two OH groups on ring A; m/z 167 – indicative of one OH and one methoxy group on ring A; and m/z 151 – a flavanol, with no ketone oxygen but two OH groups on ring A. In addition, a radical loss (loss of $\text{CH}_3\cdot$) has been observed for flavonoids having methoxy groups in their structures. Other common losses include that of H_2O and CO .

Three pairs of structural isomers are selected for detailed discussion using their MS/MS spectra and energy breakdown graphs to help proposed fragmentation pathways. It is shown that although they may share the same common losses, ring C cleavages are often diagnostic due to their distinct A and B ring structures and substitution patterns. Isorhamnetin has a loss of methanol which can be used to indicate the presence of a methoxy group at C-14 on ring B. Isorhamnetin, kaempferol and chrysin all have the same ring A structure. Hence, the presence of the product ion at m/z 153 in their spectra. Extensive ring contractions have been observed in chrysin to facilitate an unusual loss of C_3O_2 . The isoflavone daidzein also has ring contraction to lose a CO in its fragmentation pathway.

The proton affinity is calculated for kaempferol, fisetin, chrysin, daidzein, and a few other selected flavonoids. The results reveal that ketone oxygen is the most energetically favourable protonation site for most flavonoids, which agrees with the experimental observation in their mass spectra and energy breakdown graphs.

4. Polyether

4.1. Polyether Overview

Polyether ionophores are a group of natural products with typically complex structures.²⁴⁸ ‘Ionophore’ refers to an ability to transport cations across biological membranes. The reason polyether ionophores are capable of doing this lies in their unique structure. Firstly, they have two different termini, one with a carboxylic acid and one with a hydroxy group. This allows it to form a hydrogen bond at the terminus and create a cyclical structure. Secondly, they have cyclic ethers, spiroketals and beta-ketols, the oxygens atoms of which can rotate inwards to chelate with metal cations, creating a hydrophilic centre with a hydrophobic outer area to cross the membrane. Once the transportation is complete, the polyether ionophores will release the cations on the other side of the membrane, reverting to their initial linear structure. This mechanism is known as passive diffusion.²⁴⁹ Due to this unique ability, polyether ionophores play a vital role in antibiotics and chemotherapy.²⁵⁰ Interest in the studies of polyether ionophores has increased dramatically since the late 70s, and currently, over 120 naturally occurring fermentation-derived ionophores are known.²⁵¹

Salinomycin (SAL) is a polyether ionophore often used as an anti-coccidiostat in poultry farming. It was approved for use with chickens in 1983.²⁵² Coccidiosis is a parasitic disease that can occur in animals raised in a small area contaminated with coccidial oocysts. It occurs in the intestines of the infected hosts and causes high mortality. Coccidiosis is present in all poultry farms, even those with high sanitary and well-management standards.²⁵² Therefore, using an anti-coccidiostat such as salinomycin is highly beneficial in avoiding economic losses for poultry farmers. Salinomycin and other polyether ionophores could also be used as growth promoters in ruminant animals. Their antibiotic activity improves feed conversion by manipulation of the rumen flora.^{253, 254}

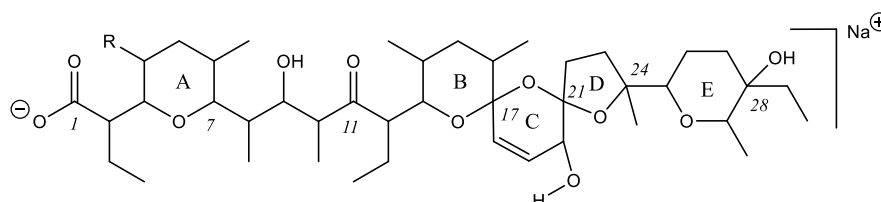
In recent years, salinomycin was found to reduce the proportion of cancer stem cells by 100-fold relative to a commonly used breast cancer chemotherapeutic drug, and treatment of mice with salinomycin showed inhibition in mammary tumour growth *in vivo*.²⁵⁵ Since the publication of this study, an influx of research has been conducted on salinomycin, especially in its cancer treatment applicability. The pathways of how salinomycin could kill cancer stem cells have been investigated.^{256, 257} Salinomycin’s antitumour mechanism is connected to its strong affinity to potassium cations.²⁵⁸ This important conclusion reflects an earlier study on the cation selectivity of salinomycin, which is $K^+ > Na^+ > Cs^+ \gg Ca^{2+}$.²⁵⁹ Salinomycin has also shown its biological activities in other anti-tumoral actions, such as

increasing the death of lung cancer stem cells, sensitising cancer cells to radiation and synergising with dichloroacetate cytotoxicity in colon cancer cells. ²⁶⁰⁻²⁶²

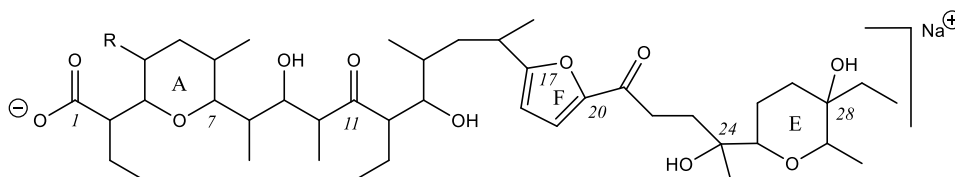
Salinomycin was first reported in 1974 in the Journal of Antibiotics. ²⁶³ It was produced by a strain of *Streptomyces albus* (ATCC 21838) in a tank fermentation. In a later patent publication, the same research team reported that the production of salinomycin from *Streptomyces albus* 80614 was improved by using fatty acid esters or oil and ammonium ions in the fermentation medium. ²⁶⁴ Investigation into salinomycin biosynthesis revealed that the polyketide chain is synthesised on an assembly line of nine polyketide synthase multienzymes. ^{265, 266} The paper also probed the mechanism of polyether ring formation via targeted deletion of the genes that control oxidative cyclisation. ²⁶⁷ This enabled potential structural mutation to salinomycin by biosynthetic methods.

The structure of salinomycin is represented in the figure below. In addition to salinomycin, narasin (NAR) is another polyether antibiotic investigated in this research. Narasin has an additional C-methyl group on the tetrahydropyran ring adjacent to the carboxyl group (carbon 4, ring A) compared to salinomycin, and it is produced by *Streptomyces aureofaciens* NRRL 5758. Narasin is also an effective anticoccidial agent in poultry and improves the efficiency of feed utilisation in ruminant animals. ²⁵²

original form



isomeric form



Salinomycin, R =H : $C_{42}H_{70}O_{11}Na^+$ $m/z = 773$

Narasin, R =Me : $C_{43}H_{72}O_{11}Na^+$ $m/z = 787$

Figure 49 The structures (linear representation) of the natural and isomeric forms of salinomycin (R = H) and narasin (R = Me) sodium salts.

4.1.1. Studies of Polyether Antibiotics by Mass Spectrometry

MS remains a popular technique in analysing polyether ionophores, particularly in determining their presence in animal-related bioproducts and human plasma.²⁶⁸⁻²⁷¹ When coupled with HPLC, MS/MS has proven powerful in the identification of polyether ionophores in previous studies. Compared to other polyether ionophore antibiotics, salinomycin and narasin fragmentation studies are lacking. There is currently no study on salinomycin or narasin metal complexes and no published data on their ion mobility mass spectra.

The first reference on the MS of salinomycin and its derivatives was published in 1976.²⁷² The paper reported three types of fragmentations, which include a characteristic cleavage occurring at ring C of the tricyclic spiroketal system (Figure 50), fragmentation occurring at the hydroxyketone group by McLafferty rearrangement²⁷³, and α positions of the cyclic ether rings. It also noted the low solubility of salinomycin sodium salt in water, resulting from its hydrophobic property at the exterior surface when a metal ion is chelated in the molecule's centre. Upon closer examination of the product ions, the paper identified that the Na^+ is not localised in the carboxylate group but chelated firmly with the residual oxygen functions of the molecules. This provided a first insight into the high affinity for metal cations from Salinomycin through MS analysis, and also agrees with our results later discussed in the chapter.

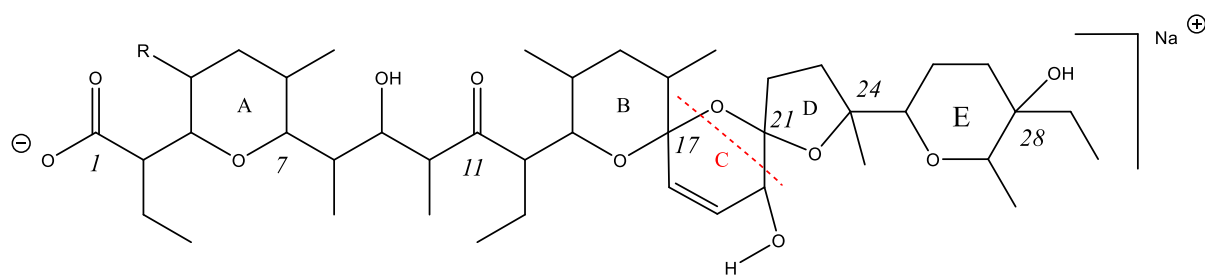


Figure 50 Characteristic cleave at ring C of salinomycin.

Later in 2003, another paper was published on the fragmentation study of salinomycin and monensin A using electrospray quadrupole time-of-flight mass spectrometry.²⁷⁴ It proposed a schematic representation of the salinomycin-sodium complex (Figure 51). However, our results showed that Na^+ binds more with the oxygens in the tricyclic spiroketal system, which was not shown in this reference. Salinomycin was observed to produce both sodiated $[(M-H+\text{Na})+H]^+$ and doubly sodiated molecular ions $[M-H+2\text{Na}]^+$. The cyclic structure of salinomycin is flexible enough to accommodate two sodium ions with one sodium on the acid end. Two types of fragmentations were observed. Product ions with

the carboxyl terminal are termed **Type A**, whereas product ions with the hydroxyl group terminal are termed **Type F**.²⁷⁵ This nomenclature is adopted in this research to study salinomycin and narasin. Product ions of sodiated and disodiated salinomycin complexes were reported, and schematic fragmentation pathways were published. The paper did not record any structural isomers of the product ions, there were no mechanisms of the fragmentation pathways, and no other metal cation complexes of salinomycin were investigated.

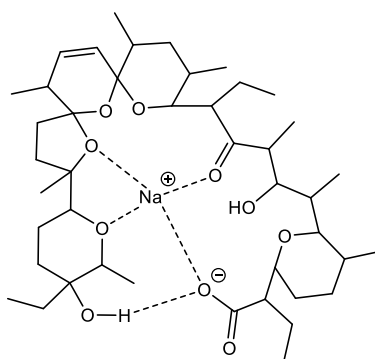


Figure 51 Structure of salinomycin-sodium complex.

The isomerisation of salinomycin was reported in 1998 and published in 1999.¹⁵¹ The structure of the salinomycin isomer is given in Figure 51. Salinomycin and narasin were reported to isomerise in water-methanol solution rapidly at ambient temperature (20-25 °C). After 11 days of storage time, salinomycin and narasin were completely isomerised. The MS data showed that the isomerised salinomycin and narasin retain their ionophore ability to chelate with metal cations. This paper established a foundation for investigating salinomycin and narasin isomers by ion mobility mass spectrometry. In addition, it is also interesting to see how chelation with different metal ions will impact the isomerisation.

4.1.2. Aims and Objectives

Salinomycin and narasin are selected in the final part of this research for their antibiotic and pharmaceutical properties and tendency to isomerise in solution. Their structures contain moieties similar to those studied previously with the α -amino acids and flavonoids. The aim is to adopt the proven methodologies developed during the analysis of the AAs and flavonoids to investigate the fragmentation of salinomycin and narasin. There is currently a lack of fragmentation studies on these polyether ionophores.

Previous studies in this research group demonstrated that chelating a polyether ionophore with alternative metal cations provokes changes in their fragmentation pathways²⁷⁶. Therefore, this part of the research also investigates salinomycin and narasin metal complexes and their effect on their structural conformation/isomerisation. The methodology is expanded to include ion mobility mass spectrometry as an additional tool in isomer differentiation and ionoform identifications. Results obtained from this research could provide a new approach to pharmaceutical drug developments and the study of other polyether ionophore antibiotics.

4.2. Experimental

4.2.1. Analytes

Salinomycin monosodium salt (SAL), from *Streptomyces albus*, $\geq 98\%$ and narasin (NAR), from *Streptomyces auriofaciens*, $\geq 98\%$ were purchased from Sigma-Aldrich (UK). Both samples were dissolved in 1:1 methanol:water sample solution to make up a standard solution of 10 mg/mL. These solutions were diluted to 1 mg/mL for analyses in the Orbitrap and the Synapt.

Metal chloride salts were dissolved in 1:1 methanol:water standard solution to make up a 10 mg/mL metal solution, and 200 μL of this solution was added to 800 μL of salinomycin solution to create a SAL+Metal mixture solution. Metal ions used for this research are: Li^+ , K^+ , Ca^{2+} , Mn^{2+} , Fe^{2+} , CO^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , Cs^+ , Ba^{2+} .

4.2.2. Instrumentation

Direct infusion positive ion mode ESI-MS and MS^n were performed on an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Hemel-Hempstead, UK). Analyte solutions were introduced into the instrument via a syringe pump at a 5 $\mu\text{L}/\text{minute}$ flow rate. The mass range was 100 – 1000 m/z for salinomycin monosodium salt and 100 – 1500 m/z for larger metal cation complexes. The scan rate was 2 scans/sec, and data were recorded for 30 seconds. The mass resolution was set to 240,000. CID-MS/MS using dry N_2 as the collision gas was performed on the isolated precursor ions and collision energy specified in the results. MS^n was also performed on the most intense product ions from the MS/MS spectrum.

Direct infusion nanospray MS/MS and IMS were performed on a Synapt G2-Si mass spectrometer (Waters, Manchester, UK) equipped with a Nanomate Triversa chip-based nanospray system (Advion Biosciences, Norwich, UK). The nanomate was set to aspirate 5 μL of sample solution to generate approximately 30 minutes of usable signal. IMS-MS/MS data of the relevant complex was obtained by isolating the precursor ion, subjecting it to IMS separation and then fragmenting post-IMS using CID-MS/MS in the transfer cell of the tri-wave ion mobility spectrometer. Collision energies were set from 30 to 50 eV, and mass spectra were collected individually along with drift time for approximately a minute per sample per collision energy used. The $[\text{SAL-H}+\text{Cu}^{\text{II}}]^+$ ion was also studied by MS/MS-IMS using pre-IMS fragmentation in the ion trap region of the tri-wave ion mobility spectrometer. In this experiment, the precursor ion was isolated and fragmented by CID MS/MS, with the product ions

entering the ion mobility chamber. The collision energy ranged from 20 – 30 eV. Mass spectra and IMS data were acquired for a minute per sample per collision energy. All data were analysed via DriftScope and exported to MassLynx. These two experiments allow the study of isomerism in two different ways. By generating product ions post-IMS, one uses the IMS cell to separate the isomers in the solution and obtain isomerically discrete MS/MS spectra. By generating product ions before IMS, you use the IMS cell to separate isomeric product ions and observe the alternative isomeric fragmentation routes separately.

4.2.3. Computational Studies

Gaussian 09¹²⁸ was used to perform all optimisation. Density functional theory (DFT) was used with exchange hybrid functional B3LYP. However, if B3LYP failed to achieve satisfactory results, mPW1PW91 was used as an alternative. This hybrid function is similar to Becke One-Parameter hybrid functional and uses the Perdew-Wang exchange as modified by the Adamo and Barone correlation.²⁷⁷ It was selected due to its success in previous literature.²⁷⁸ Due to salinomycin's larger size and complex structure, the basis set was set at a lower level for salinomycin to reduce computational costs. Based on computational costs and structure accuracy, the basis set applied to neutral salinomycin, its monosodiated molecular ion, and any product ions varied from 3-21 G to 6-311G. Gaussian 09 was later updated to Gaussian16¹²⁸ due to a system update on the advanced computer within the University. Theoretical CCS data of selective salinomycin structures and its product ions were calculated by IMoS. This was then used to compare with the experimental observation in drift time of IMS MS/MS spectra.

4.3. Results and Discussion

4.3.1. Overview

Spectra from CID MSⁿ of salinomycin and narasin monosodiated ions were investigated, and the product ions were tabulated in Table 15. Ion identities were labelled, and the route to their fragmentation pathway was also indicated. A nomenclature similar to the one in Chapter 3 was used here to clarify the product ions and the fragmentation pathways. As mentioned in the introduction, product ions from salinomycin and narasin in the positive ion mode typically result from one of two types of fragmentations. One retains the carboxyl group on the acid terminus intact and is identified as Type A fragmentation; the other retains the hydroxyl group on the alcohol terminus intact and is identified as Type F fragmentation. These two routes are combined into the updated nomenclature and written as ' $\text{M}\text{C}^{\text{X},\text{Y}}_{\text{A or F}}$ ' or ' $\text{M}_i\text{C}^{\text{X},\text{Y}}_{\text{A or F}}$ '. M means that it was initiated through a McLafferty rearrangement, predominately seen in the MSⁿ spectra of salinomycin and narasin. The lower case *i* represents the product ions were from the isomer of salinomycin or narasin. C^{X,Y} identifies which carbons (number X and Y), and where the C-C bonds were broken. Finally, the _{A or F} subscript signifies whether the product ion is produced by a type A or type F fragmentation.

Fomula	m/z	Ion Identity	Fragmentation Type	Pathway
C ₄₂ H ₇₀ O ₁₁ Na ⁺	773	[(SAL-H+Na)+H] ⁺	Monosodiated Ion	FULL SCAN
C ₄₂ H ₆₈ O ₁₀ Na ⁺	755	[773-H ₂ O] ⁺	A and F	MS/MS 773
C ₄₂ H ₆₆ O ₉ Na ⁺	737	[773-2H ₂ O] ⁺	A and F	MS/MS 773
C ₄₁ H ₇₀ O ₉ Na ⁺	729	[773-CO ₂] ⁺	F	MS/MS 773
C ₄₁ H ₆₈ O ₈ Na ⁺	711	[755-CO ₂] ⁺	F	MS/MS 773
C ₂₉ H ₄₈ O ₇ Na ⁺	531	[773-242] ⁺	M C ^{17,21} _A	MS/MS 773
	531	[773-242] ⁺	M C ^{9,11} _F	MS/MS 773
	531	[773-242] ⁺	Mi C ^{9,11} _F	MS/MS 773
C ₂₉ H ₄₆ O ₆ Na ⁺	513	[531-H ₂ O] ⁺	M C ^{9,11} _F - 18	MS/MS 773
	513	[531-H ₂ O] ⁺	Mi C ^{9,11} _F - 18	MS/MS 773
C ₂₉ H ₄₄ O ₅ Na ⁺	495	[531-2H ₂ O] ⁺	M C ^{9,11} _F - 36	MS ³ 773-755
	495	[531-2H ₂ O] ⁺	Mi C ^{9,11} _F - 36	MS ³ 773-755
C ₂₈ H ₄₈ O ₆ Na ⁺	503	[531-CO] ⁺	M C ^{17,21} _A - 28	MS ³ 773-531
C ₂₈ H ₄₈ O ₅ Na ⁺	487	[531-CO ₂] ⁺	M C ^{17,21} _A - 44	MS ³ 773-531
C ₂₃ H ₃₆ O ₆ Na ⁺	431	[531-100] ⁺	M C ^{9,11} _F - 100	MS/MS 773
	431	[773-342]	M C ^{11,13} _F	MS/MS 773
	431	[531-100] ⁺	Mi C ^{9,11} _F - 100	MS/MS 773
C ₂₃ H ₃₄ O ₅ Na ⁺	413	[431-H ₂ O] ⁺	M C ^{9,11} _F - 100 - 18	MS/MS 773
	413	[431-H ₂ O] ⁺	Mi C ^{9,11} _F - 100 - 18	MS/MS 773
C ₂₂ H ₃₆ O ₅ Na ⁺	403	[431-CO] ⁺	Mi C ^{9,11} _F - 100 - 28	MS ⁴ 773-531-431
C ₂₃ H ₃₂ O ₄ Na ⁺	395	[413-H ₂ O] ⁺	M C ^{9,11} _F - 100 - 36	MS ³ 773-413
	395	[413-H ₂ O] ⁺	Mi C ^{9,11} _F - 100 - 36	MS ³ 773-413
C ₂₀ H ₃₀ O ₅ Na ⁺	373	[513-140] ⁺	Mi C ^{9,11} _F - 18 - 140	MS ⁴ 773-531-513
C ₁₉ H ₃₄ O ₅ Na ⁺	365	[531-166] ⁺	M C ^{17,21} _A - 166	MS ³ 773-531
C ₁₉ H ₃₀ O ₅ Na ⁺	361	[431-70] ⁺	M C ^{9,11} _F - 100 - 70	MS ⁴ 773-531-431
C ₁₉ H ₃₂ O ₄ Na ⁺	347	[513-166] ⁺	M C ^{9,11} _F - 18 - 166	MS ⁴ 773-531-513
C ₁₉ H ₂₈ O ₄ Na ⁺	343	[413-70] ⁺	Mi C ^{9,11} _F - 100 - 18 - 70	MS ⁴ 773-755-513
C ₁₇ H ₂₆ O ₅ Na ⁺	333	[403-70] ⁺	Mi C ^{9,11} _F - 100 - 28 - 70	MS ⁴ 773-531-431
C ₁₆ H ₂₆ O ₃ Na ⁺	289	[487-198] ⁺	M C ^{17,21} _A - 44 - 198	MS ³ 773-531
C ₁₃ H ₂₂ O ₄ Na ⁺	265	[531-266] ⁺	M C ^{17,21} _A - 266	MS ³ 773-531
	265	[431-166] ⁺	M C ^{9,11} _F - 100 - 166	MS ³ 773-431
C ₁₃ H ₂₀ O ₃ Na ⁺	247	[265-H ₂ O] ⁺	M C ^{17,21} _A - 266 - 18	MS ⁴ 773-531-265
C ₁₃ H ₁₆ O ₃ Na ⁺	243	[343-100] ⁺	Mi C ^{9,11} _F - 100 - 18 - 70 - 100	MS ⁴ 773-755-413
C ₁₂ H ₂₂ O ₂ Na ⁺	221	[265-CO ₂] ⁺	M C ^{17,21} _A - 266 - 44	MS ⁴ 773-531-265
C ₁₀ H ₁₆ O ₃ Na ⁺	207	[265-58] ⁺	M C ^{17,21} _A - 266 - 58	MS ⁴ 773-531-265

Table 15 Table of [(SAL-H+Na)+H]⁺ precursor ion and identified product ions, their fragmentation pathways and ion identities from MSn spectra in the positive ion mode. The fragmentation pathway is colour coded based on the routes. *Peach*: Type A fragmentation leaving carboxyl end; *Blue*: Type F fragmentation leaving hydroxyl end and *Green*: Fragmentation from isomeric SAL.

4.3.2. Salinomycin Monosodiated Ion

Table 15 allows the determination of the fragmentation pathway for $[(\text{SAL-H+Na})+\text{H}]^+$. Due to the structure of SAL and its tendency to isomerise in solution, there are multiple isomeric product ions from the same precursor ion and product ions produced from the isomeric SAL (iSAL) with the same m/z . This further complicates the spectra and makes the structural elucidation more challenging. However, Table 15 summarised all identified product ions and their respective fragmentation pathways based on the route they appeared in the mass spectra.

The first trend to notice is the extensive loss of H_2O , CO and CO_2 in the overall spectra. This is expected based on the structure of SAL and its isomer, iSAL. In addition, type F fragmentation dominates the spectra of SAL and its isomer. Type A fragmentation produced more ions towards the lower mass range $m/z < 365$. Especially with m/z 265, all of the product ions of m/z 265 resulted from type A fragmentation. Moreover, the identifiable fragmentation pathway of SAL isomer is distinctively all type F. This could be explained by the structure from C1 to C13 of iSAL being identical to non-isomerised SAL. Fragmentations of this part of the molecule produce indistinctive type A product ions. Secondly, explained by the ring opening of rings B, C and D in the isomer structure offering a higher degree of freedom, it leads to entropically favoured fragmentation. These observations are useful in differentiating between the product ions of SAL and its isomer iSAL. These observations are also crucial in proposing a reliable fragmentation scheme for SAL and iSAL.

Figure 52 shows the MS/MS spectrum of SAL, precursor ion $[(\text{SAL-H+Na})+\text{H}]^+$ m/z 773. Some of the typical product ions to be investigated are m/z 531, 431 and 265. These product ions are essential in determining which SAL precursor ions they were generated from and via which type of fragmentation pathways. There are two consecutive losses of water from m/z 773, and losses of CO_2 are also present, although the intensities of water losses are much higher. Notice the 100 mass unit differences between peak m/z 531 and m/z 431 and m/z 513 and m/z 413. This is an iconic fragmentation pattern in studying the MS/MS spectra of salinomycin and will be useful in later sections when the SAL metal cation complexes are discussed.

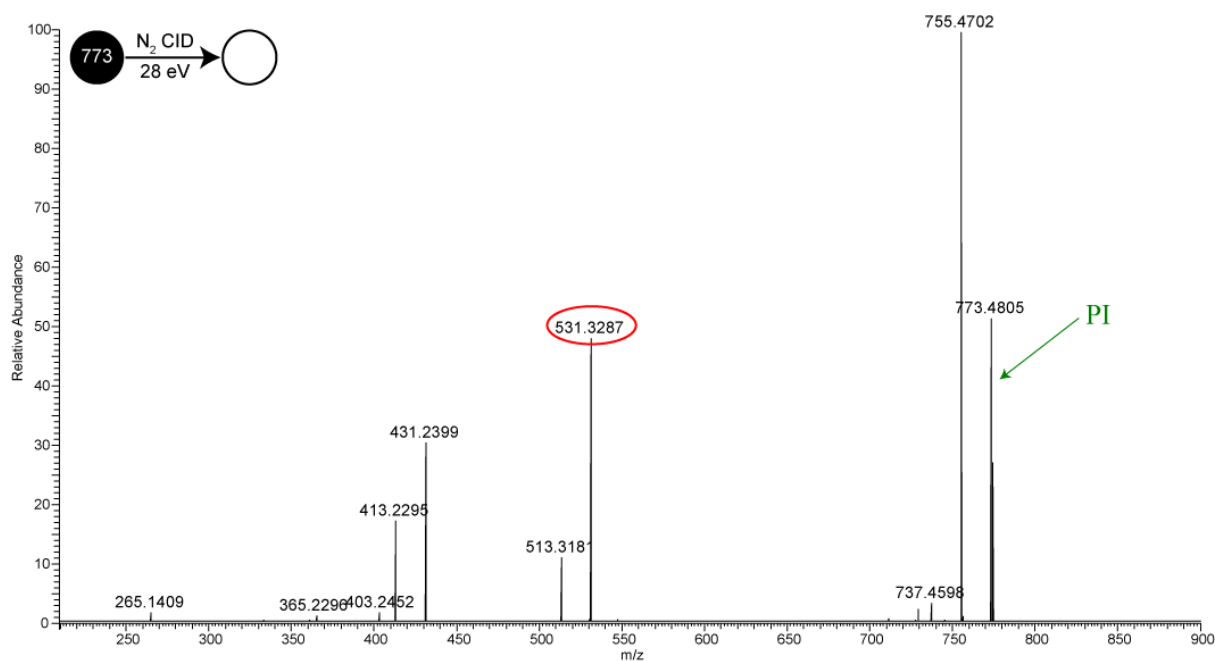
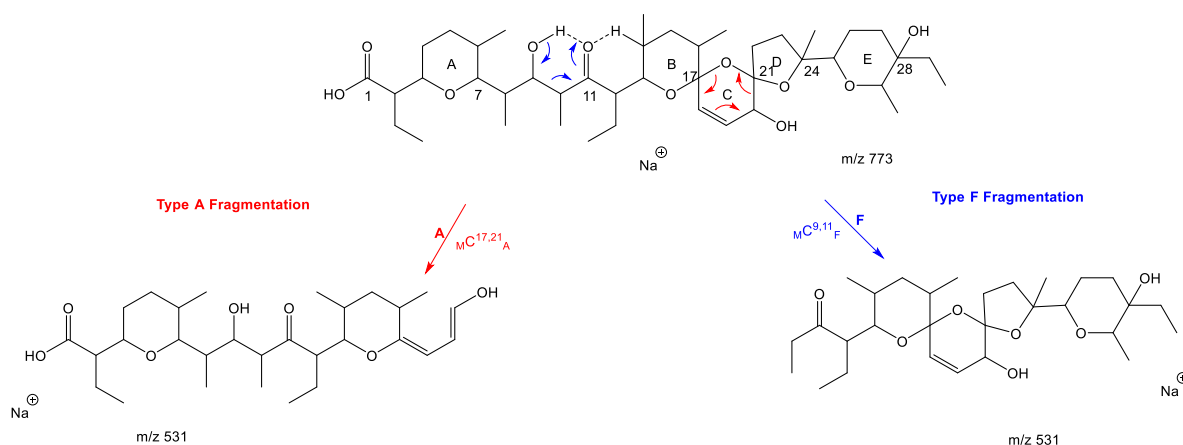


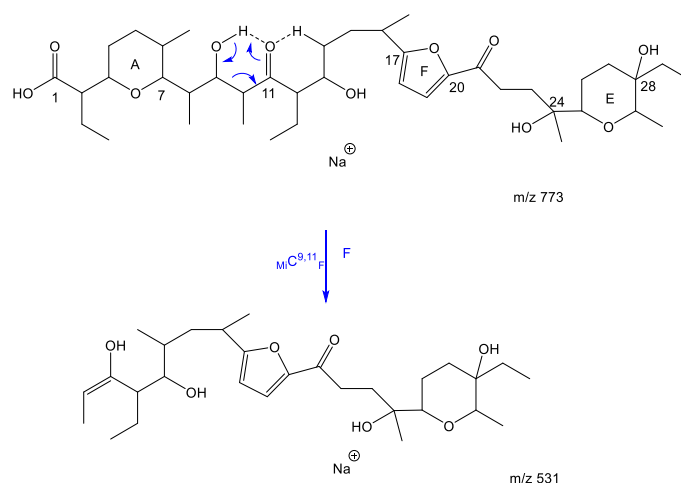
Figure 52 Positive ion ESI-CID-MS/MS spectrum of $[(\text{SAL-H}+\text{Na})+\text{H}]^+$ (precursor ion, PI, m/z 773).

The first proposed mechanism is for generating product ion m/z 531 from m/z 773. This ion was highlighted in red in Figure 52, and its fragmentation pathway was illustrated in Scheme 32. Product ion m/z 531 on the left-hand side of the schemes is the result of McLafferty rearrangement in ring C to generate a type A ion, which is marked in red. An alternative product ion m/z 531 can be produced by a similar McLafferty rearrangement but between Carbon 9 and 11 and with the carbonyl oxygen forming a cyclic structure by hydrogen bonding – this is a type F ion with the hydroxy terminus intact. This is coloured in blue on the right-hand side of the scheme. In the case of the iSAL, m/z 531 is solely produced by type F fragmentation. The reason for this is explained by the presence of the furan ring (ring F) in the iSAL structure. To obtain m/z 531 by type A fragmentation, this furan ring would have to be broken, which is extremely energetically unfavourable. All three m/z 531 product ions are structural isomers. This illustrates the fragmentation's complexity, which would be easily overlooked if only the MS/MS data were studied.

(a) [(SAL-H+Na)+H]⁺ *m/z* 531



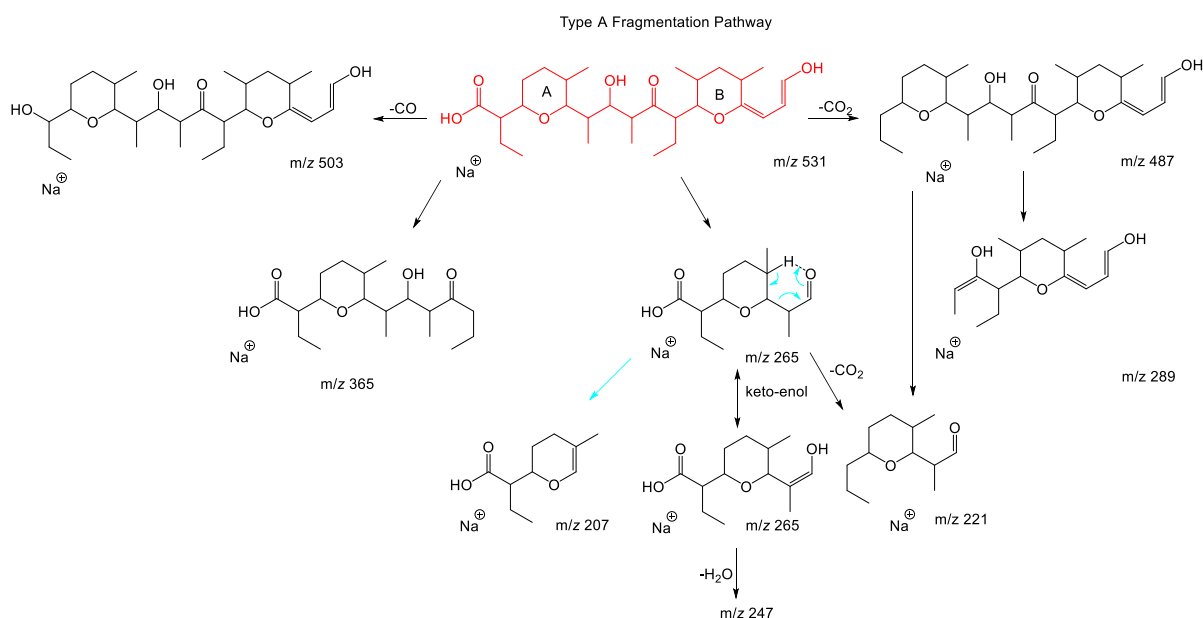
(b) [(iSAL-H+Na)+H]⁺ *m/z* 531



Scheme 32 Proposed mechanisms for generating the three isomeric product ions at *m/z* 531 of salinomycin sodiated ion and its isomer. Type A ions are marked in red, and type F ions are marked in blue.

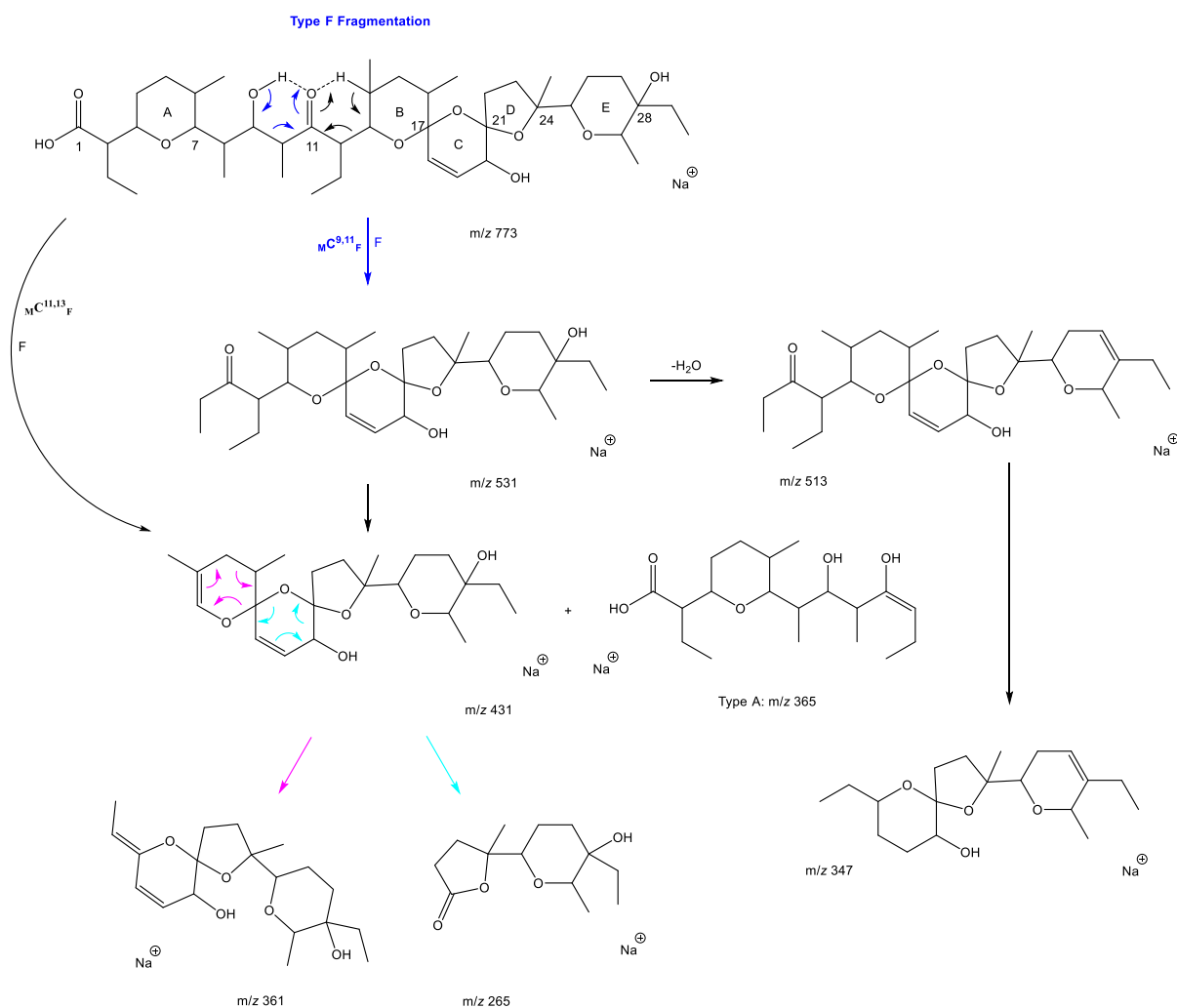
A detailed fragmentation pathway of the MSⁿ fragmentation of precursor ion *m/z* 531 is proposed in Scheme 32. Due to the multiple possible structures of this ion, discussions are spread over three schemes based on their types of fragmentation. Scheme 33 is for type A fragmentation, which includes only ions produced via the type A fragmentation from *m/z* 531. The first CO and CO₂ were expected with the carboxyl terminus in *m/z* 531. Product ion *m/z* 365 is a direct breakage at Carbon 12 and 13, leaving ring B propan-2-en-1-ol as a neutral loss. The interesting observation is ion *m/z* 265, which carried on producing sub-fragmentations *m/z* 247, *m/z* 221 and *m/z* 201. The chemical nature of *m/z* 265 allows it to alternate its structure from a keto to an enol by shifting the hydrogen onto the carboxyl oxygen, thus facilitating a loss of H₂O to produce ion *m/z* 247. The sub-fragmentations observed for product ion *m/z*

265 are a key tool in identifying which route it was generated from and in determining the structure of its precursor ion m/z 531. As m/z 265 could also be produced via type F fragmentation, it could not produce further sub-fragmentations due to its different structure, shown in the next scheme. This observation and further data provide evidence that MS^n is a powerful tool in differentiating structural isomers.



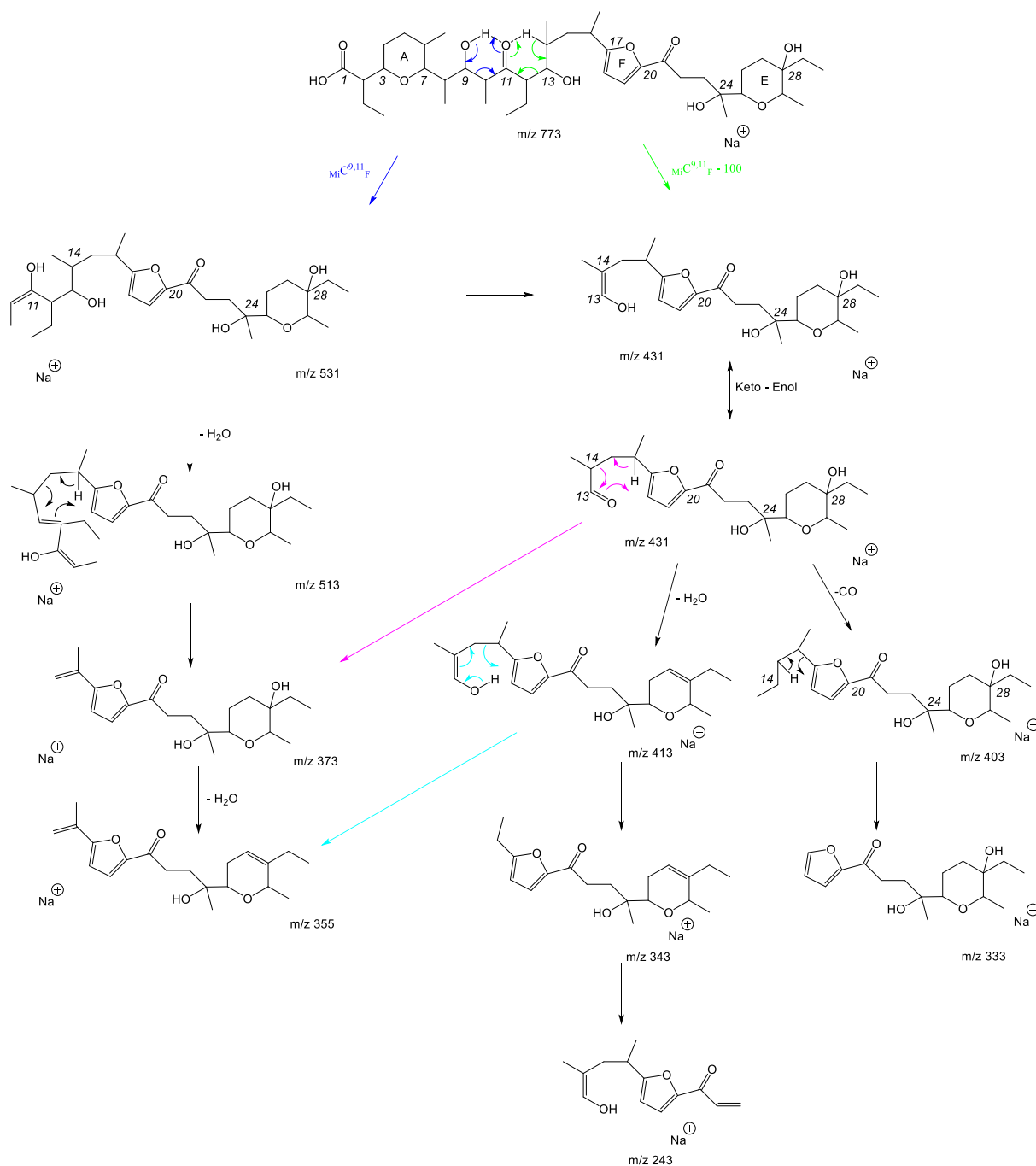
Scheme 33 Proposed mechanism for MS^n fragmentation of precursor ion m/z 531 following the type A fragmentation pathway.

Scheme 34 is the proposed mechanism for the MS^n fragmentation of m/z 531 following the type F pathway. The mechanism is marked in blue and black to represent that the product ions have the hydroxyl terminus. The initial bond breakage resulting in m/z 531 and m/z 431 are labelled with their nomenclature for clarity. Product ion m/z 531 underwent a McLafferty rearrangement, resulting in the breakage of C-C bonds and m/z 431 via pericyclic fragmentation. Product ion m/z 431 could also be from m/z 531 by losing the ketone. Multiple H_2O could be lost from m/z 531 and its sub-product ions. These are listed in Table 15 and are not included in Scheme 34 for clarity. Product ions m/z 361 and m/z 265 were also from m/z 531. These are differentiated by the colour of their proposed mechanisms. However, m/z 265 does not go on to have further losses, unlike what was observed for product ion m/z 265 in the type A fragmentation pathway. The cyclic structure of m/z 265 on the type F route is likely preventing its further fragmentation, i.e., the chelated sodium cation stabilises it, thus preventing further fragmentation.



Scheme 34 Proposed mechanism for the MS^n fragmentation of m/z 773 and m/z 531 following the type F fragmentation pathway.

Scheme 35 is the proposed mechanism for the MS^n fragmentation of iSAL following the type F pathway (see Table 15). This fragmentation pathway is in direct competition with the type A and F pathways of SAL. Other than the iconic m/z 531 and m/z 431 product ions, iSAL generates unique product ions through its open rings backbone, such as m/z 403 (loss of 18), m/z 395 (loss of 18), m/z 373, m/z 343, m/z 333 and m/z 243. These ions are only present in the MS^n of $[(iSAL-H+Na)+H]^+$. They are useful in determining the fragmentation pathway of the isomeric form of salinomycin. Due to the presence of multiple secondary and tertiary alcohols, water loss is common, and the position of each water loss defines which route the fragmentation is to happen next.



Scheme 35 Proposed mechanism for MS^n fragmentation of $[(i\text{SAL-H+Na})+\text{H}]^+$ in the positive ion mode.

4.3.3. Narasin Monosodiated Ion

The only difference between narasin and salinomycin is the extra methyl on C-4 of ring A. This would result in the spectra of $[(\text{Nar-H+Na})+\text{H}]^+$ having different peaks than that of SAL. There are three isomeric m/z 531 product ions for SAL, two from type A and F fragmentations from $[(\text{SAL-H+Na})+\text{H}]^+$. One from type F fragmentation from $[(i\text{SAL-H+Na})+\text{H}]^+$. However, in the case of NAR, the extra methyl group on ring A breaks the symmetry. Therefore, there could only be two m/z 531 generated

from type F fragmentations, one from NAR and one from isomeric narasin (iNAR). The original type A fragmentation in the case of SAL would be 14 mass units higher, generating m/z 545. These distinctive structural differences between SAL and NAR could mean quicker spectral interpretation for NAR and help elucidate the full pathway for SAL.

Fomula	m/z	Ion Identity	Fragmentation Type	Pathway
$C_{43}H_{72}O_{11}Na^+$	787	$[(NAR-H+Na)+H]^+$	monosodiated ion	Full scan
$C_{43}H_{70}O_{10}Na^+$	769	$[787-H_2O]^+$	A and F	MS/MS 787
$C_{43}H_{68}O_9Na^+$	751	$[787-2H_2O]^+$	A and F	MS/MS 787
$C_{42}H_{72}O_9Na^+$	743	$[787-CO_2]^+$	F	MS/MS 787
$C_{42}H_{70}O_8Na^+$	725	$[769-CO_2]^+$	F	MS/MS 787
$C_{30}H_{50}O_7Na^+$	545	$[787-242]^+$	$M C^{17,21}_A$	MS/MS 787
$C_{29}H_{48}O_7Na^+$	531	$[787-256]^+$	$M C^{9,11}_F$	MS/MS 787
	531	$[787-256]^+$	$M_i C^{9,11}_F$	MS/MS 787
$C_{30}H_{48}O_6Na^+$	527	$[545-H_2O]^+$	$M C^{17,21}_A - 18$	MS ³ 787-545
$C_{29}H_{46}O_6Na^+$	513	$[531-H_2O]^+$	$M C^{9,11}_F - 18$	MS/MS 787
	513	$[531-H_2O]^+$	$M_i C^{9,11}_F - 18$	MS/MS 787
$C_{29}H_{50}O_5Na^+$	501	$[545 - CO_2]^+$	$M C^{17,21}_A - 44$	MS/MS 787
$C_{30}H_{48}O_4Na^+$	495	$[531-2H_2O]^+$	$M C^{9,11}_F - 36$	MS ³ 787-751
	495	$[531-2H_2O]^+$	$M_i C^{9,11}_F - 36$	MS ³ 787-751
$C_{23}H_{36}O_6Na^+$	431	$[531-100]^+$	$M C^{9,11}_F - 100$	MS/MS 787
	431	$[787-356]$	$M C^{11,13}_F$	MS/MS 787
	431	$[531-100]^+$	$M_i C^{9,11}_F - 100$	MS/MS 787
$C_{23}H_{34}O_5Na^+$	413	$[431-H_2O]^+$	$M C^{9,11}_F - 100 - 18$	MS/MS 787
	413	$[431-H_2O]^+$	$M_i C^{9,11}_F - 100 - 18$	MS/MS 787
$C_{22}H_{36}O_5Na^+$	403	$[431 - CO]^+$	$M_i C^{9,11}_F - 100 - 28$	MS ⁴ 787-531-431
$C_{20}H_{36}O_5Na^+$	379	$[545-166]^+$	$M C^{12,13}_A$	MS ³ 787-545
$C_{20}H_{30}O_5Na^+$	373	$[513-140]^+$	$M_i C^{9,11}_F - 18 - 140$	MS ³ 787-513
$C_{20}H_{34}O_4Na^+$	361	$[431-70]^+$	$M C^{9,11}_F - 100-70$	MS ³ 773-431
	361	$[379 - H_2O]^+$	$M C^{12,13}_A - 18$	MS ³ 787-545
$C_{19}H_{36}O_3Na^+$	335	$[379 - CO_2]^+$	$M C^{12,13}_A - 44$	MS ⁴ 787-545-379
$C_{14}H_{24}O_4Na^+$	279	$[545-266]^+$	$M C^{9,10}_A$	MS/MS 787
$C_{13}H_{22}O_4Na^+$	265	$[431-166]^+$	$M C^{9,11}_F - 100 - 166$	MS ³ 787-431
$C_{13}H_{24}O_2Na^+$	235	$[279-CO_2]^+$	$M C^{9,10}_A - 44$	MS ³ 787-545
$C_{11}H_{18}O_3Na^+$	221	$[279-58]^+$	$M C^{7,8}_A$	MS ⁴ 787-545-279
$C_{10}H_{16}O_2Na^+$	191	$[235-44]^+$	$M C^{9,10}_A - 44 - 44$	MS ³ 787-279
$C_7H_{12}O_3Na^+$	167	$[279-112]^+$	$M C^{9,10}_A - 112$	MS ³ 787-279

Table 16 Table of $[(NAR-H+Na)+H]^+$ precursor ion and identified product ions, their fragmentation pathways, and ion identities from MSn spectra in the positive ion mode. The fragmentation pathway is colour-coded based on the routes. Peach: Type A fragmentation leaving carboxyl end; Blue: Type F fragmentation leaving hydroxyl end; and Green: Fragmentation from isomeric SAL.

The first thing to be seen from Table 16 is the shift of 14 mass for product ions generated via the Type A fragmentation pathway. The additional c-methyl on ring A increases all the product ions masses by 14 and can be seen throughout the MSⁿ spectra. As the rest of the molecule remains unchanged, the Type F fragmentation pathways remain identical to that seen for salinomycin. For NAR, there are only two isomers of *m/z* 531 and one at *m/z* 265. Both are generated from the type F fragmentation of NAR and iNAR. Since the mechanisms were proposed before for the type F fragmentation pathways and little change was observed, the next part focuses on the type A fragmentation pathway for NAR only.

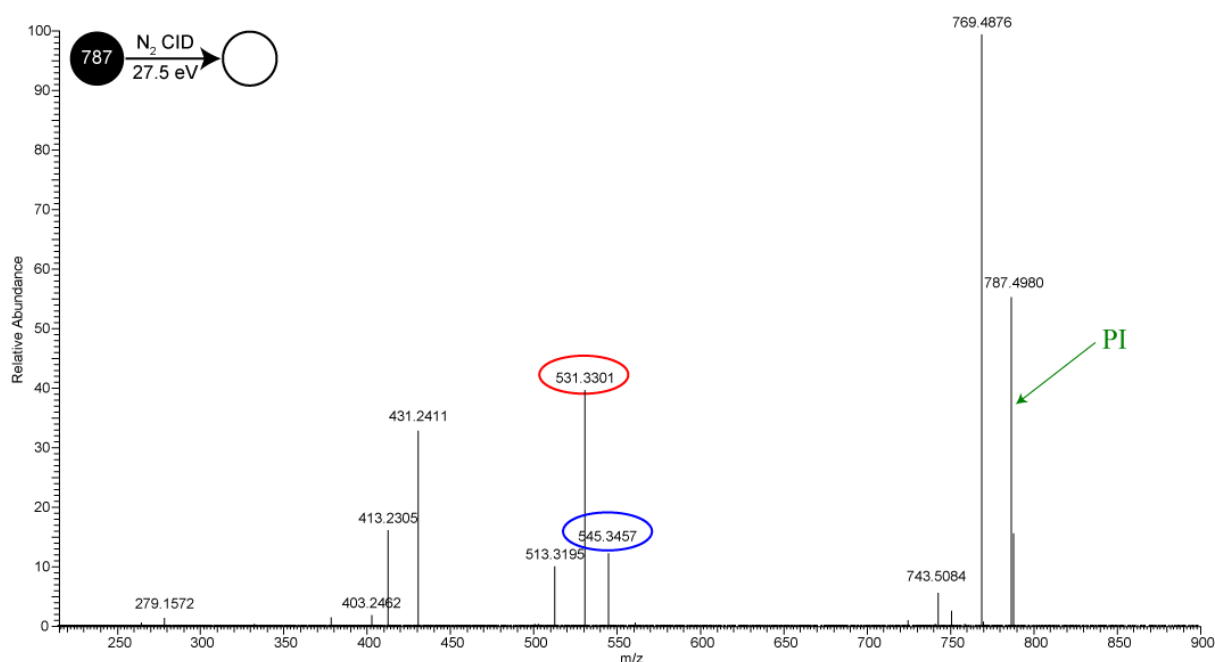
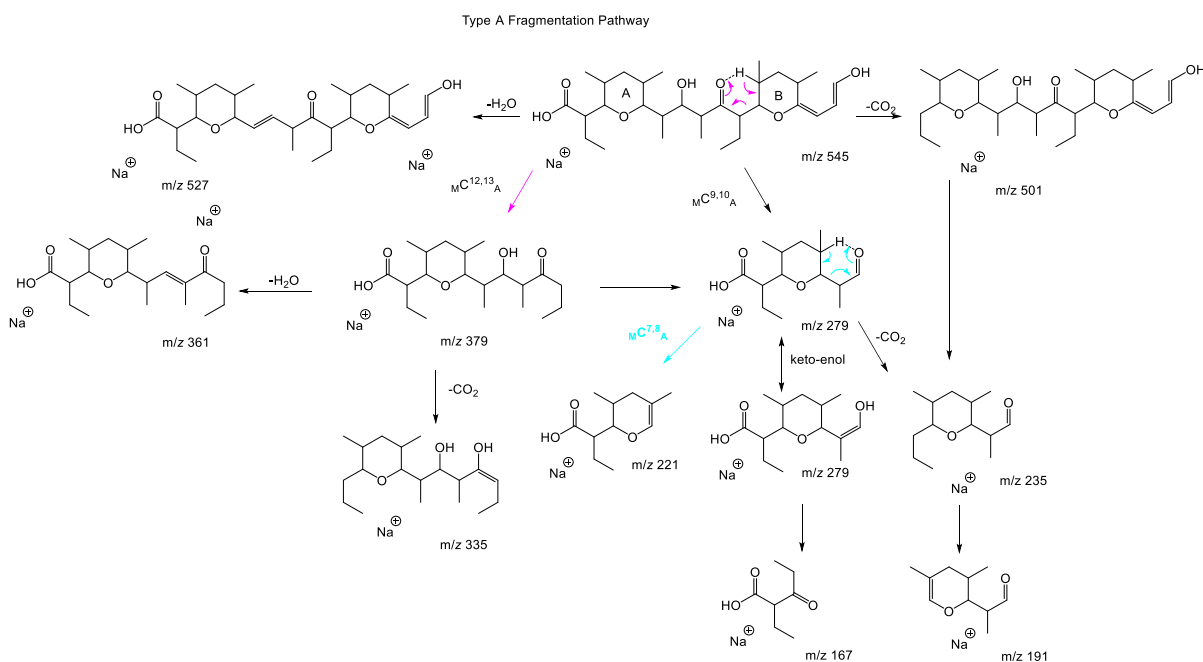


Figure 53 Positive ion ESI-CID-MS/MS spectrum of $[(\text{NAR-H}+\text{Na})+\text{H}]^+$ (precursor ion, PI, *m/z* 787).

Figure 53 is the positive ion MS/MS spectrum of $[(\text{NAR-H}+\text{Na})+\text{H}]^+$, precursor ion *m/z* 787. Compared to the equivalent spectrum for SAL, the differences to be noted are the addition of a peak at *m/z* 545 and the changes to the lower mass range. Product ion *m/z* 265 is less intense compared to *m/z* 279. This suggests that the type A fragmentation to generate *m/z* 279 is the preferred route. Product ions *m/z* 531, 513, 431, 413, and 403 remained in the spectrum, showing that the type F fragmentation pathway is unchanged for NAR and iNAR. Compared to the intensity of the two peaks at *m/z* 545 and *m/z* 513, type F fragmentation is more energetically favourable and produces more intense product ion peaks.



Scheme 36 Proposed mechanism for MS^n fragmentation of m/z 545 of $[(NAR-H+Na)+H]^+$ by the type A fragmentation pathway.

The proposed mechanism for m/z 545 of NAR is illustrated in Scheme 36. The fragmentation pathway was similar to that of SAL m/z 531 from type A but with all product ions 14 mass units higher. The first difference is that m/z 545 does not lose a CO like m/z 531 of SAL. Instead, m/z 545 loses CO_2 and H_2O . In addition, product ion m/z 379 also loses water or CO_2 and can go on sub-fragmentation to generate ion m/z 279. Product ion m/z 279 is intense enough to be investigated by MS^4 and generates more ions at the lower mass range. Overall, NAR produces similar fragmentation pathways compared to SAL. However, there are a few differences especially in its type A fragmentation route. The addition of a C-methyl group on ring A could provide more stabilisation by donating more electron density to ring A. Therefore making the resulting product ions more energetically stable structures.

4.3.4. Salinomycin Metal Complexes

Due to the previous success in using metal cations to generate novel fragmentation pathways,^{276,279} this research also investigated the effect of metal complexation on the fragmentation pathways observed for salinomycin. Among the 11 selected metal cations, the most distinctive results are from Cu^{II} . This required a more thorough investigation and will be discussed in more detail later. Overall, the ionophore properties of salinomycin enable it to chelate with all of the metal cations selected for this study, even with metal cations that are 1.5 times larger than Na^+ , such as Cs^+ . This was expected since salinomycin can generate doubly sodiated and doubly potassiated metal cations by the backbone oxygens and its acid terminus in its MS/MS studies. The high affinity for metal cations was already well established in the literature.²⁵⁹

Figure 54 demonstrates the metal ion binding affinities represented by their mass peak intensities of SAL. Overall, Cs^+ being the largest ion, has the lowest binding affinity when in competition with Na^+ and K^+ , which naturally exist in the original SAL solution. Salinomycin has a high affinity to bind two metal cations, in the case of Li^+ , K^+ , and Na^+ . The observation of their relative peak intensities supports this theory. Among the doubly charged metal cations, Ca^{2+} produces the most intense peak. This agrees with previous literature on the binding affinity of SAL with Ca^{2+} . It also agreed with our unpublished in-group research on lasalocid, another ionophore. All transition metals have similar relative peak intensities, which could be explained by the similarity in their cation sizes.

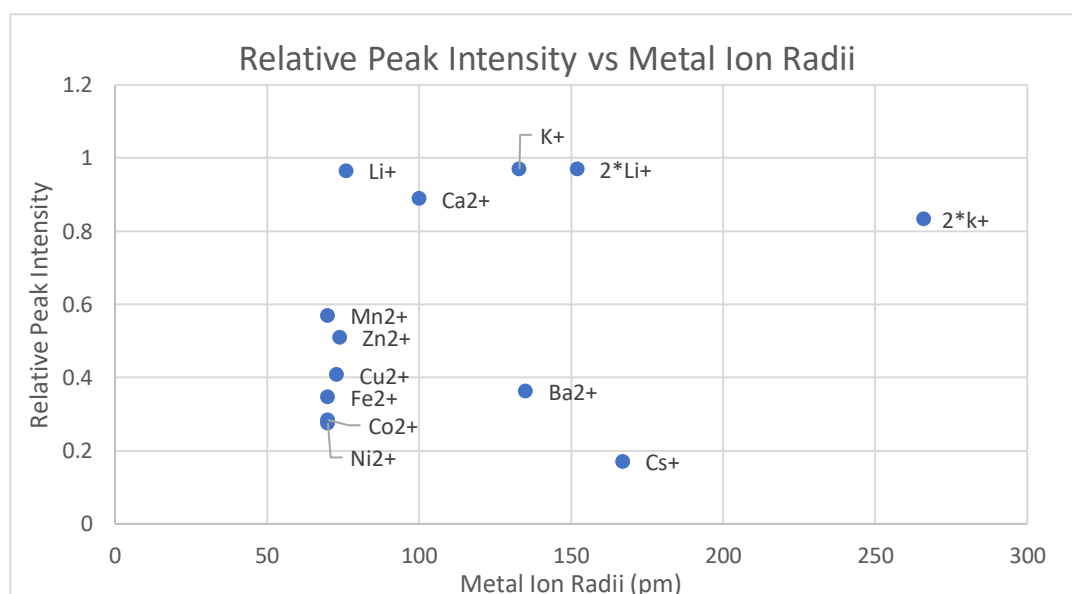


Figure 54 Relative peak intensity of $[\text{SAL}+\text{Metal}]^+ : [(\text{SAL}-\text{H}+\text{Na})+\text{H}]^+$ and metal cation radii. Each dot represents a metal cation. Peak intensities are normalised with the intensity of $[(\text{SAL}-\text{H}+\text{Na})+\text{H}]^+$. Disodiated and di-potassiated peaks are also included.

i. Lithium

Lithium is the smallest singly charged metal cation used in this experiment. Both single Li^+ and double Li^+ have great binding affinity to salinomycin and replace the sodium cation. The spectra of $[\text{SAL}+\text{Li}-\text{H}_2\text{O}]^+$ and $[\text{SAL}+2\text{Li}-\text{H}-\text{H}_2\text{O}]^+$ were investigated and compared in Figure 55. These spectra do not provide much additional information when compared to their sodiated alternatives. The consecutive losses of H_2O are readily seen, and the classic peaks of m/z 515 and m/z 415 with 100 mass unit differences are also present. However, after a single water loss, more product ions give additional insights into the fragmentation pathways of the single and double Li^+ complexes.

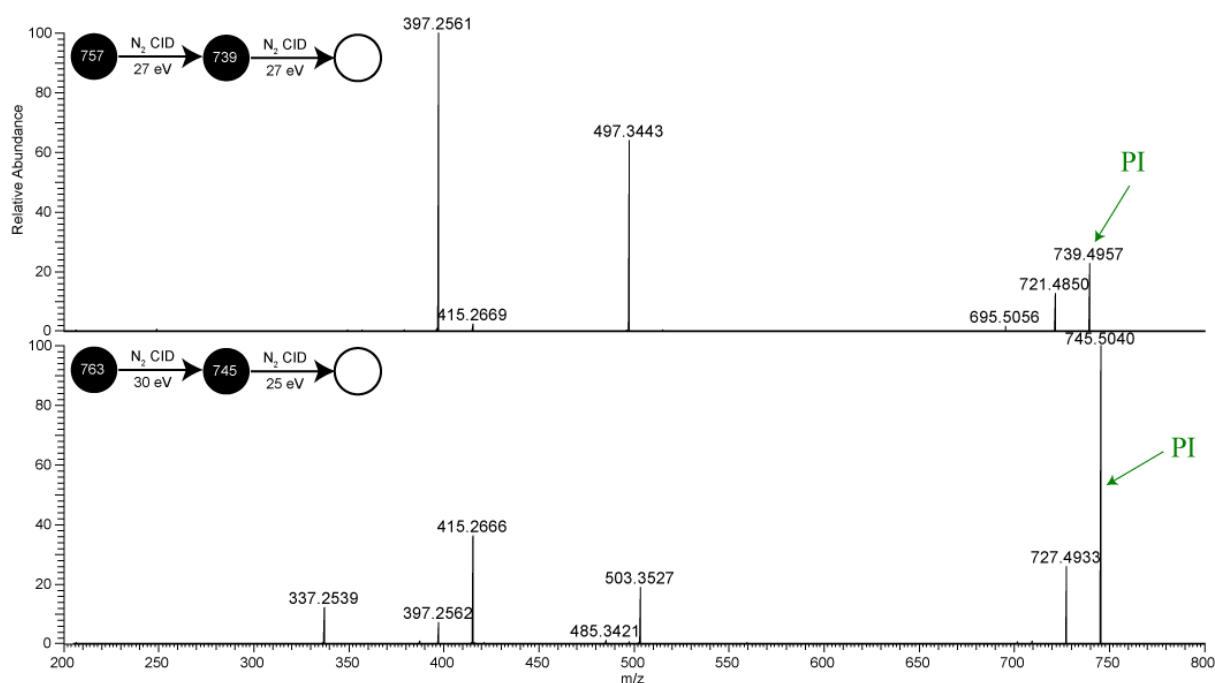
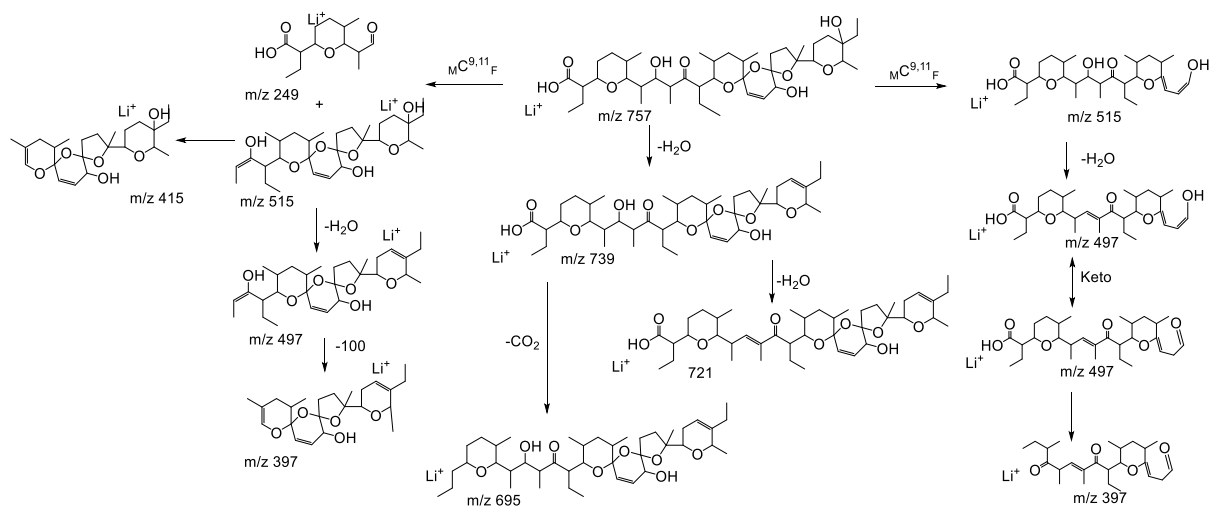
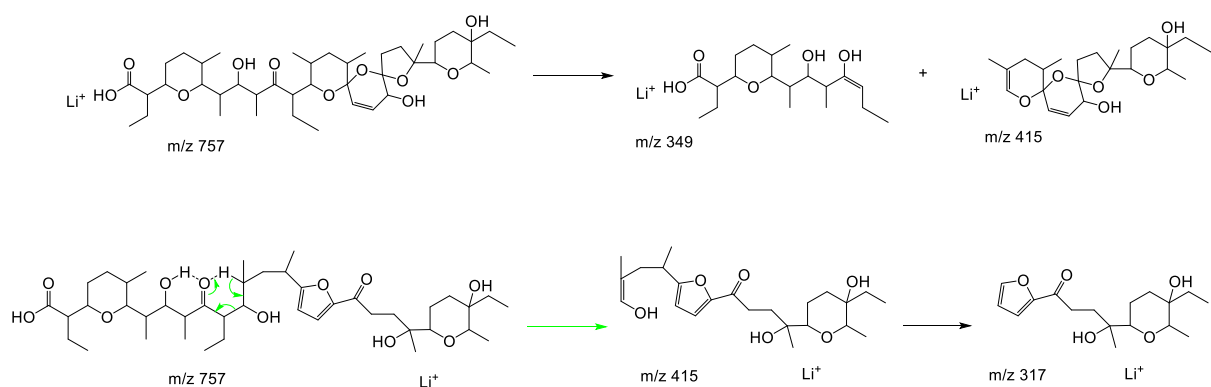


Figure 55 Positive ion MS³ spectra for (top) $[\text{SAL}+\text{Li}]^+$ (precursor ions 757 \rightarrow 739) and (bottom) $[\text{SAL}-\text{H}+2\text{Li}]^+$ (precursor ions 763 \rightarrow 745).

Proposed mechanisms were illustrated in Scheme 37 and Scheme 38 based on observations in the spectra of $[\text{SAL}+\text{Li}]^+$. The product ions were similar to $[(\text{SAL}-\text{H}+\text{Na})+\text{H}]^+$. Product ions resulting from the McLafferty rearrangement are labelled with their nomenclature along the carbons associated with the ring cleavages. The two types of fragmentations, A and F, are both observed. In addition, production ions of $[\text{iSAL}+\text{Li}]^+$ were also present in the spectrum.

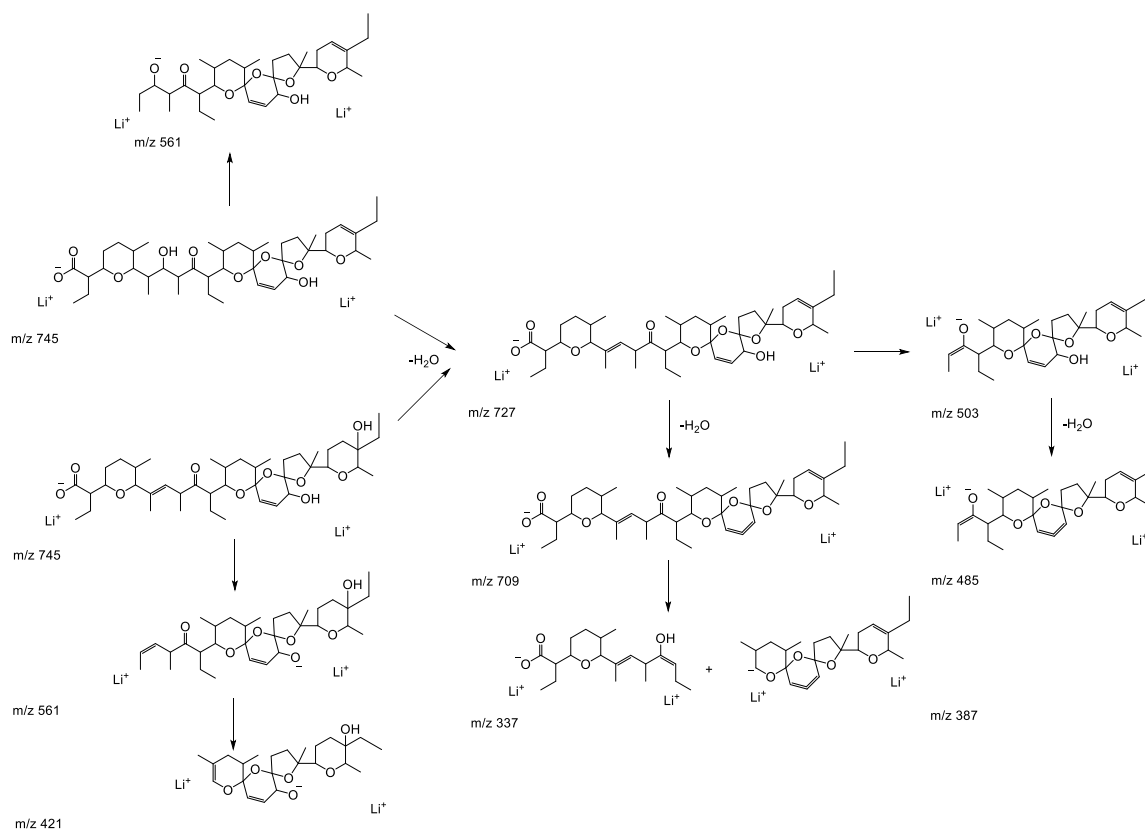


Scheme 37 Proposed mechanism for $[SAL+Li]^+$ fragmentation in positive ion mode.



Scheme 38 Additional proposed mechanism for generating product ion m/z 349 and m/z 317 of $[SAL+Li]^+$ in positive ion mode.

Spectra of $[SAL+2Li-H]^+$ and $[SAL+2Li-H-H_2O]^+$ demonstrated the ability of SAL to chelate double lithium cations. In the spectrum of $[SAL+2Li-H-H_2O]^+$, as shown in Figure 55, some of the product ion peaks remain the same as that of $[SAL+Li]^+$ which has a single Li^+ . However, peak m/z 745 and its consecutive water losses, peaks m/z 561, 503, 485, 421, 387, and 337, resulting from the fragmentation of the double lithium cations SAL complex. After the consecutive water losses, the lithium cations are expected to chelate with the oxygens on either end of SAL. The abundant number of oxygens in SAL facilitates this proposed chelation. Scheme 39 demonstrates the proposed fragmentation pathways for these product ion peaks. A McLafferty rearrangement is responsible for any ring cleavages observed with other SAL product ions and the production of ions m/z 421 and 337, for example.



Scheme 39 Proposed mechanisms for $[Sali+2Li-H-H_2O]^+$ in positive ion mode. Only product ions with double lithium metal cations are shown.

ii. Potassium

Similar to Li^+ and Na^+ , K^+ forms double cation complexation with SAL. The spectrum of $[SAL+K]^+$ is almost identical to that of $[(SAL-H+Na)+H]^+$. However, the spectrum and product ions of $[SAL+2K-H]^+$ offer variations in its fragmentation pathways. These are investigated and compared with that of $[SAL+2Li-H]^+$.

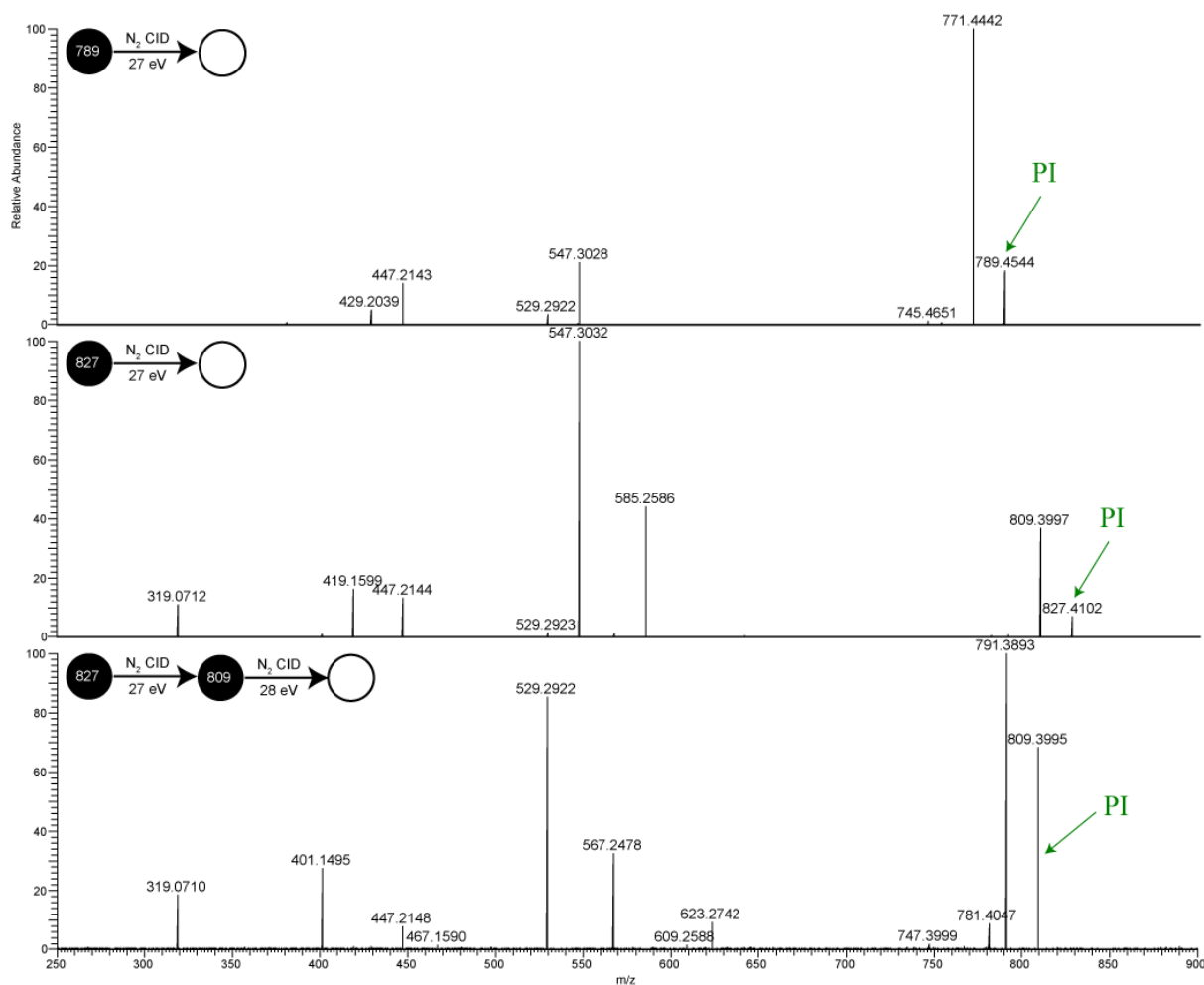
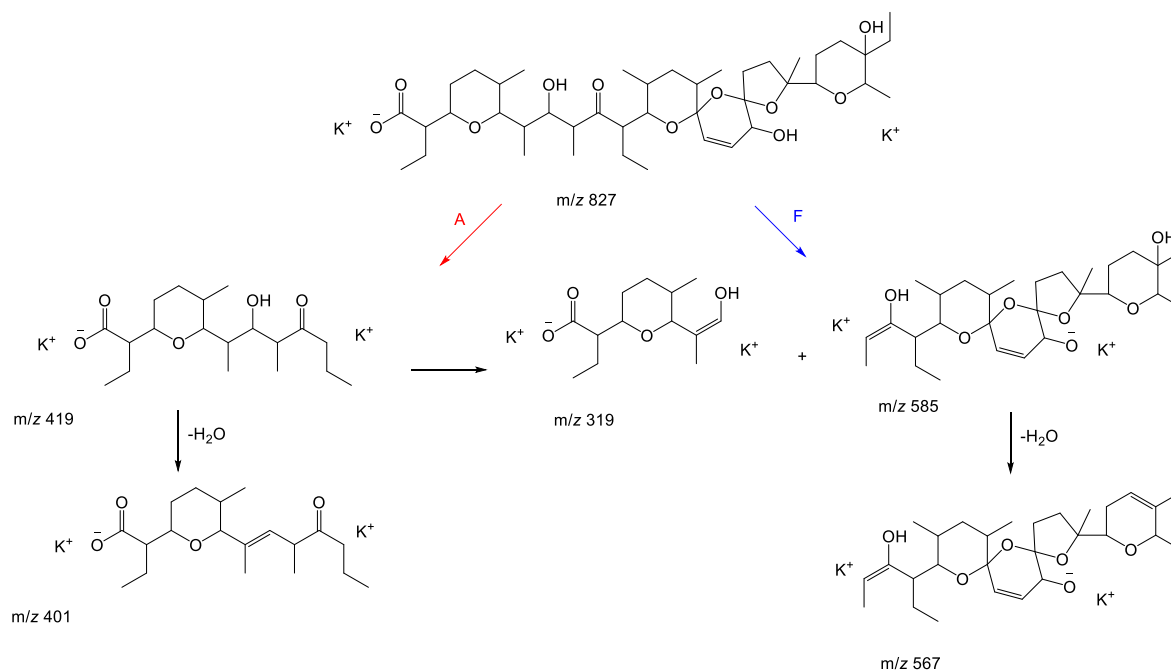


Figure 56 Positive ion MS^n spectra of salinomycin with potassium. MS/MS of $[SAL+K]^+$ (top) precursor ion m/z 789; MS/MS of $[SAL+2K-H]^+$ (middle) precursor ion m/z 827; and MS^3 of $[SAL+2K-H]^+$ (bottom) precursor ions 827 \rightarrow 809.

From Figure 56, the spectrum of MS/MS of $[SAL+K]^+$ looks nearly identical to the $[(SAL-H+Na)+H]^+$. However, the spectra of SAL's double K metal cation complex are different. The identifiable type A and F fragmentation product ions at m/z 547 and 447, m/z 529 and 429 are still present, along with the expected consecutive water losses. However, the product ion peak at m/z 585 is 38 mass units higher than m/z 547. This is also observed for peak m/z 567 and m/z 529. This represents two fragmentation pathways occurring in the same spectrum. One with a single K^+ adduct leaves as a neutral acid, and the other with a double K^+ . Since the proposed mechanisms for the original routes to form product ions $M\mathbf{C}^{17,21}_A$ and $M\mathbf{C}^{9,11}_F$ have been established, only the mechanism for the doubly potassiumated SAL complex product ions is discussed here.



Scheme 40 Proposed mechanism for the fragmentation of [SAL+2K-H]⁺ in positive ion mode.

Compared to Scheme 39, there are fewer product ions with double K⁺. This could be because potassium binds strongly with salinomycin, preventing further fragmentations.²⁸⁰ The fragmentation pathway is preferred with a single K⁺ adduct. However, type A and type F fragmentation pathways are still seen. The McLafferty rearrangement has not been shown to avoid repetitiveness. The lack of product ions with double potassium could be explained by the larger size and the increasing steric challenge for the oxygens to chelate with two potassium cations simultaneously.

iii. Iron

Among the transition metal cations, fragmentation pathways of the [SAL+Metal]⁺ typically follow the same fragmentation pathways of type A and F to [(SAL-H+Na)+H]⁺. However, there is one product ion peak that does not occur in the spectra of [(SAL-H+Na)+H]⁺ and any other cations. It is the product ion *m/z*. 705 from a loss of 100 mass units of [SAL+Fe^{II}-H]⁺. In Figure 57, this product ion peak is highlighted in red. Its product ion undergoes sub-fragmentation to generate product ions *m/z*. 687, 463 and 445 (see Scheme 41).

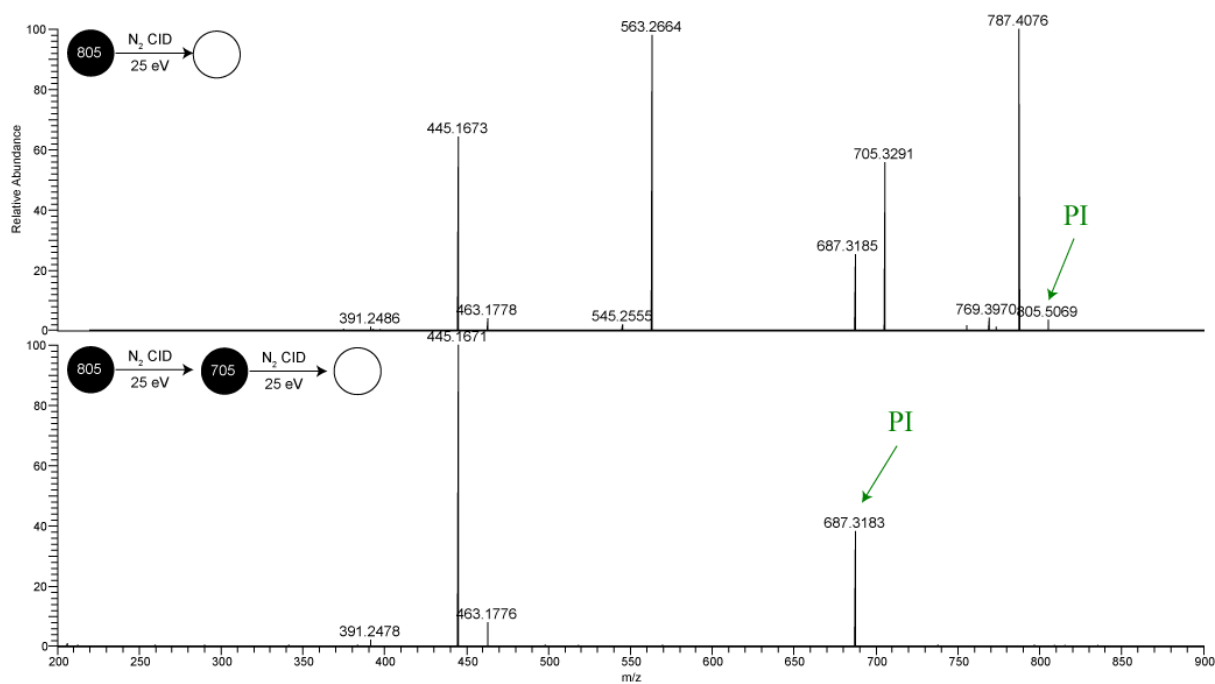
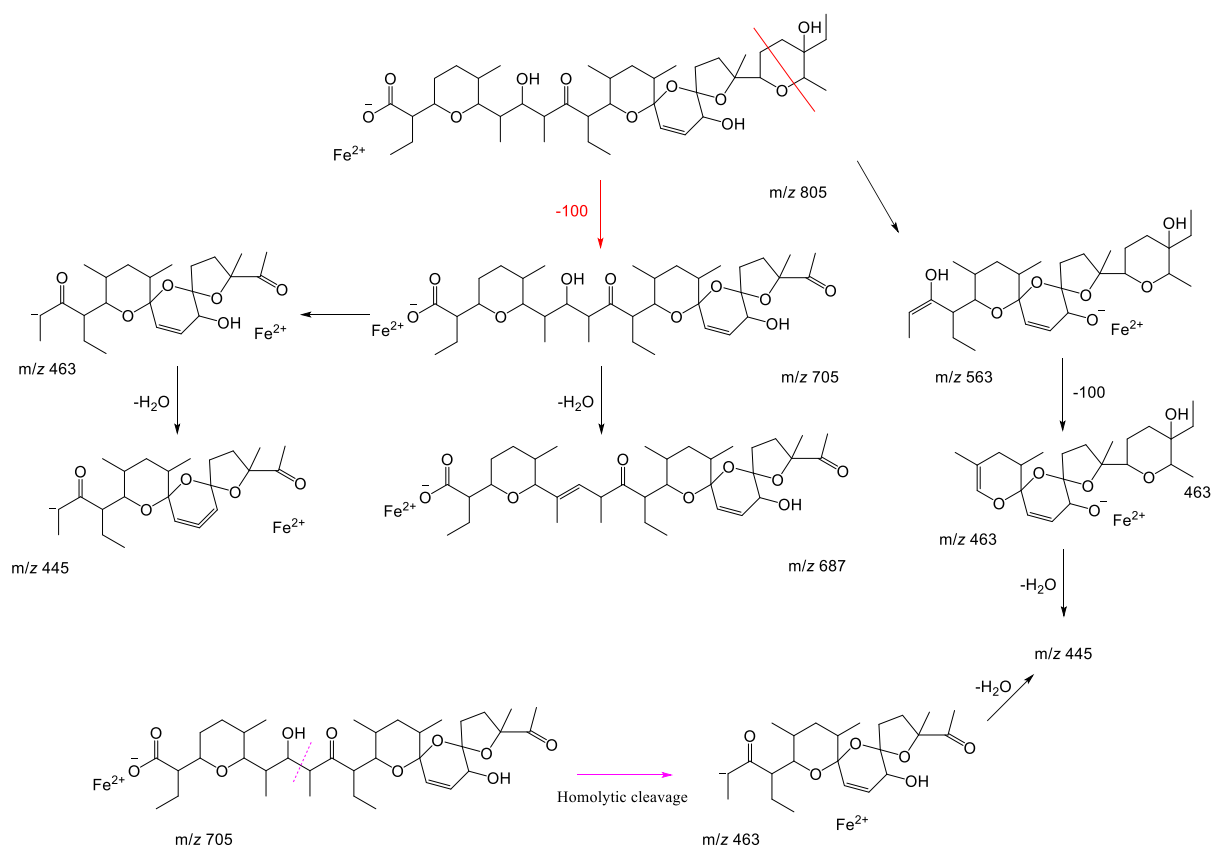


Figure 57 Positive ion MSⁿ spectra of salinomycin with iron, MS/MS of [SAL+Fe^{II}-H]⁺ (top) precursor ion m/z 805 and MS³ of [SAL+Fe^{II}-H]⁺ precursor ions 805→705 (bottom).

The fragmentation pathway is proposed in the scheme below. Product ion *m/z* 705 is generated from the breakage of ring E at carbon locations highlighted in red. This type of fragmentation has not been seen with other metal cations. In addition, this ion could also be produced by iSAL, which would be at the same carbon locations on ring E. This unique type of ion could be explained by Fe²⁺ interacting with the tetrahydropyran oxygen, weakening the bond between carbon 28 and the oxygen inducing a ring cleavage. Two fragmentation pathways were proposed for the generation of product ion *m/z* 463.



Scheme 41 Proposed mechanism for loss of 100 from product ion m/z 705 and its subsequent losses.

iv. Copper

Chelation of ionophores by copper (II) metal cations has been reported in the literature to generate novel fragmentation pathways in the gas phase due to its low redox potential and ability to act as a Lewis acid.²⁷⁶ In this section, the spectra of $[\text{SAL-H}+\text{Cu}^{\text{II}}]^+$ are investigated and discussed. Product ions observed in these spectra are unique and are not present for any other metal cation complexes of SAL. As previously reported in the literature, Cu^{2+} can reduce to Cu^+ in the gas phase when chelated and stabilised by (for example) an ionophore. This, in turn, can promote a new type of fragmentation pathway. Figure 58 is the positive ion MS/MS spectrum of $[\text{SAL-H}+\text{Cu}^{\text{II}}]^+$. The precursor ion is at m/z 812.5. The spectrum consisted of odd and even mass product ions, signifying the presence of two entirely different fragmentation pathways. For a molecule that does not contain a nitrogen atom, this must indicate chemistry involving hydrides and/or hydride abstraction. This has been previously published with ionophores^{247,276} and nucleosides, described as Radical Transfer Dissociation (RTD²⁷⁹). The classic type A and F fragmentations with product ions with peaks of 100 mass units apart were absent from this spectrum, meaning the McLafferty rearrangement is no longer observed.

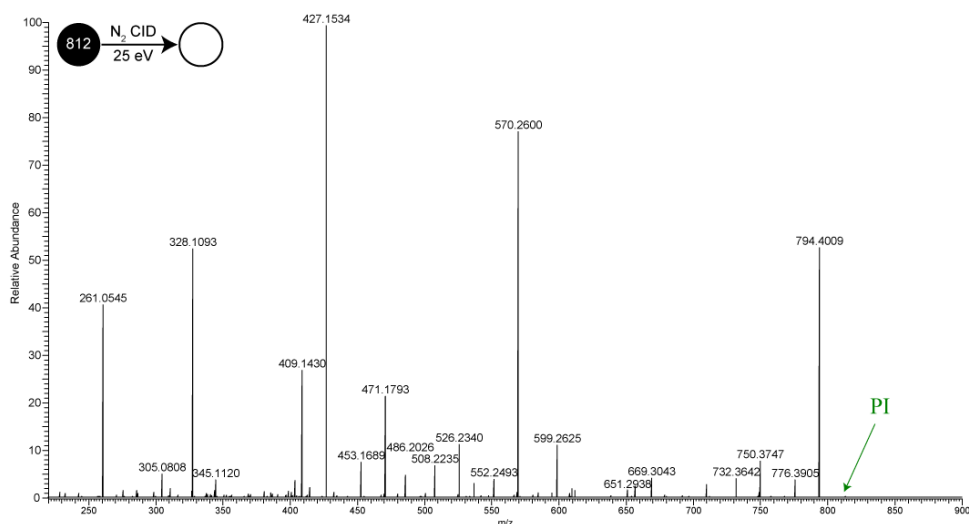


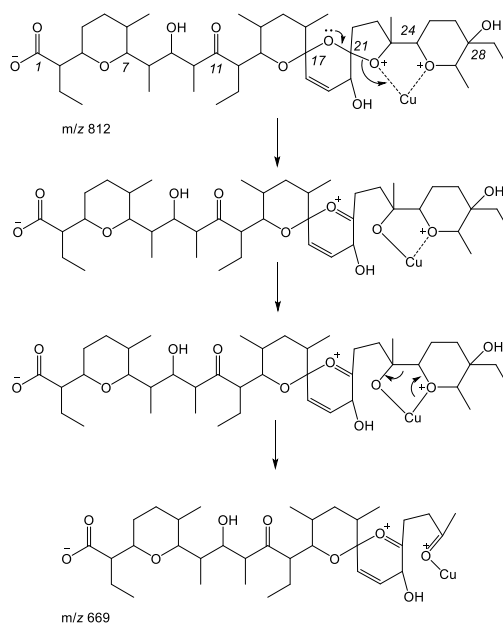
Figure 58 Positive ion MS/MS spectrum of $[SAL-H+Cu^{II}]^+$. The precursor ion is m/z 812, labelled as PI.

m/z	ID
812	$[SAL-H+Cu^{II}]^+$; PI
794	PI- H_2O
776	PI-2 H_2O
750	PI- H_2O - CO_2
732	750 - H_2O
669	$I C^{24,25}_A$
651	669- H_2O
599	$I C^{8,9}_F$
570	$II C^{9,11}_F$
552	570 - H_2O
526	570 - CO_2
508	526 - H_2O
486	508 - H_2O
471	599 - 128
453	471 - H_2O
427	599 - 172
409	427 - H_2O
404	$II C^{12,13}_A$
373	427-54
357	404 - 47
345	409 - 64
328	471-143
305	$I C^{9,10}_A$
261	305 - 44

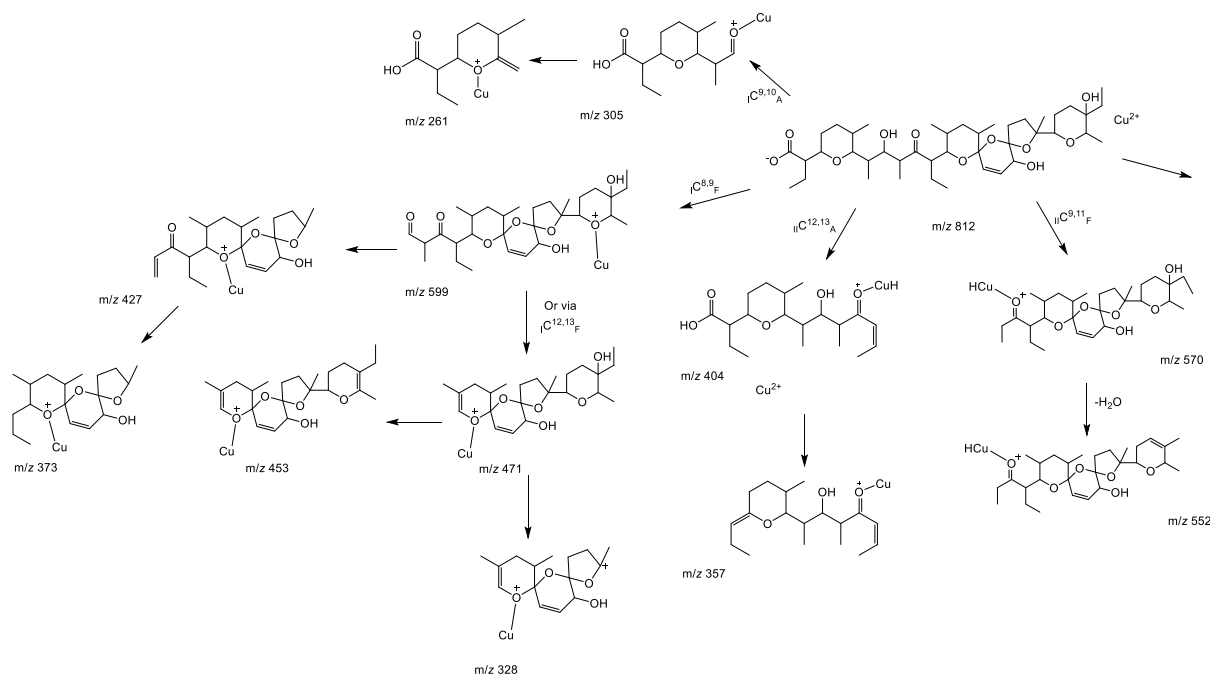
Table 17 Table of product ions observed in the positive ion MS/MS of $[SAL-H+Cu^{II}]^+$. Product ions are differentiated in odd masses (green) and even masses (yellow).

Table 17 is populated by the product ions observed for MS/MS of $[\text{SAL-H}+\text{Cu}^{\text{II}}]^+$. The product ions are coloured based on their masses. Their fragmentation pathway was also identified by using the same nomenclature previously. In the label ‘ $\text{I or II C}^{\text{x,y}}_{\text{A or F}}$ ’, I or II represents the oxidate state of Cu and whether the product ion was generated from a hydride abstraction from Cu^{II} or a redox reduction from Cu^{II} to Cu^{I} . The rest of the labels remain the same as before, and the number demonstrates the carbon location of the bond cleavage and its type of fragmentation.

Product ions with even masses were from a hydride abstraction by a Lewis acid mechanism. In contrast, product ions with odd masses were from a redox reaction to reduce Cu^{II} to Cu^{I} . This agrees with previous literature related to this study. One anomaly is ion m/z 328, which was produced from the fragmentation of odd-mass product ion m/z 471. Scheme 42 illustrates the proposed mechanism for the redox reaction to Cu^{I} . The oxygens on the pyran ring E facilitate this reduction. It was very intriguing to see these two different fragmentation pathways, which will be further investigated in the IMS-MS/MS.



Scheme 42 Proposed mechanism for generating product ion m/z 669 from a Cu^{II} reduction to Cu^{I} observed in the MS/MS of $[\text{SAL-H}+\text{Cu}^{\text{II}}]^+$.



Scheme 43 Proposed mechanism for product ions generated for the positive ion MS/MS of $[\text{Sal-H}+\text{Cu}^{\text{II}}]^+$. The even masses resulted from a hydride abstraction via a Lewis acid mechanism, and the odd masses were from the gas-phase redox reaction of $\text{Cu}(\text{II})$ to $\text{Cu}(\text{I})$.

4.3.5. IMS-MS/MS

Ion mobility studies on natural products have been increasing in popularity in recent research. In combination with MS and high-resolution MS/MS, IMS provides an additional tool for structure analysis and qualitative investigation of natural products.^{122, 281, 282} This section aims to provide IMS data on salinomycin and narasin due to their tendency to isomerise in solution. IMS data is useful in determining whether the isomers could be distinguished by their drift time and mass spectra.

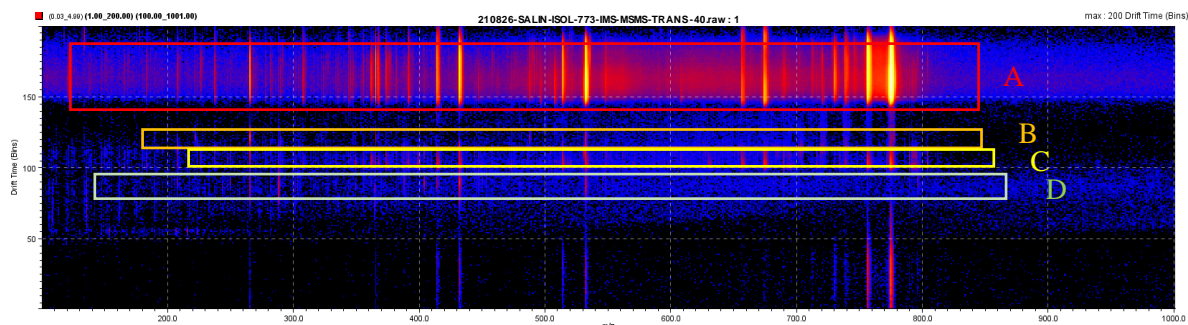
Ionoform – ‘different forms of ions’- is defined as metal cation complexes with the same mass and molecular formula. However, the various available chelating positions of the metal cation result in differences in their 3D complex structures. We hypothesise this is the case for salinomycin metal complexes. Salinomycin's unique high metal binding affinity and its ability to bind with two metal cations, Li^+ and K^+ , demonstrated the possibility of the metal cation being bound with salinomycin in slightly different positions. For example, the metal cation could be chelating in the middle of Salinomycin on rings B, C and D or towards either end of the molecule. Here we aim to use IMS-MS/MS data and computational methods to determine and distinguish the ionoforms of salinomycin monosodium cation. In addition, IMS data from narasin was also compared with SAL.

IMS-MS/MS data was captured for transfer energies ranging from 30eV to 55eV. This is when MS/MS is performed post-IMS to produce isomerically distinct MS/MS data. Figure 59 (a) shows the drift time against the mass for the IMS spectrum of m/z 773 (m/z 773 precursor mass at 40 eV collision energy in the transfer zone). The interesting observation is the number of $[(\text{SAL-H+Na})+\text{H}]^+$ peaks and the separation in drift time for product ion peaks m/z 531, m/z 431 and m/z 265. The separated $[(\text{SAL-H+Na})+\text{H}]^+$ are due to the isomerisation of SAL in solution and its ionoform, where the sodium cation can chelate to different parts of the molecule. Different rows of drift scope data were labelled and marked with letters to represent their converted mass spectra.

The separation of m/z 531 agrees with our previous observations in the tandem mass spectra. Multiple fragmentation routes can only be rationalised if SAL has multiple isomers or ionoforms in the gas phase. Based on the MS^n data and proposed mechanism of $[(\text{SAL-H+Na})+\text{H}]^+$, two m/z 531 product ions are generated from the original $[(\text{SAL-H+Na})+\text{H}]^+$ via type A and F fragmentations. Then iSAL can also generate a type F m/z 531. However, due to the structure of iSAL, type A m/z 531 was not observed in its MS^n spectra.

In addition, Figure 59 (b) also differentiates the iSAL spectrum from SAL by product ion peak m/z 403, which is unique to the isomer. Therefore, the order of the spectrum is arranged as top spectrum D: iSAL and is demonstrated by its unique product ion peak at m/z 403. Second spectrum C: $[(\text{SAL-H+Na})+\text{H}]^+$ type F, sodium will most likely chelate with salinomycin neutral molecule at the hydroxy ring end. Third spectrum B: $[(\text{SAL-H+Na})+\text{H}]^+$ type A, sodium is also likely to chelate at the ring end but could be closer to the carboxyl group at the A terminus. Bottom spectrum A: which belongs to the $[(\text{SAL-H+Na})+\text{H}]^+$.

(a) Salinomycin IMS-MS/MS (m/z 773 precursor ion) vs Drift Scope Drift Time



(b) Salinomycin IMS-MS/MS mass spectra (m/z 773 precursor ion) at Transfer Collision Energy 40 eV.

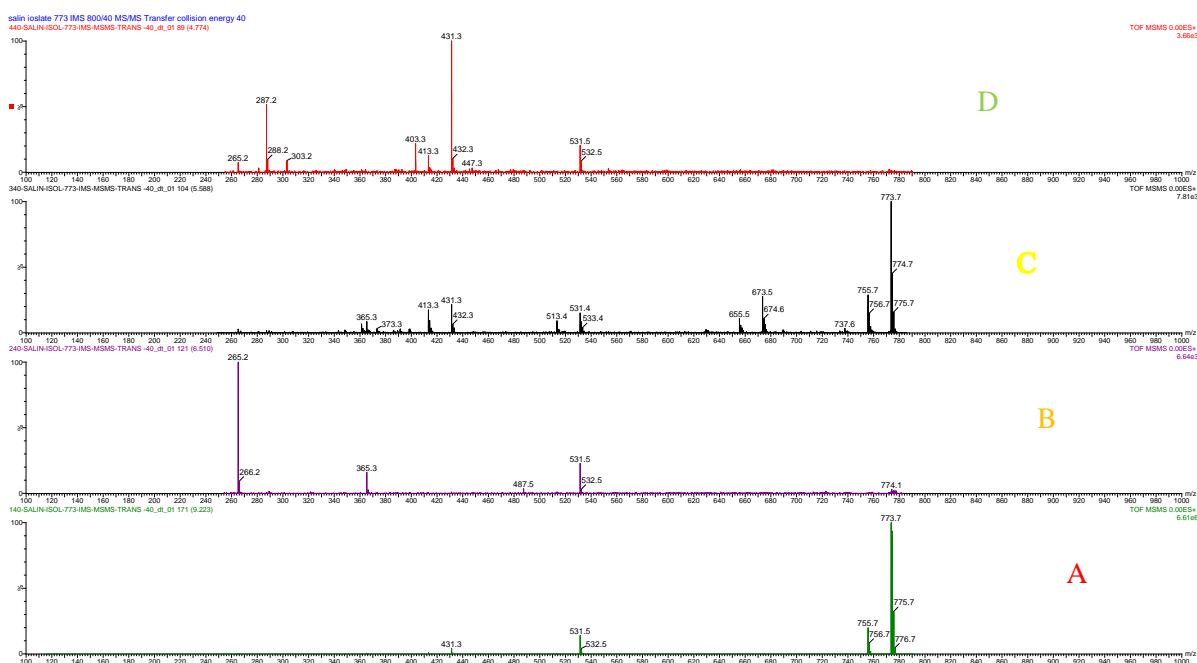


Figure 59 IMS-MS/MS data for $[(\text{SAL-H+Na})+\text{H}]^+$ at transfer energy 40 eV in positive ion mode. (a) Drift scope plots of mass vis drift time and (b) Mobility separated tandem mass spectra of m/z 773.

Theoretical CCS values were calculated on computational models of $[(\text{SAL-H+Na})+\text{H}]^+$ and $[(\text{iSAL-H+Na})+\text{H}]^+$ to aid a more detailed interpretation of the IMS data as well as the three types of product ion m/z 531, type A and F from original SAL, and type F from iSAL. $[(\text{SAL-H+Na})+\text{H}]^+$ ionoforms were modelled as sodium in the middle and sodium at the ring end. The optimised structures of the two ionoforms of $[(\text{SAL-H+Na})+\text{H}]^+$ are shown in Figure 60.

From the optimised structure (a), the sodium ion is chelating to the oxygens on rings B, C, and D, and it also looks to interact with the carboxyl oxygen at the A terminus. In structure (b), the sodium ion is at the ring end and chelates with the pyran oxygen and the OH group at the F terminus.

(a) [(SAL-H+Na)+H]⁺ sodium in the middle (b) [(SAL-H+Na)+H]⁺ sodium at the ring end

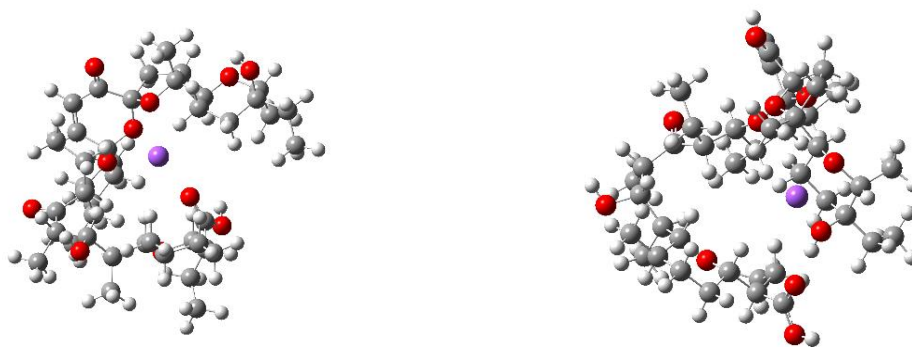


Figure 60 Gaussian modelling of optimised [(SAL-H+Na)+H]⁺ 3D structures. Grey: Carbon; White: Hydrogen; Red: Oxygen; and Purple: Sodium.

Name	Position of the Metal Cation	Calculated CCS (Å ²)
[(SAL-H+Na)+H] ⁺	Middle	283.58
[(SAL-H+Na)+H] ⁺	Ring End	277.28
[(iSAL-H+Na)+H] ⁺	Middle	285.54
531 Na iF	-	261.25
531 Na A	-	241.59
531 Na F	-	225.07

Table 18 Calculated theoretical CCS data for [(SAL-H+Na)+H]⁺, its isomer and ionoform, and three product ion m/z 531 types.

Table 18 shows the calculated CCS values for salinomycin, and this data is purely theoretical and used for comparison with the experiment observation solely. Combining these results with Figure 61 established a form of agreement. Ring end sodium [(SAL-H+Na)+H]⁺ has a lower CCS and, therefore, is expected to have a lower drift time. The separation in the drift time figure does not feature the [(iSAL-H+Na)+H]⁺ as a precursor ion. However, the similarities between the CCS values for SAL and iSAL monosodium ion could mean no experimental separation. This supports our IMS-MS/MS data for no observation of the [(iSAL-H+Na)+H]⁺ molecular ion. The relative peak intensities of selected product

ions m/z 531, 403, and 265 are compared in Figure 61, alongside the precursor ion m/z 773 from different ionoforms. It can be observed that as the collision energy increases, the majority of the ions lose intensity. Product ion peaks from the iSAL have the highest peak intensities at 45 eV. Type A and F fragmentation pathways compete. Type F m/z 531 is more intense and increases peak intensity, whereas type A m/z 531 is the opposite.

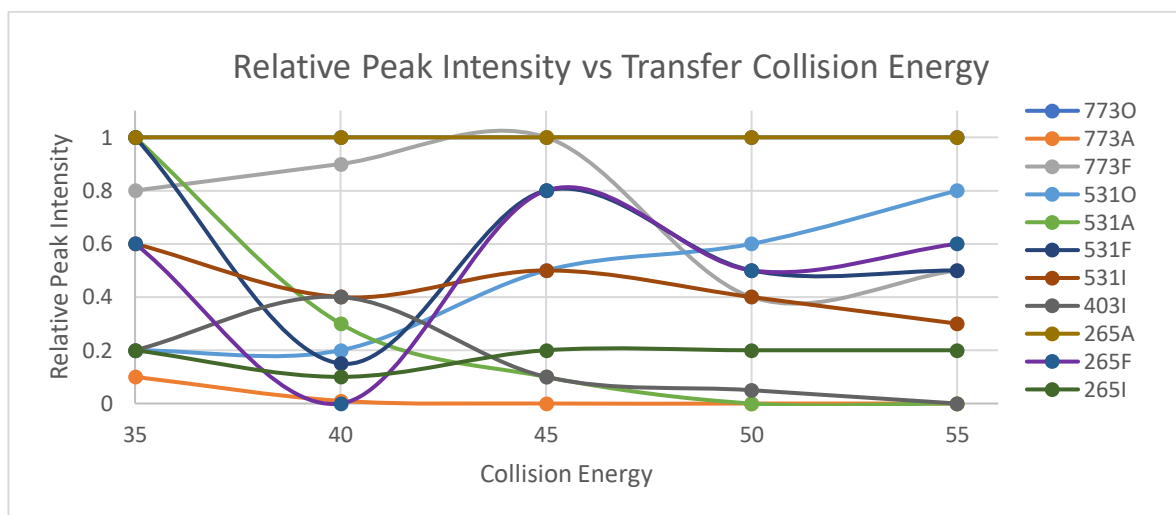
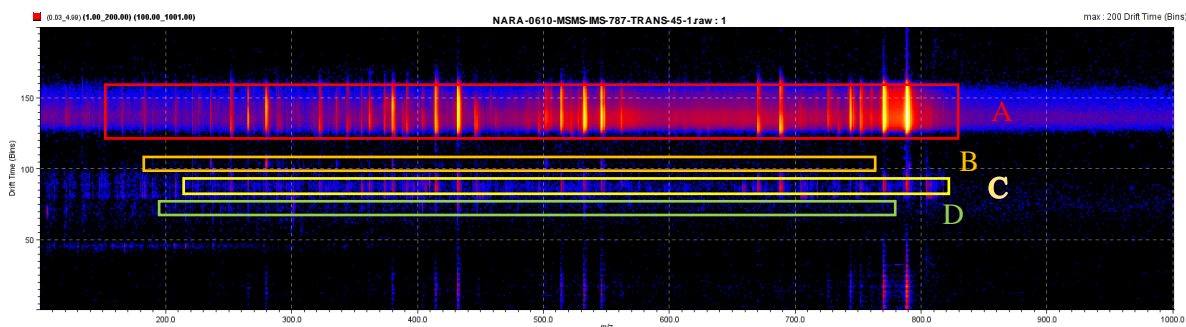


Figure 61 Relative peak intensities vs. collision energy obtained from the IMS-MS/MS spectra. The peaks selected are labelled in colours. O: original SAL; A: A type fragmentation; F: F type fragmentation and I: isomeric SAL.

Compared to IMS-MS/MS data for SAL, NAR has fewer molecular ions and product ions separations in its IMS spectrum. This is expected due to the additional C-methyl group on ring A. There are only two product ions at m/z 531, both from the type A fragmentation of the original NAR and isomeric NAR. In addition, two types of m/z 545 demonstrate the two type A fragmentations from NAR and iNAR. The observations agree with our results in the earlier section.

(a) Narasin IMS-MS/MS (m/z 787 precursor ion) vs Drift Scope Drift Time



(b) Narasin IMS-MS/MS mass spectra (m/z 787 precursor ion) at Transfer Collision Energy 45 eV. Spectra arranged in reverse order of the IMS drift scope data.

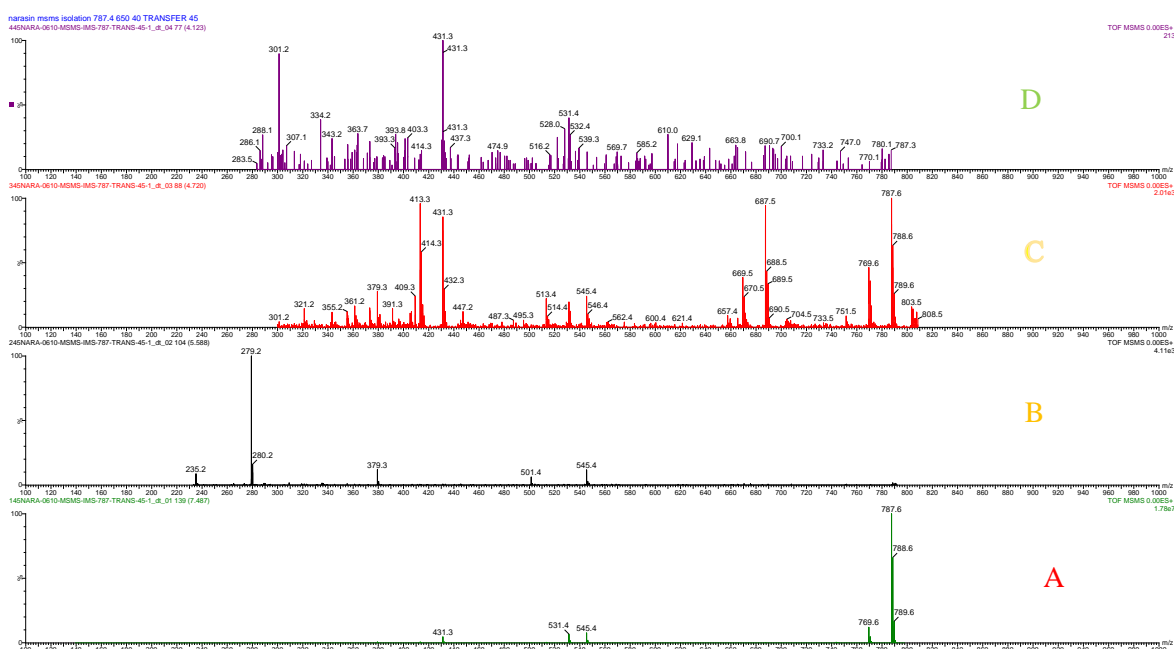
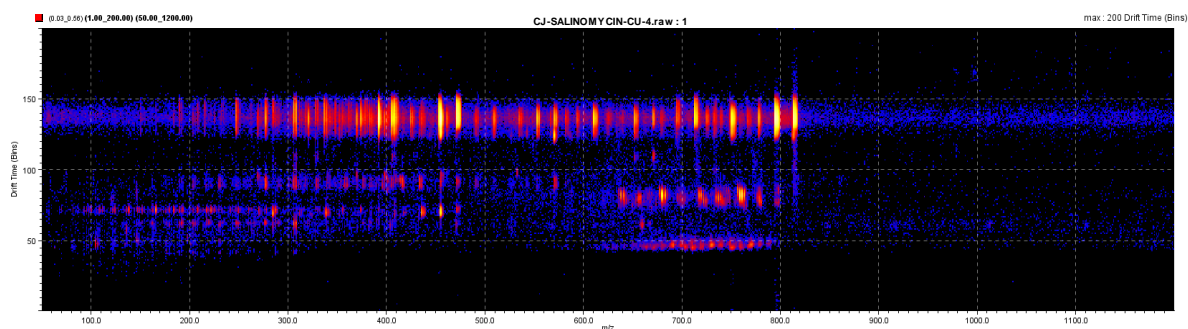


Figure 62 IMS-MS/MS drift time drift scope data and MS/MS spectra for $[(\text{NAR-H}+\text{Na})+\text{H}]^+$ at transfer collision energy 45eV in positive ion mode.

IMS-MS/MS data for $[\text{SAL-H}+\text{Cu}^{\text{II}}]^+$ is shown in Figure 63 below. The Gaussian optimisation for the SAL+Cu complex is also provided in Figure 64. The Cu^{II} cation is likely to chelate to rings B, C, D and E. This is supported by the experimental observation in its spectra that the product ions were predominately generated from Cu-O bonding. Unlike what was seen in the IMS-MS/MS data of $[(\text{SAL-H}+\text{Na})+\text{H}]^+$, there is no separation of the precursor Cu adduct ion m/z 812. This could be explained by the fact that Cu^{II} has a higher affinity for oxygen and prefers to chelate in a more specific way than a sodium cation. $[\text{SAL-H}+\text{Cu}^{\text{II}}]^+$ has more separations in its lower mass range. By investigating the

MS/MS data below, most of the separated product ions are odd masses. This suggests that the redox reaction of Cu^{II} reducing to Cu^{I} is more energetically favourable. Product ion m/z 305 has a distinctive separation in its IMS data. This could result from the iSAL. Due to the complexity of this data, additional computational modelling is required. However, the current costs for optimising a structure such as SAL+Cu are very high. Therefore, these could be the potential for interesting future work.

(a) Salinomycin Cu Complex IMS-MS/MS (m/z 812 precursor ion) vs Drift Scope Drift Time



(b) Salinomycin Cu Complex IMS-MS/MS mass spectra (m/z 812 precursor ion) at Transfer Collision Energy 45 eV. Spectra arranged in reverse order of the IMS drift scope data.

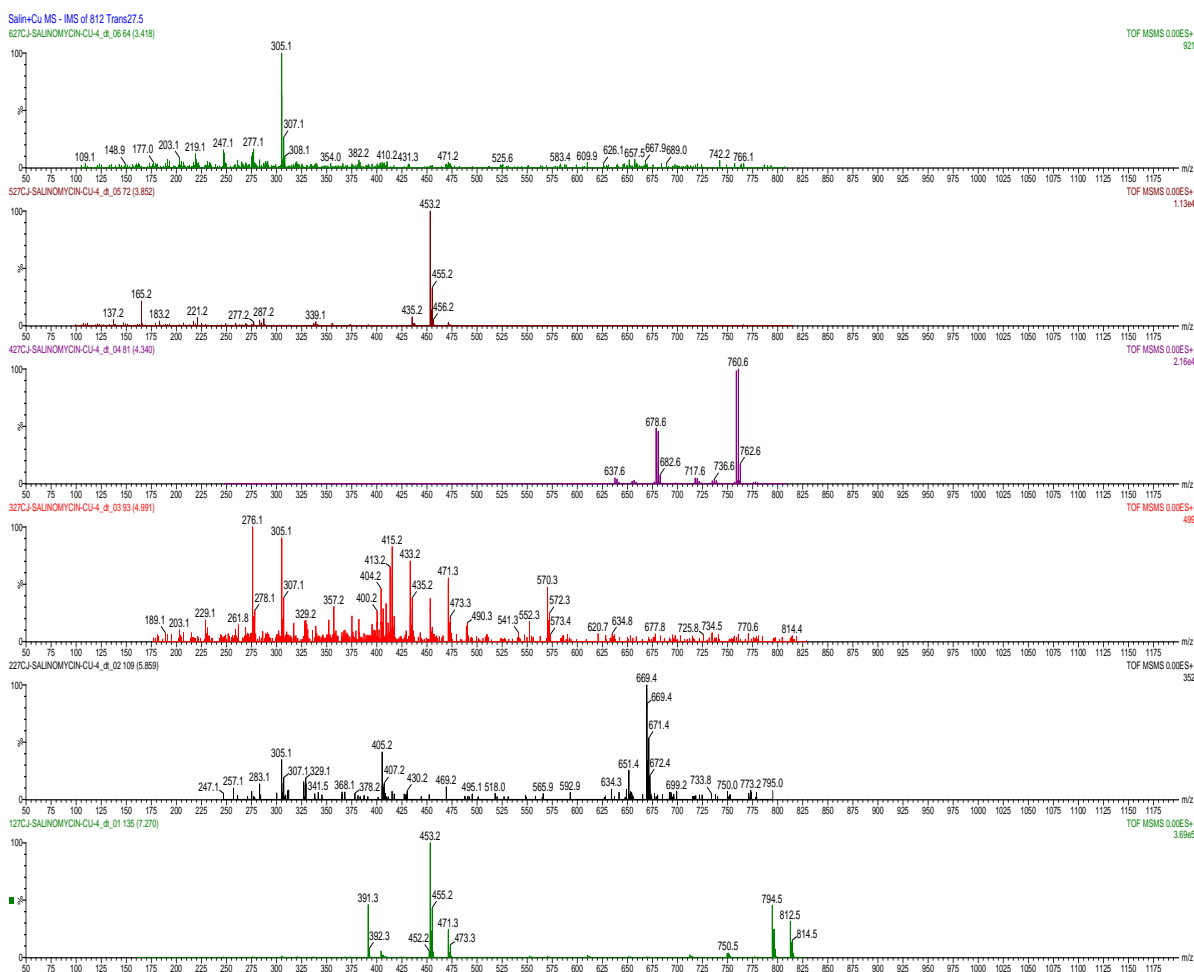


Figure 63 IMS-MS/MS drift time drift scope data and MS/MS spectra for $[\text{SAL-H}+\text{Cu}]^+$ at transfer collision energy 27.5 eV in positive ion mode.

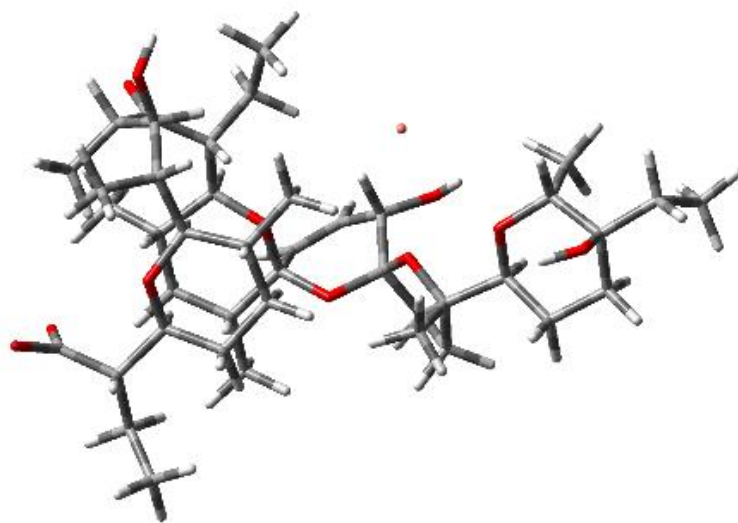


Figure 64 Gaussian modelling of an optimised structure for $[SAL+Cu^{II}-H]^+$. The optimised model is represented by tubes in different colours for clarity. The copper ion is shown in orange in the middle, chelating with the oxygens on the spiroketal rings. Oxygen atoms are represented in red, hydrogen atoms are white, and carbon is in light grey. Bonding is shown in dark grey.

4.4. Conclusion

This chapter focused on investigating the related polyether ionophores, salinomycin, and narasin. These two polyethers were selected based on their antibiotic and pharmaceutical benefits and their tendency to isomerise in solution and to produce isomeric products during MS/MS. The methodologies developed for analysing amino acids and flavonoids in the earlier chapters were applied to these two polyether ionophores. Both salinomycin and narasin gave satisfactory results and proved the efficacy of combining mass spectrometry with computational methods in structure elucidation analysis and isomer differentiation. Detailed mechanisms were proposed for salinomycin, its isomer, and the narasin sodium ion.

Metal cation complexes of salinomycin have also been studied, and novel fragmentation pathways were discovered when Fe^{II} and Cu^{II} were used as the chelating cation. Complete fragmentation pathways and mechanisms were proposed for the selected metal cation complexes. The binding affinity of salinomycin for various metal cations was established. Salinomycin has also shown the ability to chelate with multiple lithium, sodium, and potassium ions. These ‘double metal’ adduct ions also provided distinctive fragmentation pathways. These findings are useful in future structural elucidation analysis of other natural products.

The methodology was expanded to include high-resolution IMS-MS/MS as an additional tool in isomer differentiation. We have introduced the term ‘ionoform’ to define metal cation complexes of salinomycin, which are different only in the metal binding locations. It is theorised that the fragmentation pathway would be altered due to the changes in the metal cation binding position. This has been proven with IMS-MS/MS and computational data obtained in this research. By utilising ion mobility and theoretical CCS calculations, ionoforms could be distinguished and rationalised. Overall, results obtained from this research could provide new analytical insight into other disciplines, such as pharmaceutical drug development. The methodology developed should be robust and adaptable to other natural product related research.

5. Conclusions

Mass spectrometry (MS) is a fundamental analytical tool in nearly every discipline within the applied chemical sciences, which is used for both qualitative and quantitative studies in association with chromatography. The studies of natural products by ESI-MS/MS have attracted a vast amount of interest due to the advances in instrumentation and their robust analytical power. Understanding the fragmentation reactions is crucial to elucidate the structure and to identify natural products successfully.

This research used three groups of natural products to systematically analyse their fragmentations using tandem mass spectrometry and computational methods. It developed a robust methodology that obtained results that could be applied to future related studies.

Systematically studying the α -amino acids confirmed that the side chain functionalities play a significant role in influencing their fragmentation pathways. Both the positive and negative ion mode CID data were collected and studied. Differentiation in structural isomers such as leucine and isoleucine was achieved. HCD studies were also used to generate additional information on the fragmentation pathways of the aromatic amino acids. Complete mechanisms were provided for all the amino acids studied in positive and negative ion modes. Energy breakdown graphs were plotted for the fragmentations of the protonated amino acids. Energy breakdown graphs correlate the peak intensities of the different product ions to the collision energy used in the tandem mass spectrometry experiment. By analysing these graphs, the competition between neutral losses could be examined.

Computational analysis has been utilised in this study to investigate the fragmentation pathways of the 20 α -amino acids. Calculations performed by Gaussian 09 provided insight into the most likely sites of protonation, proton affinity, and gas phase basicity. In addition, IRC calculations in Gaussian 09 made locating transition states and following reaction paths achievable. The reaction profile of serine has been created, as well as a simulation of its fragmentation pathway. This is a great way of visualising fragmentations observed in tandem mass spectra and can be further applied to a larger variety of natural products. It has been shown that Gaussian is a very powerful analytical tool and can complement studies performed by tandem mass spectrometry.

The methodologies developed during the first part of the research have proven their applicability to related small natural products. It was then applied to a class of compounds with considerable isomerism and more competing fragmentation pathways - flavonoids. Chapter 3 investigated the MS/MS spectra for 30 flavonoids in positive and negative ion modes. Common losses are summarised and compared. From these, three diagnostic product ions were identified, these are m/z 153 – indicative of two OH

groups on ring A; m/z 167 – indicative of one OH and one methoxy group on ring A; and m/z 151 – a flavanol, with no ketone oxygen but two OH groups on ring A. Additionally, a methyl radical loss has been observed for flavonoids having methoxy groups in their structures. Other common losses, include loss of H₂O and CO. Differentiation of the isomeric flavonoids rhamnetin and isorhamnetin, was achieved in the negative ion mode by an observation of the highly unusual loss of CH₄. The proton affinity of selected flavonoids was calculated to reveal that ketone oxygen is the most energetically favourable protonation site. This agrees with the experimental observation in their mass spectra and energy breakdown graphs.

The final result chapter utilised the improved methodologies for analysing the fragmentation of natural products in the study of polyether ionophores salinomycin and narasin. These two polyethers were selected based on their antibiotic and pharmaceutical benefits and their tendencies to isomerise in solution. Both salinomycin and narasin gave satisfactory results and proved the efficacy of combining mass spectrometry with computational methods in structure elucidation analysis and isomer differentiation. Detailed mechanisms were proposed for salinomycin, its isomer, and the narasin sodium ion complex.

Metal cation complexes of salinomycin were also studied, and novel fragmentation pathways were discovered when Fe^{II} and Cu^{II} were used as chelating cations. Complete fragmentation mechanisms were proposed for the selected metal cation complexes. The binding affinity of salinomycin for various metal cations was established, and salinomycin was shown to chelate with two alkali metal cations. These doubly metallated adduct ions also provided distinctive fragmentation pathways. These findings will be useful in future structural elucidation analyses of other natural products.

The methodology was expanded to include high-resolution IMS-MS/MS to separate isomers in the gas phase. Ionoforms (i.e. metal cation complexes, which are different only in the metal binding locations) were observed due to alterations in fragmentation pathways that could not be rationalised otherwise. These results proved the power of combining IMS-MS/MS and computational data to study highly complex molecules and their isomerism. By utilising ion mobility and theoretical CCS calculations, ionoforms could be distinguished and rationalised. Overall, results obtained from this research could provide new analytical insight into other disciplines, such as pharmaceutical drug development. The methodology developed was robust and adaptable and should be highly applicable to other natural product related research.

For future work, the study of more complex natural products using the developed methodologies from this research should be undertaken. This includes increasing the library of known flavonoids to be studied to discover unknown flavonoids in substances such as plants or beverages. Additionally, isomer

differentiation for more secondary metabolites through the existing method is another potential research area. Data from the study of salinomycin iron and copper complexes offered a high degree of complexity, requiring more detailed investigations. Results from this could be applied to similar ionophores or other polyethers.

To conclude, this thesis presented a systematic compilation of the gas-phase fragmentation reactions of three groups of natural products and increased our understanding of how molecules fragment. High-end high-resolution accurate mass techniques were used to perform the experiments to obtain reliable data and accurate results. CID-MSⁿ and IMS-MS/MS were applied to extend the range of product ions observed to lower masses. We also improved the methodologies for identifying and differentiating structural isomers, especially combining the routine mass spectrometry method with IMS when separating isomers and isobars. Lastly, identifying novel product ions of selected natural products by combining mass spectrometry and computational methods was achieved.

REFERENCES

1. E. Goldstein, *Berl. Ber.*, 1886, **39**, 601.
2. W. Wien, *Annalen der Physik*, 1898, **301**, 440-452.
3. J. J. Thomson, Longmans, Green, 1913, ch. 1, p. 20.
4. S. Meyerson, *Org. Mass Spectrom.*, 1986, **21**, 197-208.
5. R. S. Gohlke, *Anal. Chem.*, 1959, **31**, 535-541.
6. J. C. Holmes and F. A. Morrell, *Appl. Spectrosc.*, 1957, **11**, 86-87.
7. A. J. Dempster, *Physical Review*, 1921, **18**, 415-422.
8. M. S. B. Munson and F. H. Field, *Journal of the American Chemical Society*, 1966, **88**, 2621-2630.
9. N. C. Fenner and N. R. Daly, *Rev. Sci. Instrum.*, 1966, **37**, 1068-1070.
10. E. d. Hoffmann and V. Stroobant, J. Wiley, Chichester, West Sussex, England; Hoboken, NJ, 2007, ch. 1.11, pp. 43-52.
11. C. J. McNeal and R. D. Macfarlane, *Journal of the American Chemical Society*, 1981, **103**, 1609-1610.
12. M. Barber, R. S. Bordoli, R. D. Sedgwick and A. N. Tyler, *J. Chem. Soc., Chem. Commun.*, 1981, DOI: 10.1039/C39810000325, 325-327.
13. W. Aberth, K. M. Straub and A. Burlingame, *Anal. Chem.*, 1982, **54**, 2029-2034.
14. K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida and T. Matsuo, *Rapid Commun. Mass Spectrom.*, 1988, **2**, 151-153.
15. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse, *Science*, 1989, **246**, 64-71.
16. E. Hoffmann and V. Stroobant, J. Wiley, Chichester, West Sussex, England; Hoboken, NJ, 2007, ch. 2.3, pp. 122-123.
17. D. J. Harvey, in *Encyclopedia of Analytical Science (Second Edition)*, eds. P. Worsfold, A. Townshend and C. Poole, Elsevier, Oxford, 2005, DOI: <https://doi.org/10.1016/B0-12-369397-7/00348-4>, pp. 350-359.
18. D. S. Millington, in *Encyclopedia of Analytical Science (Third Edition)*, eds. P. Worsfold, C. Poole, A. Townshend and M. Miró, Academic Press, Oxford, 2019, DOI: <https://doi.org/10.1016/B978-0-12-409547-2.14534-2>, pp. 437-442.
19. T. Bonk and A. Humeny, *The Neuroscientist*, 2001, **7**, 6-12.
20. P. J. Gates, *Eur J Mass Spectrom (Chichester)*, 2021, **27**, 13-28.
21. S. Medhe, *Mass Spectrom Purif Tech*, 2018, 4:1
22. W. Bleakney, *Physical Review*, 1929, **34**, 157-160.
23. P. Gates, Electron Ionisation (EI), <http://www.chm.bris.ac.uk/ms/ei-ionisation.xhtml>, (accessed 17th June, 2014).
24. J. Baggott and J. Baggott, in *The Quantum Cookbook: Mathematical Recipes for the Foundations for Quantum Mechanics*, Oxford University Press, 2020, DOI: 10.1093/oso/9780198827856.003.0005, p. 0.
25. T. W. Bentley and R. A. W. Johnstone, in *Adv. Phys. Org. Chem.*, ed. V. Gold, Academic Press, 1970, vol. 8, pp. 151-269.
26. C. Dass, in *Fundamentals of Contemporary Mass Spectrometry*, 2007, DOI: <https://doi.org/10.1002/9780470118498.ch2>, pp. 15-65.
27. S. E. Stein and D. R. Scott, *J. Am. Soc. Mass. Spectrom.*, 1994, **5**, 859-866.
28. M. Barber, R. S. Bordoli, R. D. Sedgwick and A. N. Tyler, *Nature*, 1981, **293**, 270-275.
29. E. d. Hoffmann and V. Stroobant, J. Wiley, Chichester, West Sussex, England; Hoboken, NJ, 2007, ch. 1.4, pp. 30-33.
30. B. J. Sweetman and I. A. Blair, *Biomedical & Environmental Mass Spectrometry*, 1988, **17**, 337-340.
31. M. Barber, D. Bell, M. Eckersley, M. Morris, L. Tetler and P. J. Derrick, *Rapid Commun. Mass Spectrom.*, 1988, **2**, 18-21.
32. P. Gates, Fast Atom Bombardment (FAB) & Liquid Secondary Ion Mass Spectrometry (LSIMS), <http://www.chm.bris.ac.uk/ms/fab-ionisation.xhtml>, 2014).

33. R. D. Macfarlane and D. F. Torgerson, *Science*, 1976, **191**, 920-925.
34. S. Bouchonnet, J. P. Denhez, Y. Hoppilliard and C. Mauriac, *Anal. Chem.*, 1992, **64**, 743-754.
35. H. M. Fales, *International Journal of Mass Spectrometry and Ion Physics*, 1983, **53**, 59-67.
36. M. Irion, W. Bowers, R. Hunter, F. Rowland and R. McIver Jr, *Chem. Phys. Lett.*, 1982, **93**, 375-379.
37. F. Hillenkamp, *Springer Series in Chemical Physics.*, 1983, **25**, 190-191
38. F. Hillenkamp, *Springer Series in Chemical Physics.*, 1986, **44**, 471-472
39. G. Van Der Peyl, W. Van Der Zande and P. Kistemaker, *Int. J. Mass Spectrom. Ion Processes*, 1984, **62**, 51-71.
40. R. Conzemius, S. Zhao, R. Houk and H. Svec, *Int. J. Mass Spectrom. Ion Processes*, 1984, **61**, 277-292.
41. E. K. Fukuda and J. E. Campana, *Anal. Chem.*, 1985, **57**, 949-952.
42. M. Karas and F. Hillenkamp, *Anal. Chem.*, 1988, **60**, 2299-2301.
43. N. F. W. Ligterink, V. Grimaudo, P. Moreno-García, R. Lukmanov, M. Tulej, I. Leya, R. Lindner, P. Wurz, C. S. Cockell, P. Ehrenfreund and A. Riedo, *Scientific Reports*, 2020, **10**, 9641.
44. D. S. Peterson, *Mass Spectrom. Rev.*, 2007, **26**, 19-34.
45. M. Karas, D. Bachmann and F. Hillenkamp, *Anal. Chem.*, 1985, **57**, 2935-2939.
46. M. Karas, D. Bachmann, U. Bahr and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Processes*, 1987, **78**, 53-68.
47. K. Tanaka, *Angew. Chem. Int. Ed.*, 2003, **42**, 3860-3870.
48. M. Karas and F. Hillenkamp, *Anal. Chem.*, 1988, **60**, 2299-2301.
49. M. Karas, U. Bahr and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Processes*, 1989, **92**, 231-242.
50. R. C. Beavis and B. T. Chait, *Rapid Commun. Mass Spectrom.*, 1989, **3**, 436-439.
51. P. Gates, Matrix-assisted Laser Desorption/Ionisation (MALDI), <http://www.chm.bris.ac.uk/ms/maldi-ionisation.xhtml>).
52. R. Zenobi and R. Knochenmuss, *Mass Spectrom. Rev.*, 1998, **17**, 337-366.
53. U. Conway, A. D. Warren and P. J. Gates, *Analyst*, 2021, **146**, 5988-5994.
54. X. Wen, S. Dagan and V. H. Wysocki, *Anal. Chem.*, 2007, **79**, 434-444.
55. B. D. Glen, D. F. Leslie, F. Keith, J. Florian, L. K. Danielle, M. Thomas and F. M. Anthony, *J. Organomet. Chem.*, 2014, **751**, 482-492.
56. L. A. McDonnell and R. M. Heeren, *Mass Spectrom. Rev.*, 2007, **26**, 606-643.
57. C. S. Ho, C. W. Lam, M. H. Chan, R. C. Cheung, L. K. Law, L. C. Lit, K. F. Ng, M. W. Suen and H. L. Tai, *Clin Biochem Rev*, 2003, **24**, 3-12.
58. M. Dole, L. L. Mack, R. L. Hines, R. C. Mobley, L. D. Ferguson and M. B. Alice, *J. Chem. Phys.*, 1968, **49**, 2240-2249.
59. M. Yamashita and J. B. Fenn, *J. Phys. Chem.*, 1984, **88**, 4451-4459.
60. J. B. Fenn, *Angew. Chem. Int. Ed.*, 2003, **42**, 3871-3894.
61. P. Gates, Electrospray Ionisation (ESI), <http://www.chm.bris.ac.uk/ms/esi-ionisation.xhtml>).
62. G. Taylor, *Proc. R. Soc. London, Ser. A*, 1964, **280**, 383-397.
63. A. Gomez and K. Tang, *Physics of Fluids*, 1994, **6**, 404-414.
64. J. V. Iribarne and B. A. Thomson, *J. Chem. Phys.*, 1976, **64**, 2287-2294.
65. M. A. Kelly, M. M. Vestling, C. C. Fenselau and P. B. Smith, *Org. Mass Spectrom.*, 1992, **27**, 1143-1147.
66. T. Covey, in *Biochemical and Biotechnological Applications of Electrospray Ionization Mass Spectrometry*, American Chemical Society, 1996, vol. 619, ch. 2, pp. 21-59.
67. M. Wilm and M. Mann, *Anal. Chem.*, 1996, **68**, 1-8.
68. P. Kebarle and L. Tang, *Anal. Chem.*, 1993, **65**, 972A-986A.
69. G. J. Van Berkel and F. Zhou, *Anal. Chem.*, 1995, **67**, 2916-2923.
70. J. B. Fenn, J. Rosell, T. Nohmi, S. Shen and F. J. Banks, Jr., in *Biochemical and Biotechnological Applications of Electrospray Ionization Mass Spectrometry*, American Chemical Society, 1996, vol. 619, ch. 3, pp. 60-80.
71. D. I. Carroll, I. Dzidic, R. N. Stillwell, K. D. Haegele and E. C. Horning, *Anal. Chem.*, 1975, **47**, 2369-2373.

72. C. J. Proctor and J. F. J. Todd, *Org. Mass Spectrom.*, 1983, **18**, 509-516.
73. P. J. Gates, *Eur. J. Mass Spectrom.*, 2021, **27**, 13-28.
74. C. H. Hocart, in *Comprehensive Natural Products II*, eds. H.-W. Liu and L. Mander, Elsevier, Oxford, 2010, DOI: <https://doi.org/10.1016/B978-008045382-8.00187-8>, pp. 327-388.
75. T. Řezanka, *J. High. Resolut. Chromatogr.*, 2000, **23**, 338-342.
76. Z. Takáts, J. M. Wiseman, B. Gologan and R. G. Cooks, *Science*, 2004, **306**, 471-473.
77. M. W. Towers, T. Karancsi, E. A. Jones, S. D. Pringle and E. Claude, *J. Am. Soc. Mass. Spectrom.*, 2018, **29**, 2456-2466.
78. J. M. Wiseman, D. R. Ifa, Q. Song and R. G. Cooks, *Angew. Chem. Int. Ed.*, 2006, **45**, 7188-7192.
79. E. Claude, E. A. Jones and S. D. Pringle, *Methods Mol Biol*, 2017, **1618**, 65-75.
80. R. G. Hemalatha and T. Pradeep, *J. Agric. Food. Chem.*, 2013, **61**, 7477-7487.
81. D. Parrot, S. Papazian, D. Foil and D. Tasdemir, *Planta Med.*, 2018, **84**, 584-593.
82. J. I. Zhang, N. Talaty, A. B. Costa, Y. Xia, W. A. Tao, R. Bell, J. H. Callahan and R. G. Cooks, *Int. J. Mass spectrom.*, 2011, **301**, 37-44.
83. X. Li, H. Hu, R. Yin, Y. Li, X. Sun, S. K. Dey and J. Laskin, *Anal. Chem.*, 2022, **94**, 9690-9696.
84. S. Santagata, L. S. Eberlin, I. Norton, D. Calligaris, D. R. Feldman, J. L. Ide, X. Liu, J. S. Wiley, M. L. Vestal, S. H. Ramkisson, D. A. Orringer, K. K. Gill, I. F. Dunn, D. Dias-Santagata, K. L. Ligon, F. A. Jolesz, A. J. Golby, R. G. Cooks and N. Y. R. Agar, *Proceedings of the National Academy of Sciences*, 2014, **111**, 11121-11126.
85. L. Hänel, M. Kwiatkowski, L. Heikaus and H. Schlüter, *Future Science OA*, 2019, **5**, FSO373.
86. K. H. Kingdon, *Phys. Rev.*, 1923, **21**, 408-418.
87. A. Makarov, *Anal. Chem.*, 2000, **72**, 1156-1162.
88. E. Denisov, E. Damoc and A. Makarov, *Int. J. Mass spectrom.*, 2021, **466**, 116607.
89. R. H. Perry, R. G. Cooks and R. J. Noll, *Mass Spectrom. Rev.*, 2008, **27**, 661-699.
90. A. Makarov and M. Scigelova, *J. Chromatogr. A*, 2010, **1217**, 3938-3945.
91. A. C. Peterson, G. C. McAlister, S. T. Quarmbly, J. Griep-Raming and J. J. Coon, *Anal. Chem.*, 2010, **82**, 8618-8628.
92. B. B. Misra, *Analytical Methods*, 2021, **13**, 2265-2282.
93. T. B. Angerer, J. Bour, J.-L. Biagi, E. Moskovets and G. Frache, *J. Am. Soc. Mass. Spectrom.*, 2022, **33**, 760-771.
94. N. E. Manicke, A. L. Dill, D. R. Ifa and R. G. Cooks, *Journal of mass spectrometry : JMS*, 2010, **45**, 223-226.
95. J. J. Coon, J. E. P. Syka, J. C. Schwartz, J. Shabanowitz and D. F. Hunt, *Int. J. Mass spectrom.*, 2004, **236**, 33-42.
96. Z. Takáts, J. M. Wiseman, B. Gologan and R. G. Cooks, *Science*, 2004, **306**, 471-473.
97. G. Hopfgartner and E. Varesio, *TrAC, Trends Anal. Chem.*, 2005, **24**, 583-589.
98. A. U. Jackson, A. Tata, C. Wu, R. H. Perry, G. Haas, L. West and R. G. Cooks, *Analyst*, 2009, **134**, 867-874.
99. M. Friia, V. Legros, J. Tortajada and W. Buchmann, *J. Mass Spectrom.*, 2012, **47**, 1023-1033.
100. Q. Hu, N. Talaty, R. J. Noll and R. G. Cooks, *Rapid Commun. Mass Spectrom.*, 2006, **20**, 3403-3408.
101. B. D. Musselman, *Biomedical & Environmental Mass Spectrometry*, 1989, **18**, 942-942.
102. P. Gates, Sequential (Multistage) Mass Spectrometry (MSn), <http://www.chm.bris.ac.uk/ms/msn.shtml>, (accessed 04.03.2023, 2023).
103. M. L. Vestal and J. M. Campbell, *Methods Enzymol.*, 2005, **402**, 79-108.
104. A. K. Shukla and J. H. Futrell, *J Mass Spectrom.*, 2000, **35**, 1069-1090.
105. I. V. Chernushevich, A. V. Loboda, B. A. Thomson, *J. Mass Spectrom.*, 2001, **36**, 849-865.
106. R. A. Yost and C. G. Enke, *Journal of the American Chemical Society*, 1978, **100**, 2274-2275.
107. L. Sleno and D. A. Volmer, *J. Mass Spectrom.*, 2004, **39**, 1091-1112.
108. J. V. Olsen, O. Macek B Fau - Lange, A. Lange O Fau - Makarov, S. Makarov A Fau - Horning, M. Horning S Fau - Mann and M. Mann.
109. M. P. Jedrychowski, W. Huttlin El Fau - Haas, M. E. Haas W Fau - Sowa, R. Sowa Me Fau - Rad, S. P. Rad R Fau - Gygi and S. P. Gygi.

110. M. P. Jedrychowski, E. L. Huttlin, W. Haas, M. E. Sowa, R. Rad and S. P. Gygi, *Mol Cell Proteomics*, 2011, **10**, M111.009910.
111. M. P. Jedrychowski, E. L. Huttlin, W. Haas, M. E. Sowa, R. Rad and S. P. Gygi, *Molecular & Cellular Proteomics*, 2011, **10**, M111.009910.
112. G. A. Eiceman and Z. Karpas, *Ion mobility spectrometry*, CRC Press, Boca Raton, 1994.
113. Richardson K.; Giles K.; Ashcroft A. E.; Sobott F.. Travelling Wave Ion Mobility Separation: Basics and Calibration. In *New Developments in Mass Spectrometry*; Royal Society of Chemistry, 2021, **4**, 83–104.
114. K. Giles, K. R. Pringle, D. Worthington, J. L. Little, R. H. Wildgoose, R. H. Bateman, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 2401-2401.
115. S. D. Pringle, K. Giles, J. L. Wildgoose, J. P. Williams, S. E. Slade, K. Thalassinou, R. H. Bateman, M. T. Bowers and J. H. Scrivens, *Int. J. Mass spectrom.*, 2007, **261**, 1-12.
116. D. M. compiled by Alan and W. Andrew, *Compendium of chemical terminology : IUPAC recommendations*, Second edition. Oxford [Oxfordshire] ; Malden, MA : Blackwell Science, 1997., 1997.
117. M.P.Langevin, *Ann. Chim. Phys.*, 1905, **8**, 245-288.
118. F. Lanucara, S. W. Holman, C. J. Gray and C. E. Eyers. *Nat Chem.*, 2014, **6**, 281-294.
119. V. Domalain, M. Hubert-Roux, V. Tognetti, L. Joubert, C. M. Lange, J. Rouden and C. Afonso, *Chemical Science*, 2014, **5**, 3234-3239.
120. P. M. Lalli, B. A. Iglesias, H. E. Toma, G. F. de Sa, R. J. Daroda, J. C. Silva Filho, J. E. Szulejko, K. Araki and M. N. Eberlin, *J. Mass Spectrom.*, 2012, **47**, 712-719.
121. Z. Hall, A. Politis and Carol V. Robinson, *Structure*, 2012, **20**, 1596-1609.
122. K. Masike, M. A. Stander and A. de Villiers, *J. Pharm. Biomed. Anal.*, 2021, **195**, 113846.
123. A. Petrone, P. Cimino, G. Donati, H. P. Hratchian, M. J. Frisch and N. Rega, *Journal of Chemical Theory and Computation*, 2016, **12**, 4925-4933.
124. W. Deng, J. R. Cheeseman and M. J. Frisch, *Journal of Chemical Theory and Computation*, 2006, **2**, 1028-1037.
125. K. Wang, G. Luo, J. Hong, X. Zhou, L. Weng, Y. Luo and L. Zhang, *Angew. Chem. Int. Ed.*, 2014, **53**, 1053-1056.
126. A. V. Marenich, C. J. Cramer and D. G. Truhlar, *The Journal of Physical Chemistry B*, 2009, **113**, 6378-6396.
127. M. Kaupp, V. G. Malkin, O. L. Malkina and D. R. Salahub, *Journal of the American Chemical Society*, 1995, **117**, 1851-1852.
128. Gaussian 16, Revision C.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.
129. P. R. S. G.Frenking, *J. Comput. Chem.*, 2004, **25**, v-viii.
130. M. J. Frisch, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M.

- Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox., Gaussian 09, <https://gaussian.com/g09citation/>).
131. I. N. Levine, *Quantum chemistry*, Seventh edition. Boston : Pearson, [2009], 2009.
 132. N. Harrison, *Computational materials science*, 2003, **187**, 45.
 133. J. B. Foresman and A. Frisch, Gaussian, Pittsburgh, PA, 2nd ed. edn., 1996, ch. 6, pp. 115-120.
 134. E. Fermi, *Rend. Accad. Naz. Lincei*, 1927, **6**, 32.
 135. L. H. Thomas, *Mathematical Proceedings of the Cambridge Philosophical Society*, 1927, **23**, 542-548.
 136. J. C. Slater, *Physical Review*, 1951, **81**, 385-390.
 137. C. J. Cramer, Wiley, Chichester, 2nd ed. edn., 2004, ch. 7, p. 252.
 138. P. Hohenberg and W. Kohn, *Physical Review*, 1964, **136**, B864-B871.
 139. J. B. Foresman and A. Frisch, Gaussian, Pittsburgh, PA, 2nd ed. edn., 1996, ch. Appendix A, pp. 272-275.
 140. A. D. Becke, *Physical Review A*, 1988, **38**, 3098-3100.
 141. C. Lee, W. Yang and R. G. Parr, *Physical Review B*, 1988, **37**, 785-789.
 142. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
 143. A. D. Becke, *The Journal of Chemical Physics*, 1988, **88**, 1053-1062.
 144. S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **59**, 1200.
 145. J. P. Perdew and Y. Wang, *Physical Review B*, 1992, **45**, 13244-13249.
 146. N. Chéron, D. Jacquemin and P. Fleurat-Lessard, *Physical Chemistry Chemical Physics*, 2012, **14**, 7170-7175.
 147. W. Yu, L. Liang, Z. Lin, S. Ling, M. Haranczyk and M. Gutowski, *J. Comput. Chem.*, 2009, **30**, 589-600.
 148. J. B. Foresman and A. Frisch, Gaussian, Inc., Wallingford, CT, Third edition edn., 2015, ch. 5, pp. 97-103.
 149. R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1971, **54**, 724-728.
 150. T. Shoeib, A. C. Hopkinson and K. W. M. Siu, *The Journal of Physical Chemistry B*, 2001, **105**, 12399-12409.
 151. A. L. Davis, J. A. Harris, C. A. Russell and J. P. Wilkins, *Analyst*, 1999, **124**, 251-256.
 152. C. Larriba and C. J. Hogan Jr, *Journal of Computational Physics*, 2013, **251**, 344-363.
 153. C. Larriba and C. J. Hogan Jr, *The Journal of Physical Chemistry A*, 2013, **117**, 3887-3901.
 154. C. Larriba-Andaluz, M. Nahin and V. Shrivastav, *Aerosol Sci. Technol.*, 2017, **51**, 887-895.
 155. H. Ouyang, C. Larriba-Andaluz, D. R. Oberreit and C. J. Hogan Jr, *J. Am. Soc. Mass. Spectrom.*, 2013, **24**, 1833-1847.
 156. C. Larriba-Andaluz and C. J. H. Jr., *The Journal of Chemical Physics*, 2014, **141**, 194107.
 157. A. A. Shvartsburg and M. F. Jarrold, *Chem. Phys. Lett.*, 1996, **261**, 86-91.
 158. M. Mesleh, J. Hunter, A. Shvartsburg, G. Schatz and M. Jarrold, *The Journal of Physical Chemistry A*, 1997, **101**, 968-968.
 159. V. Shrivastav, M. Nahin, C. J. Hogan and C. Larriba-Andaluz, *Journal of the American Society for Mass Spectrometry*, 2017, **28**, 1540-1551.
 160. C. Larriba, IMoSPedia, <https://www.imospedia.com/imos/>).
 161. EDITORIAL, *Nature Chemical Biology*, 2007, **3**, 351-351.
 162. R. A. Maplestone, M. J. Stone and D. H. Williams, *Gene*, 1992, **115**, 151-157.
 163. D. P. Demarque, A. E. M. Crotti, R. Vessecchi, J. L. C. Lopes and N. P. Lopes, *Natural Product Reports*, 2016, **33**, 432-455.
 164. P. Zhang, W. Chan, I. L. Ang, R. Wei, M. M. T. Lam, K. M. K. Lei and T. C. W. Poon, *Scientific Reports*, 2019, **9**, 6453.
 165. C. Jiang, C. J. Arthur and P. J. Gates, *Analyst*, 2020, **145**, 6632-6638.
 166. B. Harper, E. K. Neumann, S. M. Stow, J. C. May, J. A. McLean and T. Solouki, *Anal. Chim. Acta*, 2016, **939**, 64-72.
 167. P. Shi, Q. He, Y. Song, H. Qu and Y. Cheng, *Anal. Chim. Acta*, 2007, **598**, 110-118.
 168. N. Chhabra, M. L. Aseri and D. Padmanabhan, *Int J Appl Basic Med Res*, 2013, **3**, 16-18.
 169. I. Wagner and H. Musso, *Angewandte Chemie International Edition in English*, 1983, **22**, 816-828.

170. J. Jones, *Amino acid and peptide synthesis*, Oxford University Press, Oxford ; New York, 2nd edn., 2002.
171. J. Clayden, Oxford University Press, Oxford ; New York, 2001, ch. 6, p. 389.
172. G. C. Barrett and D. T. Elmore, *Amino acids and peptides*, Cambridge University Press, Cambridge ; New York, 1998.
173. M. J. Smith, E. P. Mazzola, T. J. Farrell, J. A. Sphon, S. W. Page, D. Ashley, S. R. Sirimanne, R. H. Hill and L. L. Needham, *Tetrahedron Lett.*, 1991, **32**, 991-994.
174. A. D. Bravenec, K. D. Ward and T. J. Ward, *J. Sep. Sci.*, 2018, **41**, 1489-1506.
175. J. L. Bada, in *Methods Enzymol.*, Academic Press, 1984, vol. 106, pp. 98-115.
176. B. S. Meldrum, *J Nutr*, 2000, **130**, 1007s-1015s.
177. P. M. Doraiswamy, *Curr Neurol Neurosci Rep*, 2003, **3**, 373-378.
178. J. P. Greenstein and M. Winitz, *Chemistry of the amino acids*, Krieger, Malabar, Fla., 1984.
179. J. Jones, in *Oxford chemistry primers*, Oxford University Press, Oxford ; New York, 2nd edn., 2002, ch. 2, pp. 8-12.
180. F. W. McLafferty, R. Kornfeld, W. F. Haddon, K. Levsen, I. Sakai, P. F. Bente, S.-C. Tsai and H. D. R. Schuddemage, *Journal of the American Chemical Society*, 1973, **95**, 3886-3892.
181. A. W. Coulter and C. C. Fenselau, *Org. Mass Spectrom.*, 1972, **6**, 105-111.
182. C. Tsang and A. Harrison, *Journal of the American Chemical Society*, 1976, **98**, 1301-1308.
183. K. Biemann, *Biomed Environ Mass Spectrom*, 1988, **16**, 99-111.
184. W. Heerma and W. Kulik, *Biomed Environ Mass Spectrom*, 1988, **16**, 155-159.
185. N. N. Dookeran, T. Yalcin and A. G. Harrison, *J. Mass Spectrom.*, 1996, **31**, 500-508.
186. S. Bouchonnet, J. P. Denhez, Y. Hoppilliard and C. Mauriac, *Anal. Chem.*, 1992, **64**, 743-754.
187. M. Piraud, C. Vianey-Saban, K. Petritis, C. Elfakir, J. P. Steghens, A. Morla and D. Bouchu, *Rapid Commun. Mass Spectrom.*, 2003, **17**, 1297-1311.
188. P. I. Shek, J. Zhao, Y. Ke, K. M. Siu and A. C. Hopkinson, *The journal of physical chemistry A*, 2006, **110**, 8282-8296.
189. S. Koenig, in *Biomacromolecular mass spectrometry yearbook series*, Nova Science Publishers, New York, 2011, ch. 3, pp. 211-220.
190. S.-S. Choi and O.-B. Kim, *Int. J. Mass spectrom.*, 2013, **338**, 17-22.
191. L. Rodriguez-Santiago, M. Sodupe and J. Tortajada, *The Journal of Physical Chemistry A*, 2001, **105**, 5340-5347.
192. R. N. Grewal, H. El Aribi, A. G. Harrison, K. W. M. Siu and A. C. Hopkinson, *The Journal of Physical Chemistry B*, 2004, **108**, 4899-4908.
193. T. Shoeib, A. Cunje, A. C. Hopkinson and K. W. M. Siu, *J. Am. Soc. Mass. Spectrom.*, 2002, **13**, 408-416.
194. A. Gil, S. Simon, L. Rodríguez-Santiago, J. Bertrán and M. Sodupe, *Journal of Chemical Theory and Computation*, 2007, **3**, 2210-2220.
195. Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.
196. S. Koenig, *Biomacromolecular mass spectrometry yearbook*, Nova Science Publishers, New York, 2011.
197. R. L. Thurlkill, G. R. Grimsley, J. M. Scholtz and C. N. Pace, *Protein Sci.*, 2006, **15**, 1214-1218.
198. C. N. Pace, G. R. Grimsley and J. M. Scholtz, *J. Biol. Chem.*, 2009, **284**, 13285-13289.

199. H. El Aribi, G. Orlova, A. C. Hopkinson and K. W. M. Siu, *The Journal of Physical Chemistry A*, 2004, **108**, 3844-3853.
200. G. S. Gorman, J. P. Speir, C. A. Turner and I. J. Amster, *Journal of the American Chemical Society*, 1992, **114**, 3986-3988.
201. A. G. Harrison, *Mass Spectrom. Rev.*, 1997, **16**, 201-217.
202. J. B. F. a. Æ. Frisch, *Exploring Chemistry with Electronic Structure Methods*, Gaussian, Inc.: Wallingford, CT, 2015.
203. A. N. Panche, A. D. Diwan and S. R. Chandra, *J Nutr Sci*, 2016, **5**, e47.
204. A. M. M. Careri, M. Musci, *J. Chromatogr. A*, 1998, **794**, 263-297.
205. F. Ververidis, E. Trantas, C. Douglas, G. Vollmer, G. Kretzschmar and N. Panopoulos, *Biotechnol. J.*, 2007, **2**, 1214-1234.
206. S. Kumar and A. K. Pandey, *Sci. World J.*, 2013, **2013**, 162750.
207. E. Middleton, in *Flavonoids in the living system*, Springer, 1998, pp. 175-182.
208. G. M. Gil-Izquierdo A, Ferreres F, Tomás-Barberán FA., *J. Agric. Food. Chem.*, 2001, **49**, 1035-1041.
209. B. S. Fuhrman B, Vaya J, Belinky PA, Coleman R, Hayek T., *Am J Clin Nutr.*, 1997, **66**, 267-275.
210. A. Cassidy and A.M. Minihane, *Am J Clin Nutr*, 2017, **105**, 10-22.
211. T. P. Nguyen, D. T. Mai, T. H. T. and Do, N.M. Phan, *Nat Prod Res.*, 2017, **12**, 1061-1063.
212. C. S. Wang Y, Yu O., *Appl Microbiol Biotechnol.*, 2011, **91**, 949.
213. P. M. Dewick, in *Medicinal Natural Products*, 2009, vol. 137, p. 86.
214. D. G. Kingston and H. M. Fales, *Tetrahedron*, 1973, **29**, 4083-4086.
215. P. J. Gates and N. P. Lopes, *Int J Anal Chem*, 2012, **2012**, 259217.
216. S. C. Sahu and G. C. Gray, *Food Chem. Toxicol.*, 1997, **35**, 443-447.
217. J. Yinon, D. Issachar and H. G. Boettger, *Org. Mass. Spectrom.*, 1978, **13**, 167-171.
218. M. Stobiecki, *Phytochemistry* 2000, **54**, 237-256.
219. R. J. Hughes, T. R. Croley, C. D. Metcalfe and R. E. March, *Int. J. Mass. Spectrom.*, 2001, **210**, 371-385.
220. D. Tsimogiannis, M. Samiotaki, G. Panayotou and V. Oreopoulou, *Molecules*, 2007, **12**, 593-606.
221. S. Barnes, M. Kirk and L. Coward, *J Agr Food Chem*, 1994, **42**, 2466-2474.
222. P. L. Mauri, L. Iemoli, C. Gardana, P. Riso, P. Simonetti, M. Porrini and P. G. Pietta, *Rapid Commun. Mass Spectrom.*, 1999, **13**, 924-931.
223. D. G. I. KINGSTON, *Pergamon Press*, 1971, **27**, 2691-2700.
224. N. Fabre, I. Rustan, E. de Hoffmann and J. Quetin-Leclercq, *J. Am. Soc. Mass. Spectrom.*, 2001, **12**, 707-715.
225. F. Cuyckens and M. Claeys, *J. Mass Spectrom.*, 2004, **39**, 1-15.
226. M. T. Fernandez, M. L. Mira, M. H. Florencio and K. R. Jennings, *J. Inorg. Biochem.*, 2002, **92**, 105-111.
227. M. Satterfield and J. S. Brodbelt, *Anal. Chem.*, 2000, **72**, 5898-5906.
228. E. Hvattum and D. Ekeberg, *J. Mass Spectrom.*, 2003, **38**, 43-49.
229. C. P. G. Butcher, P. J. Dyson, B. F. G. Johnson, P. R. R. Langridge-Smith, J. S. McIndoe and C. Whyte, *Rapid Commun. Mass Spectrom.*, 2002, **16**, 1595-1598.
230. S. Mörlein, C. Schuster, M. Paal and M. Vogeser, *Clinical Mass Spectrometry*, 2020, **18**, 48-53.
231. T. Murakami, Y. Iwamuro, R. Ishimaru, S. Chinaka, N. Kato, Y. Sakamoto, N. Sugimura and H. Hasegawa, *J. Mass Spectrom.*, 2019, **54**, 205-212.
232. Y. L. Ma, Q. M. Li, H. Van den Heuvel and M. Claeys, *Rapid Commun. Mass Spectrom.*, 1997, **11**, 1357-1364.
233. S. Habtemariam, in *Medicinal Foods as Potential Therapies for Type-2 Diabetes and Associated Diseases*, 2019, DOI: 10.1016/b978-0-08-102922-0.00016-x, pp. 551-578.
234. A. Baracco, G. Bertin, E. Gnocco, M. Legorat, S. Sedocco, S. Catinella, D. Favretto and P. Traldi, *Rapid Commun. Mass Spectrom.*, 1995, **9**, 427-436.
235. D. B. H. a. J. M. H. Seema Bhagwat, *Journal*, 159.
236. K. Sak, *Pharmacogn Rev*, 2014, **8**, 122-146.

237. M. Shields, in *Pharmacognosy*, eds. S. Badal and R. Delgoda, Academic Press, Boston, 2017, DOI: <https://doi.org/10.1016/B978-0-12-802104-0.00014-7>, pp. 295-313.
238. Y. Zhang, A. Y. Chen, M. Li, C. Chen and Q. Yao, *J Surg Res*, 2008, **148**, 17-23.
239. B.-W. Kim, E.-R. Lee, H.-M. Min, H.-S. Jeong, J.-Y. Ahn, J.-H. Kim, H.-Y. Choi, H. Choi, E. Y. Kim, S. P. Park and S.-G. Cho, *Cancer Biol Ther*, 2008, **7**, 1080-1089.
240. P. Maher, *Front Biosci (Schol Ed)*, 2015, **7**, 58-82.
241. S. Har Bhajan and K. A. Bharati, in *Handbook of Natural Dyes and Pigments*, eds. S. Har Bhajan and K. A. Bharati, Woodhead Publishing India, 2014, DOI: <https://doi.org/10.1016/B978-93-80308-54-8.50006-X>, pp. 33-260.
242. E. M. Aldred, C. Buck and K. Vall, in *Pharmacology*, eds. E. M. Aldred, C. Buck and K. Vall, Churchill Livingstone, Edinburgh, 2009, DOI: <https://doi.org/10.1016/B978-0-443-06898-0.00038-4>, pp. 299-306.
243. J. Kang, L. A. Hick and W. E. Price, *Rapid Commun. Mass Spectrom.*, 2007, **21**, 857-868.
244. J. Schmidt, *J. Mass Spectrom.*, 2016, **51**, 33-43.
245. L. Zhang, L. Xu, S. S. Xiao, Q. F. Liao, Q. Li, J. Liang, X. H. Chen and K. S. Bi, *J. Pharm. Biomed. Anal.*, 2007, **44**, 1019-1028.
246. P. J. Gates and N. P. Lopes, *International Journal of Analytical Chemistry*, 2012, **2012**, 259217.
247. N. P. Lopes, T. Fonseca, J. P. Wilkins, J. Staunton and P. J. Gates, *Chem. Commun.*, 2003, 72-73.
248. S. Lin, H. Liu, E. B. Svenningsen, M. Wollesen, K. M. Jacobsen, F. D. Andersen, J. Moyano-Villameriel, C. N. Pedersen, P. Nørby, T. Tørring and T. B. Poulsen, *Nature Chemistry*, 2021, **13**, 47-55.
249. R. S. Cockrell, E. J. Harris and B. C. Pressman, *Biochemistry*, 1966, **5**, 2326-2335.
250. J. W. Westley, M. Dekker, New York, 1982, ch. 1, p. 7.
251. C. J. Dutton, B. J. Banks and C. B. Cooper, *Natural Product Reports*, 1995, **12**, 165-181.
252. A. Anadón and M. R. Martínez-Larrañaga, in *Encyclopedia of Food Safety*, ed. Y. Motarjemi, Academic Press, Waltham, 2014, DOI: <https://doi.org/10.1016/B978-0-12-378612-8.00246-8>, pp. 63-75.
253. J. Rutkowski and B. Brzezinski, *Biomed Res Int*, 2013, **2013**, 162513.
254. D. A. Kevin Ii, D. A. Meujo and M. T. Hamann, *Expert Opin Drug Discov*, 2009, **4**, 109-146.
255. P. B. Gupta, T. T. Onder, G. Jiang, K. Tao, C. Kuperwasser, R. A. Weinberg and E. S. Lander, *Cell*, 2009, **138**, 645-659.
256. H. Wang, H. Zhang, Y. Zhu, Z. Wu, C. Cui and F. Cai, *Frontiers in Oncology*, 2021, **11**.
257. T. Koltai, S. J. Reshkin and S. Harguindey, in *An Innovative Approach to Understanding and Treating Cancer: Targeting pH*, eds. T. Koltai, S. J. Reshkin and S. Harguindey, Academic Press, 2020, DOI: <https://doi.org/10.1016/B978-0-12-819059-3.00015-0>, pp. 335-359.
258. C. Naujokat and R. Steinhart, *Journal of Biomedicine and Biotechnology*, 2012, **2012**.
259. B. C. Pressman, *Annu. Rev. Biochem*, 1976, **45**, 501-530.
260. L. Larzabal, N. El-Nikhely, M. Redrado, W. Seeger, R. Savai and A. Calvo, *PLOS ONE*, 2013, **8**, e79798.
261. W. K. Kim, J.-H. Kim, K. Yoon, S. Kim, J. Ro, H. S. Kang and S. Yoon, *Investigational new drugs*, 2012, **30**, 1311-1318.
262. A. Skeberdytė, I. Sarapinienė, J. Aleksander-Krasko, V. Stankevičius, K. Sužiedėlis and S. Jarmalaitė, *Scientific reports*, 2018, **8**, 1-13.
263. Y. Miyazaki, M. SHIBUYA, H. SUGAWARA, O. KAWAGUCHI, C. HIROSE, J. NAGATSU and S. ESUMI, *The Journal of Antibiotics*, 1974, **27**, 814-821.
264. A. S. T. Y. Y. Miyazaki, M. Hara, K. Hara, S. Yoneda, H. Kasahara, Y. Nakamura, *European Patent Office*, 1978, **Publication Number EP 0 000 037 A1**, 23.
265. S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson and L. Katz, *Science*, 1991, **252**, 675-679.
266. J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevitt and P. F. Leadlay, *Nature*, 1990, **348**, 176-178.
267. M. E. Yurkovich, P. A. Tyrakis, H. Hong, Y. Sun, M. Samborsky, K. Kamiya and P. F. Leadlay, *ChemBioChem*, 2012, **13**, 66-71.

268. Y. Li, J. Fang, S. Wu, K. Ma, H. Li, X. Yan and F. Dong, *Analytical and bioanalytical chemistry*, 2010, **398**, 955-961.
269. W. J. Blanchflower and D. G. Kennedy, *Journal of Chromatography B: Biomedical Sciences and Applications*, 1996, **675**, 225-233.
270. V. Hormazábal and M. Yndestad, *J Liq Chromatogr Relat Techno.*, 2000, **23**, 1585-1598.
271. L. Mortier, E. Daeseleire and C. V. Peteghem, *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*, 2005, **19**, 533-539.
272. H. Kinashi and N. Ōtake, *Agricultural and Biological Chemistry*, 1976, **40**, 1625-1632.
273. F. W. McLafferty, *Anal. Chem.*, 1959, **31**, 82-87.
274. X. S. Miao, R. E. March and C. D. Metcalfe, *Rapid Commun. Mass Spectrom.*, 2003, **17**, 149-154.
275. Y. H. Kim, J. S. Yoo, C. H. Lee, Y. M. Goo and M. S. Kim, *J. Mass Spectrom.*, 1996, **31**, 855-860.
276. N. P. Lopes, C. B. W. Stark, J. Staunton and P. J. Gates, *Org Biomol Chem*, 2004, **2**, 358-363.
277. C. Adamo and V. Barone, *The Journal of Chemical Physics*, 1998, **108**, 664-675.
278. Y. Yang and H. Gao, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2012, **85**, 303-309.
279. G. Calderisi, H. Glasner and K. Breuker, *Angew. Chem. Int. Ed.*, 2020, **59**, 4309-4313.
280. N. P. Lopes, F. A. Almeida-Paz and P. J. Gates, *Revista Brasileira de Ciências Farmacêuticas*, 2006, **42**.
281. F. Carnevale Neto, T. N. Clark, N. P. Lopes and R. G. Linington, *J. Nat. Prod.*, 2022, **85**, 519-529.
282. S. M. Stow, N. M. Lareau, K. M. Hines, C. R. McNees, C. R. Goodwin, B. O. Bachmann and J. A. Mclean, *Natural Products Analysis: Instrumentation, Methods, and Applications*, 2014, 397-431.

Appendix

1. Table of Abbreviation

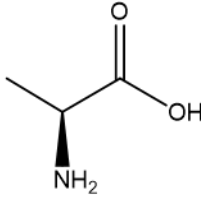
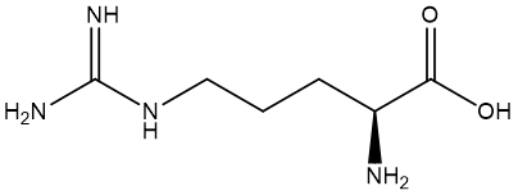
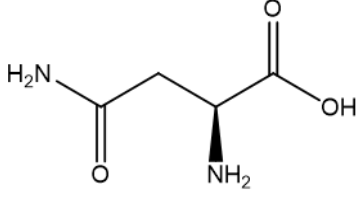
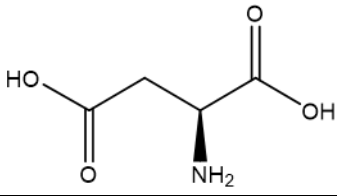
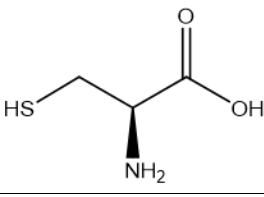
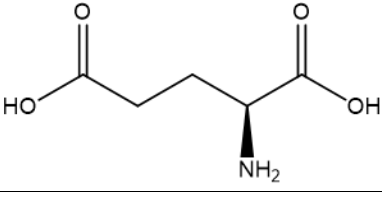
Full Name	Abbreviation (Abbr)
Mass-to-charge ratio	m/z
Mass Spectrometry	MS
Gas chromatography	GC-MS
Electron Ionisation	EI
Chemical Ionisation	CI
Fast Atom Bombardment	FAB
Plasma Desorption	PD
Electrospray Ionisation	ESI
Electrospray Ionisation Mass Spectrometry	ESI-MS
Atmospheric Pressure Chemical Ionisation	APCI
Liquid Chromatography	LC
High-Performance Liquid Chromatography	HPLC
Electron Transfer Dissociation	ETD
Linear Trap Quadrupole	LTQ
Desorption Electrospray Ionisation	DESI
Desorption Atmospheric Pressure Chemical Ionisation	DAPCI
Fourier Transform	FT
Nanospray	nESI
Tandem Mass Spectrometry	MS/MS
Sequential Mass Spectrometry	MS ⁿ
Time of Flight	ToF
Triple Quadrupole Mass Spectrometer	QqQ
Fourier-Transform Ion Cyclotron Resonance	FTICR
Ion Mobility Spectrometry	IMS
Ion Mobility-Mass Spectrometry	IM-MS

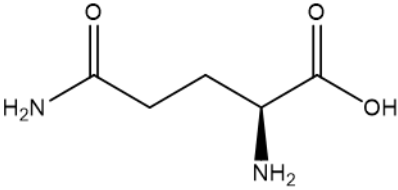
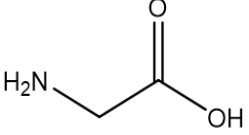
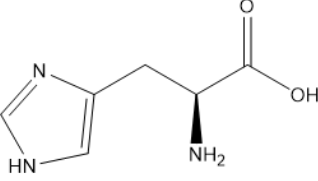
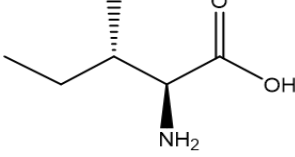
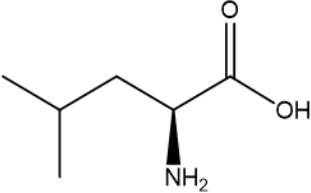
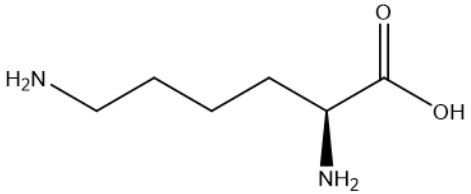
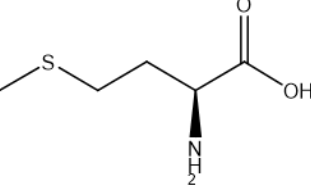
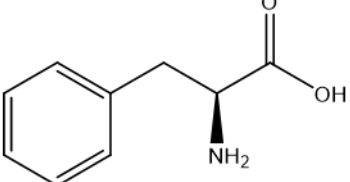
Travelling Wave Ion Mobility Spectrometry	TWIMS
Ion Mobility Spectrometry-Quadrupole-Time of Flight	IMS-Q-ToF
Collision Induced Dissociation	CID
Collision Cross Section	CCS
Higher-energy Collisional Dissociation	HCD
Hartree-Fock	HF
Self-Consistent Field	SCF
Density Function Theory	DFT
Local Density Approximation	LDA
Generalised Gradient Approximation	GGA
Lee, Yang, Parr's Gradient-Corrected Correlation Functional	LYP
Perdew-Wang 91 Correlation Functional	PW91
Ion Mobility Spectrometry Suite	IMoS
NADPH	Nicotinamide adenine dinucleotide phosphate
Amino acids [*]	AA
Salinomycin	SAL
Narain	Nar
Atmospheric pressure chemical ionisation	APCI
Neutral Molecule	M
Sodiated Molecular Ion	[M+Na] ⁺
Metal Molecular Ion Complex	[M+Metal] ⁺
Protonated Molecular Ion	[M+H] ⁺
Deprotonated Molecular Ion	[M-H] ⁻
Diode-Array Detector	DAD
Retro Diels-Alder	RDA

* For individual amino acid abbreviation refer to next page 2. Lists of Analytes

2. Lists of Analytes

2.1 Amino acids

Name	Abbr (3-letter)	Formula	Structure
Alanine	Ala	$C_3H_7NO_2$	
Arginine	Arg	$C_6H_{14}N_4O_2$	
Asparagine	Asn	$C_4H_8N_2O_3$	
Aspartic acid	Asp	$C_4H_7NO_4$	
Cysteine	Cys	$C_3H_7NO_2S$	
Glutamic acid	Glu	$C_5H_9NO_4$	

Glutamine	Gln	C ₅ H ₁₀ N ₂ O ₃	
Glycine	Gly	C ₂ H ₅ NO ₂	
Histidine	His	C ₆ H ₉ N ₃ O ₂	
Isoleucine	Ile	C ₆ H ₁₃ NO ₂	
Leucine	Leu	C ₆ H ₁₃ NO ₂	
Lysine	Lys	C ₆ H ₁₄ N ₂ O ₂	
Methionine	Met	C ₅ H ₁₁ NO ₂ S	
Phenylalanine	Phe	C ₉ H ₁₁ NO ₂	

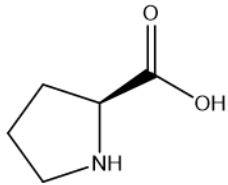
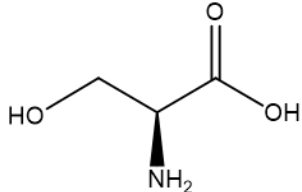
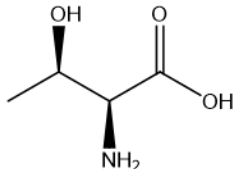
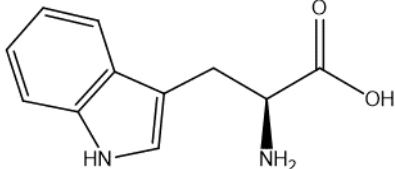
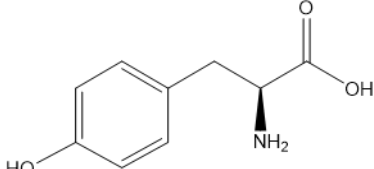
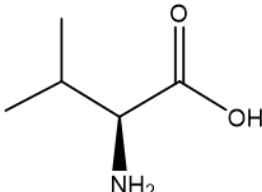
Proline	Pro	$C_5H_9NO_2$	
Serine	Ser	$C_3H_7NO_3$	
Threonine	Thr	$C_4H_9NO_3$	
Tryptophan	Trp	$C_{11}H_{12}N_2O_2$	
Tyrosine	Tyr	$C_9H_{11}NO_3$	
Valine	Val	$C_5H_{11}NO_2$	

Table 1 List of α -L-amino acids used in this research, their three-letter abbreviations, formula, and structures. For clarity, amino acids are colour coded: **Black** – Aliphatic AA; **Red** – Acidic AA; **Blue** – Basic AA; **Green** – Hydroxylic AA; **Brown** – Aromatic AA; **Purple** – Amidic AA; **Orange** – Sulfur Containing AA.

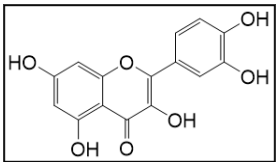
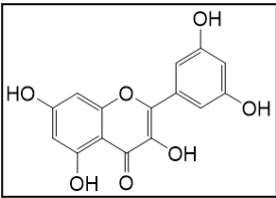
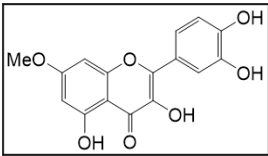
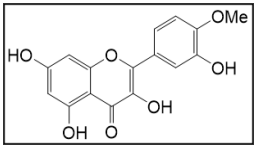
Name	Exact Mass (Da)	Protonated Mass (Da)	Sodiated Mass (Da)	Deprotonated Mass (Da)
Alanine	89.0477	90.0550	112.0369	88.0477
Arginine	174.1117	175.1190	197.1009	173.1117
Asparagine	132.0535	133.0608	155.0427	131.0535
Aspartic acid	133.0375	134.0448	156.0267	132.0375
Cysteine	121.0197	122.0270	144.0090	120.0197
Glutamic acid	147.0532	148.0604	170.0424	146.0532
Glutamine	146.0691	147.0764	169.0584	145.0691

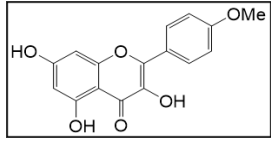
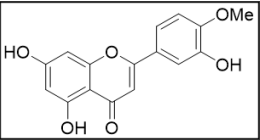
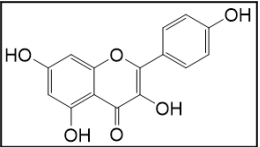
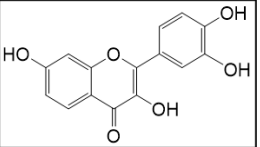
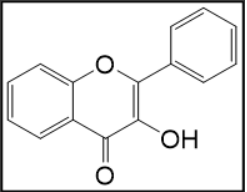
Glycine	75.0320	76.0693	98.0212	74.0320
Histidine	155.0695	156.0768	178.0587	154.0695
Isoleucine	131.0946	132.1019	154.0838	130.0946
Leucine	131.0946	132.1019	154.0838	130.0946
Lysine	146.1055	147.1128	169.0947	145.1055
Methionine	149.0510	150.0583	172.0403	148.0510
Phenylalanine	165.0790	166.0863	188.0682	164.0790
Proline	115.0633	116.0706	138.0525	114.0633

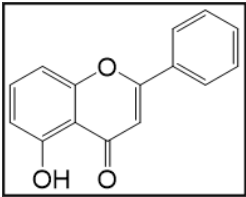
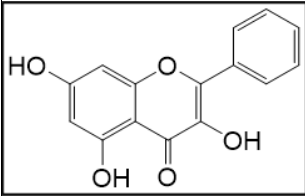
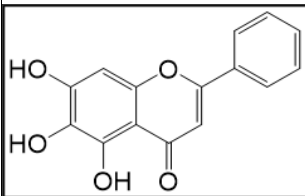
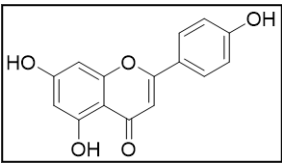
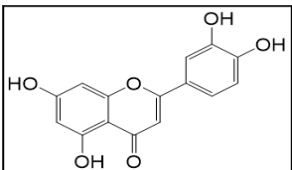
Serine	105.0426	106.0499	128.0318	104.0426
Threonine	119.0582	120.0655	142.0475	118.0582
Tryptophan	204.0899	205.0972	227.0791	203.0899
Tyrosine	181.0739	182.0812	204.0631	180.0739
Valine	117.0790	118.0863	140.0682	116.0790

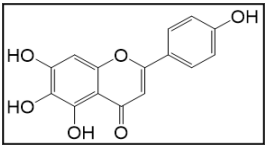
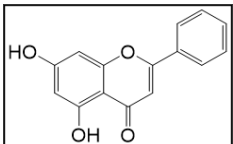
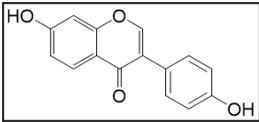
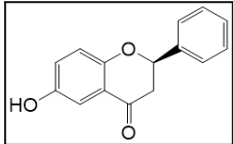
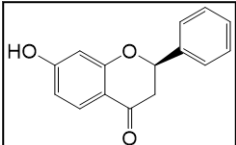
Table 2 List of α -L-amino acids, their exact, protonated, sodiated, and deprotonated masses.

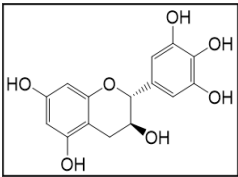
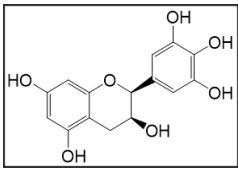
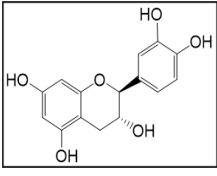
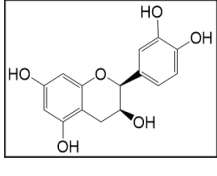
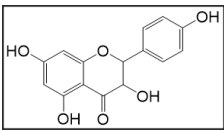
2.2 Flavonoids

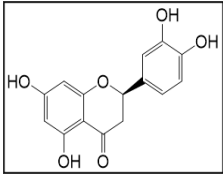
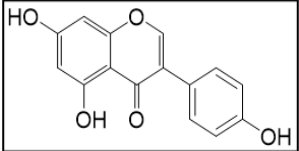
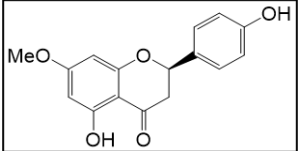
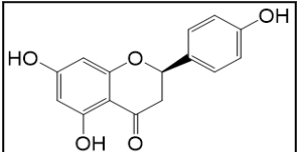
Name	MW	Structure	Protonated	Sodiated	Deprotonated
Quercetin	302		303	325	301
Morin	302		303	325	301
Rhamnetin	316		317	339	315
Isorhamnetin	316		317	339	315

Kaempferide	300		301	323	299
Diosmetin	300		301	323	299
Kaempferol	286		287	309	285
Fisetin	286		287	309	285
3-Hydroxyflavone	238		239	261	237

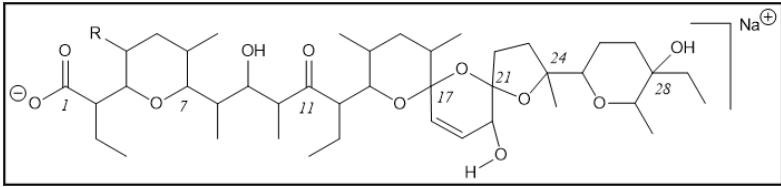
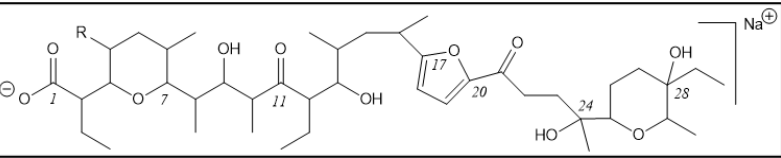
5-Hydroxyflavone	238		239	261	237
Galangin	270		271	293	269
Baicalein	270		271	293	269
Apigenin	270		271	293	269
Luteolin	286		287	309	285

Scutellarein	286		287	309	285
Chrysin	254		255	277	253
Daidzein	254		255	277	253
6-Hydroxyflavanone	240		241	263	239
7-Hydroxyflavanone	240		241	263	239

(-)-Gallocatechin	306		307	329	305
(-)-Epigallocatechin	306		307	329	305
(+) -Catechin	290		291	313	289
(-)-Epicatechin	290		291	313	289
Aromadedrin (Dihydrokaempferol)	288		289	311	287

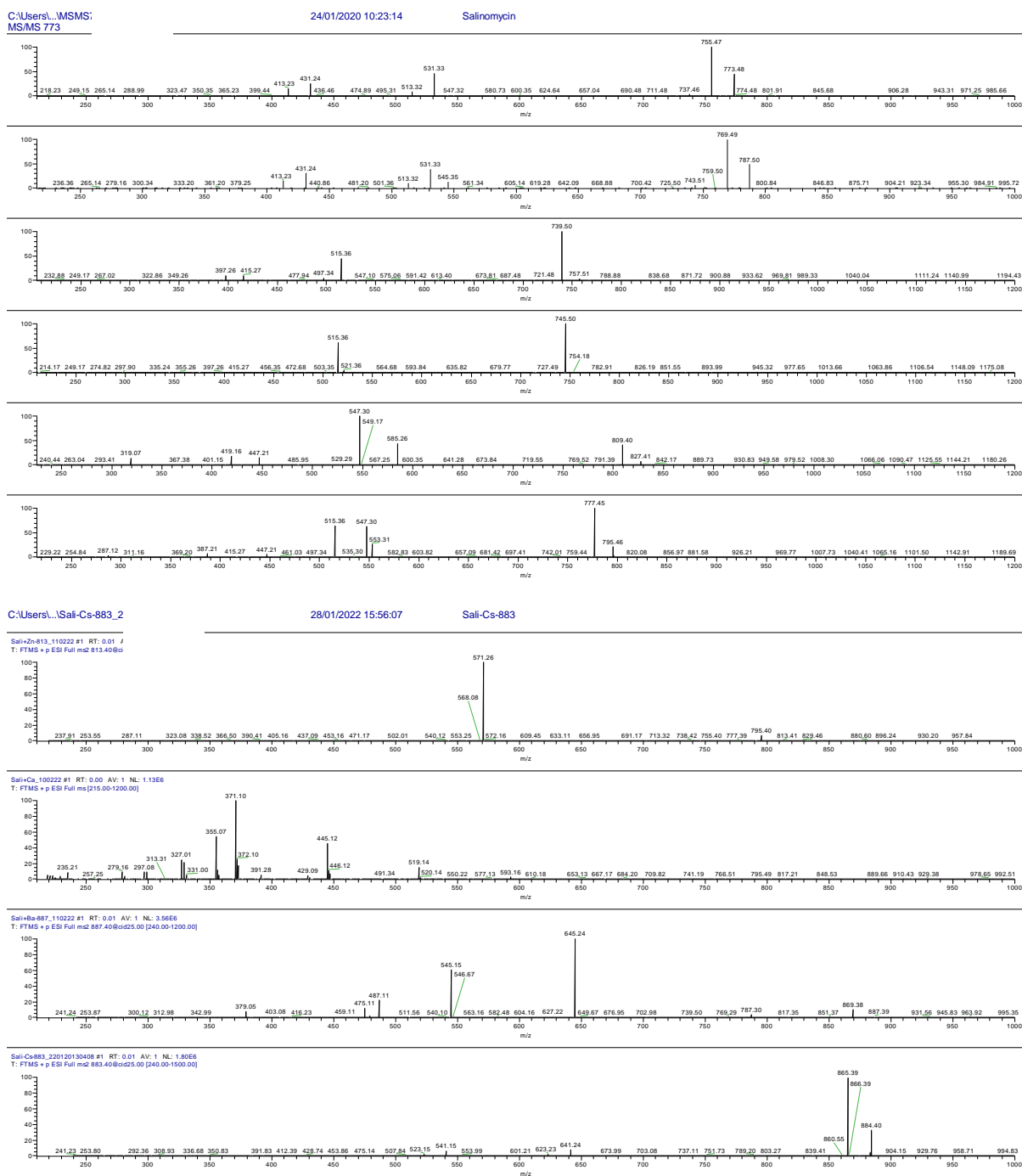
(±)-Eriodictyol	288		289	311	287
Genistein	270		271	293	269
Sakuranetin	286		287	309	285
Naringenin	272		273	295	271

2.3 Polyether

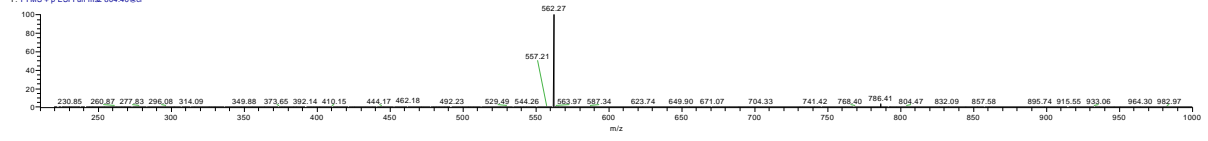
Name	Sodiated Exact Mass	Structure
Salinomycin	773	
Narasin	787	-
Salinomycin Isomer	773	

3. Spectra

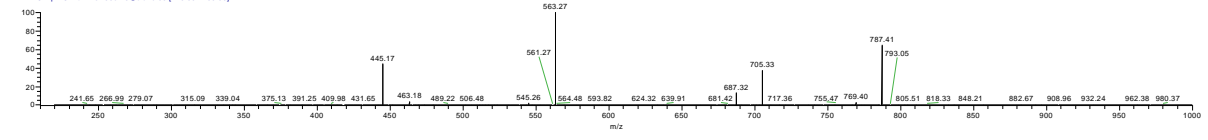
List of MS/MS spectra of salinomycin metal complex ion and narasin monosodium ion.



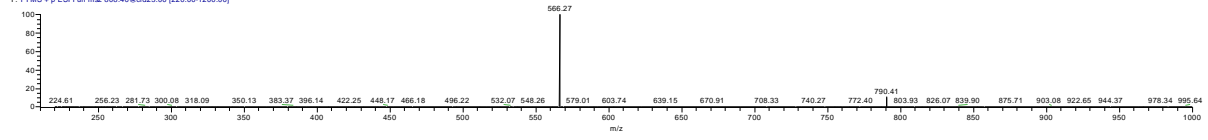
Sali+Fe-805_110222 #1 RT: 0.01 AV: 1 NL: 1.86E5
T: FTMS + p ESI Full ms2 804.40@ci



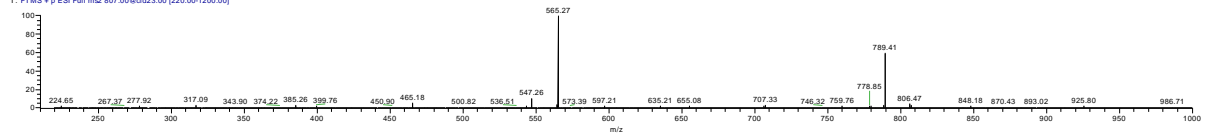
Sali+Fe-805_110222 #1 RT: 0.01 AV: 1 NL: 1.86E5
T: FTMS + p ESI Full ms2 805.40@cid25.00 [20.00-1200.00]



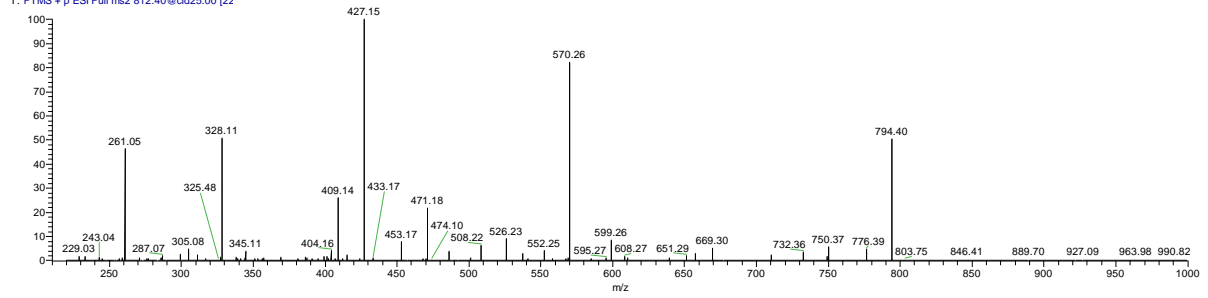
Sali+Co-808_110222 #1 RT: 0.00 AV: 1 NL: 1.07E7
T: FTMS + p ESI Full ms2 808.40@cid25.00 [20.00-1200.00]



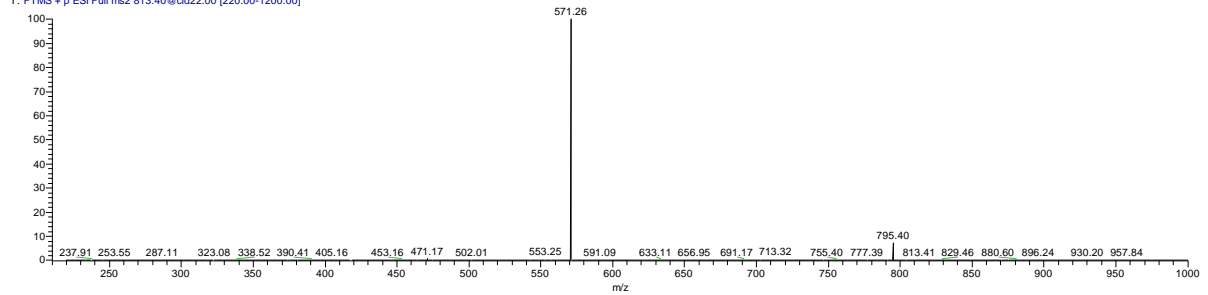
Sali+Ni-807_110222 #1 RT: 0.01 AV: 1 NL: 1.56E4
T: FTMS + p ESI Full ms2 807.00@cid23.00 [20.00-1200.00]



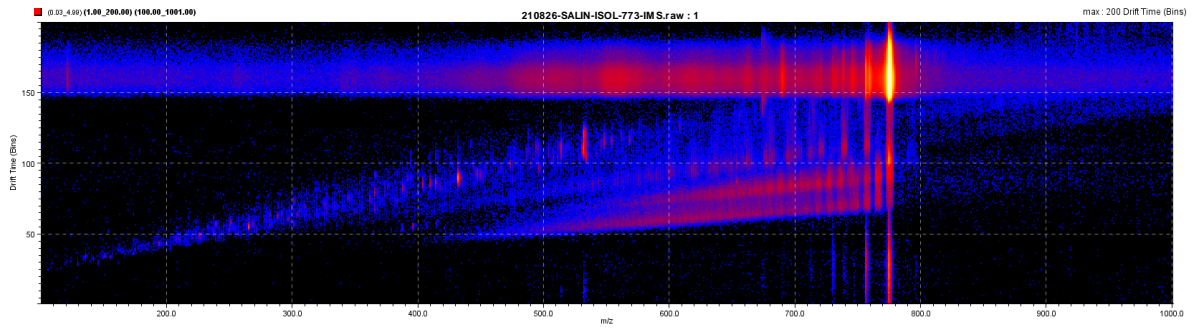
Sali+Cu-812_220120130408 #1 RT: 0.01 AV:
T: FTMS + p ESI Full ms2 812.40@cid25.00 [22



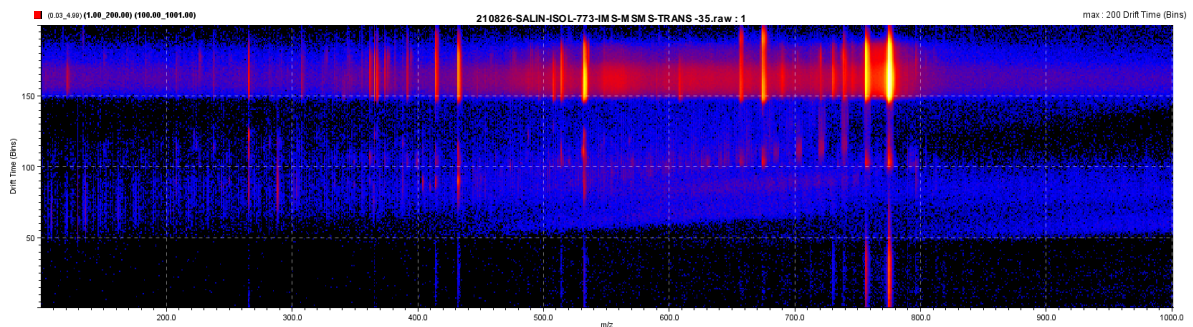
Sali+Zn-813_110222 #1 RT: 0.01 AV: 1 NL: 7.39E6
T: FTMS + p ESI Full ms2 813.40@cid22.00 [20.00-1200.00]



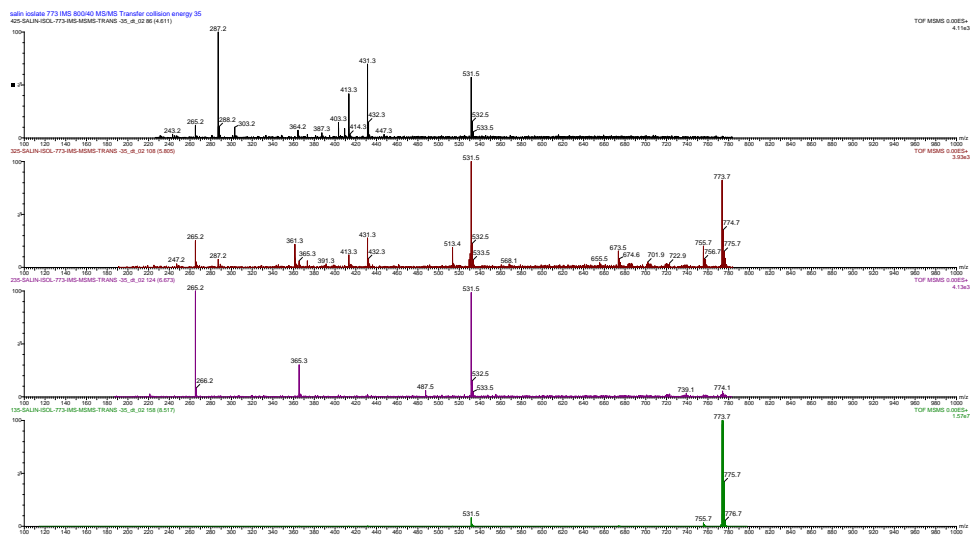
IMS-MS/MS data of salinomycin and narasin



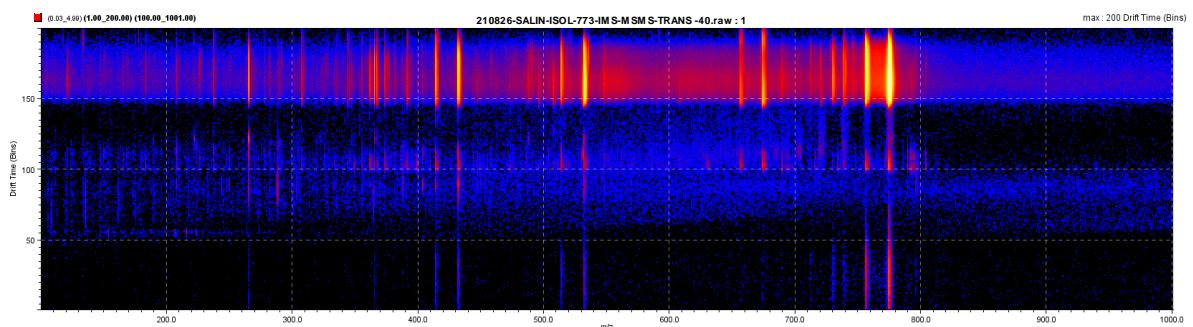
Salinomycin Isolation 773 IMS 800/40



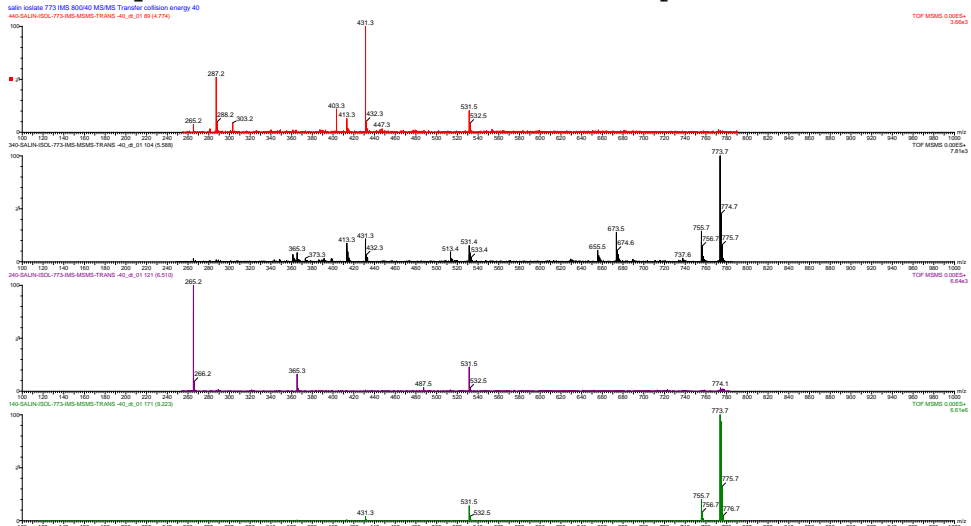
Salinomycin 773 trans 35 eV IMS 800/40 MS/MS



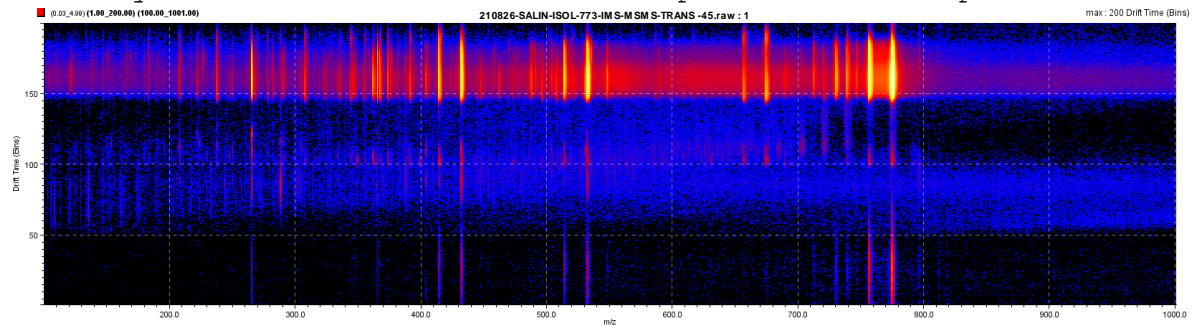
Salinomycin 773 trans 35 eV spec bottom drift time to top



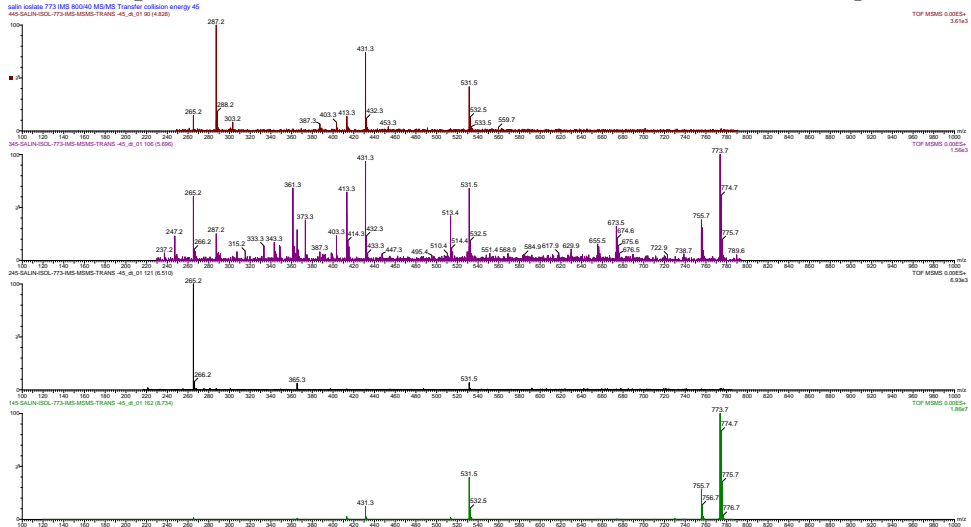
Salinomycin 773 trans 40 eV drift scope drift time



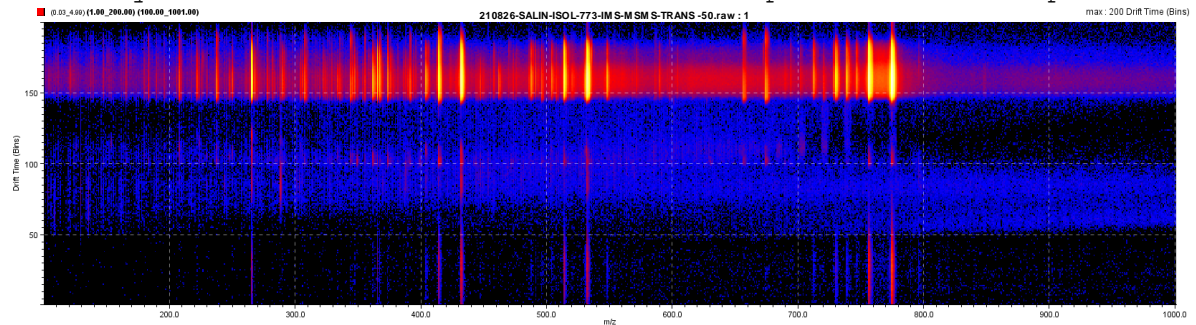
Salinomycin 773 trans 40 eV drift vs spec bottom to top



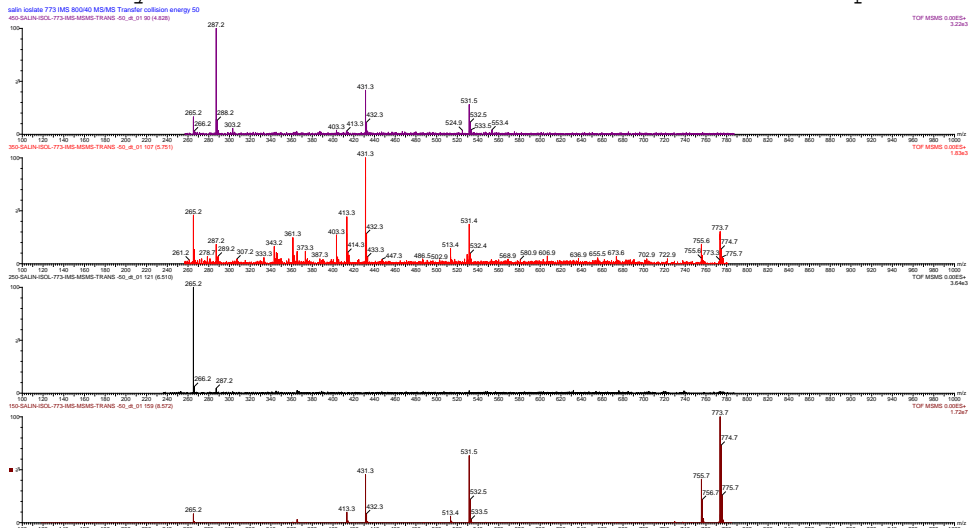
Salinomycin 773 trans 45 eV drift time drift scope



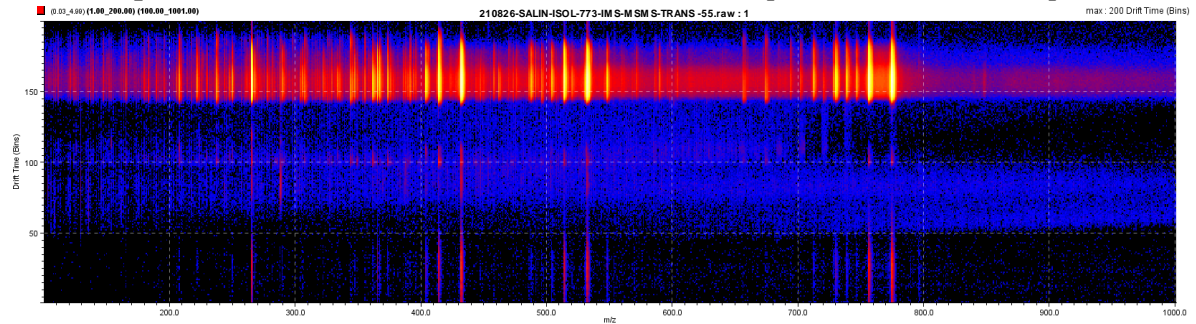
Salinomycin 773 trans 45 eV drift time vs spec bottom to top



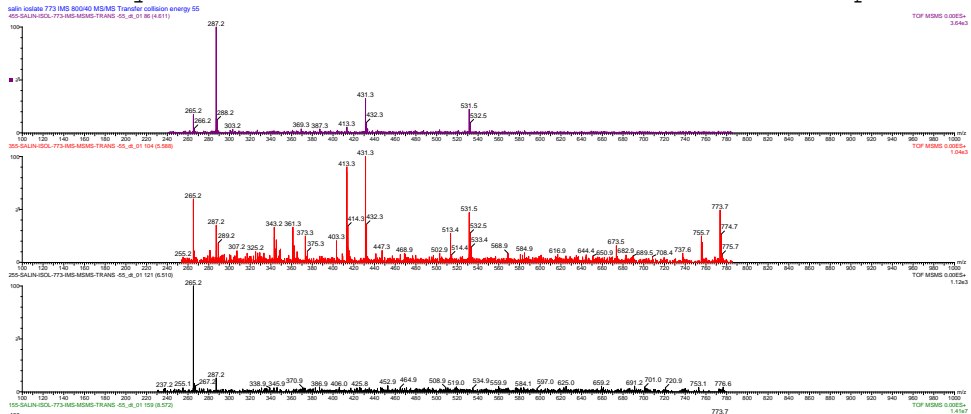
Salinomycin 773 trans 50 eV drift time drift scope



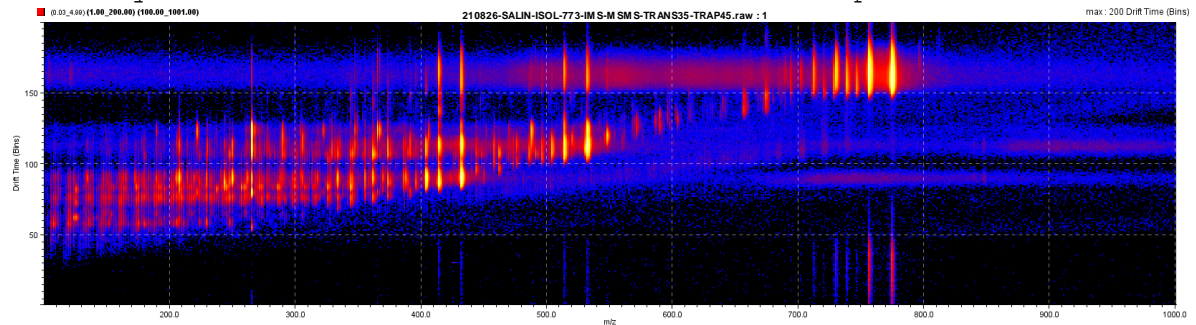
Salinomycin 773 trans 50 eV drift time vs spec bottom to top



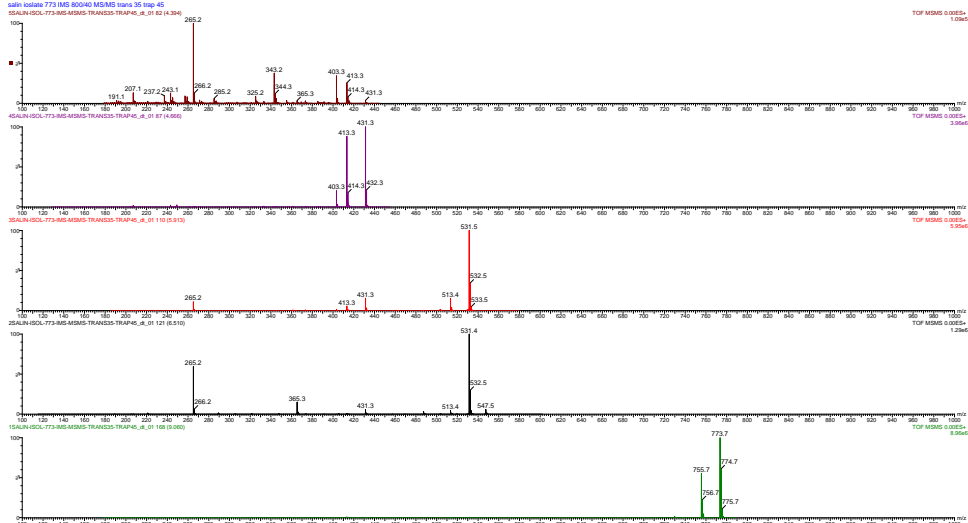
Salinomycin 773 trans 55 eV drift time drift scope



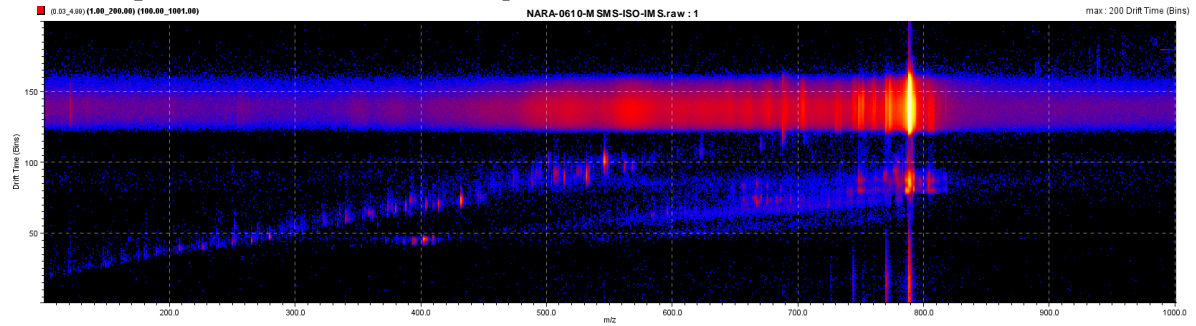
Salinomycin 773 trans 55 eV drift time drift scope



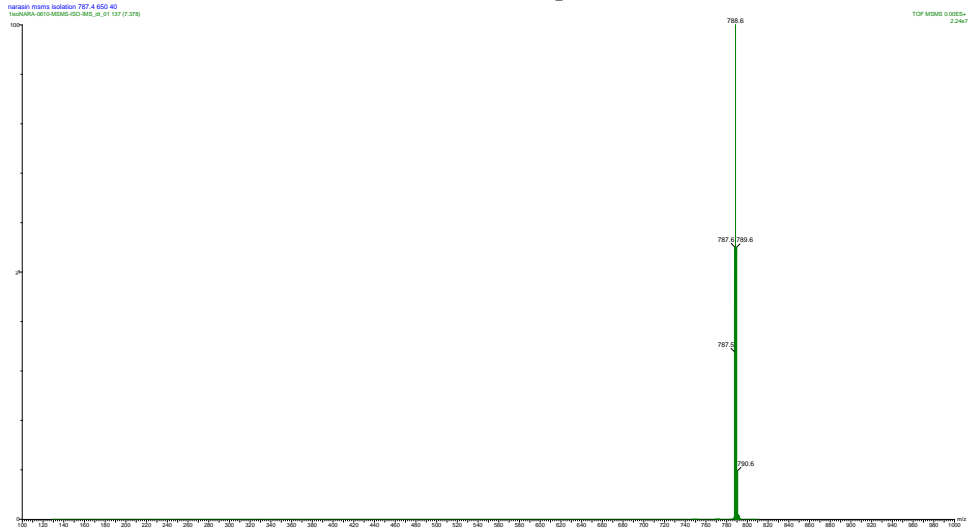
Salinomycin 773 trans 35 trap 45



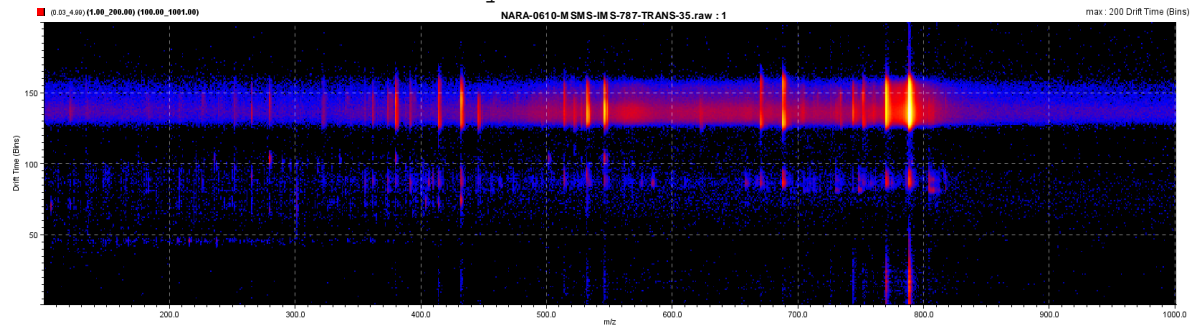
Salinomycin 773 trans 35 trap 45



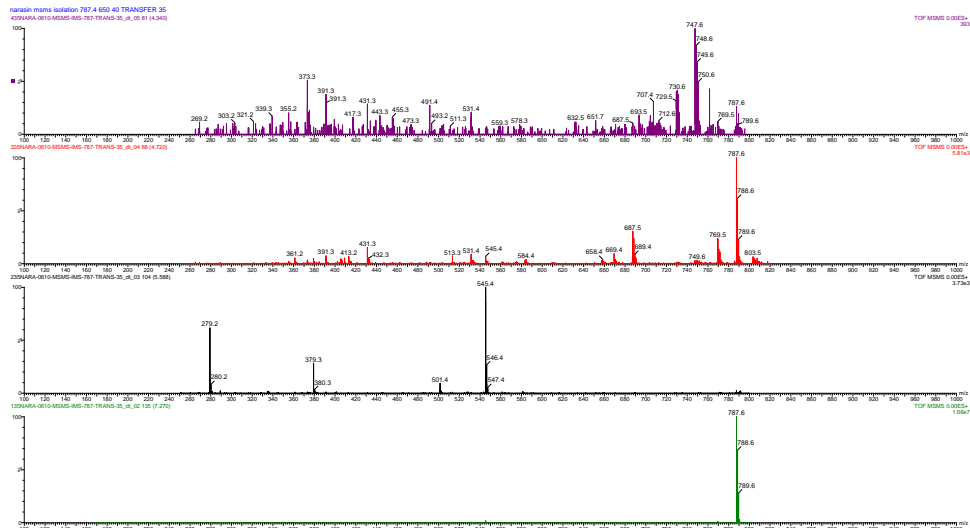
Narasin 787 isolation drift scope



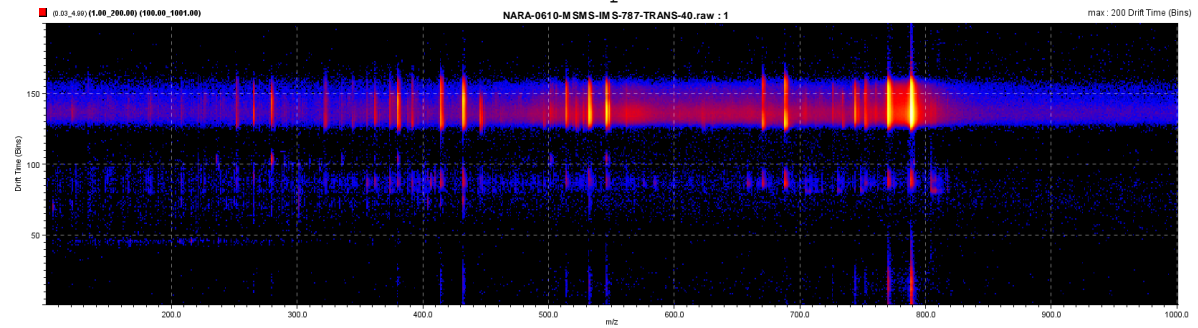
Narasin 787 isolation MassLynx



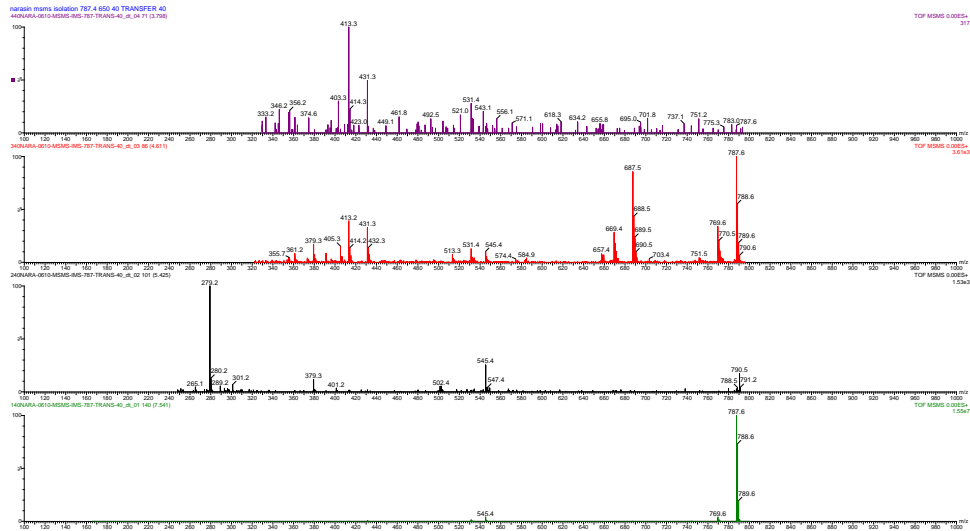
Narasin 787 trans 35 eV



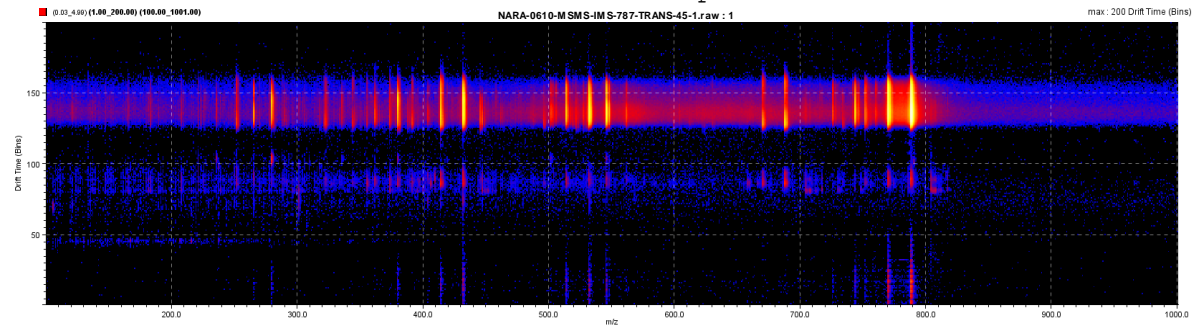
Narasin 787 trans 35 eV drift vs spec



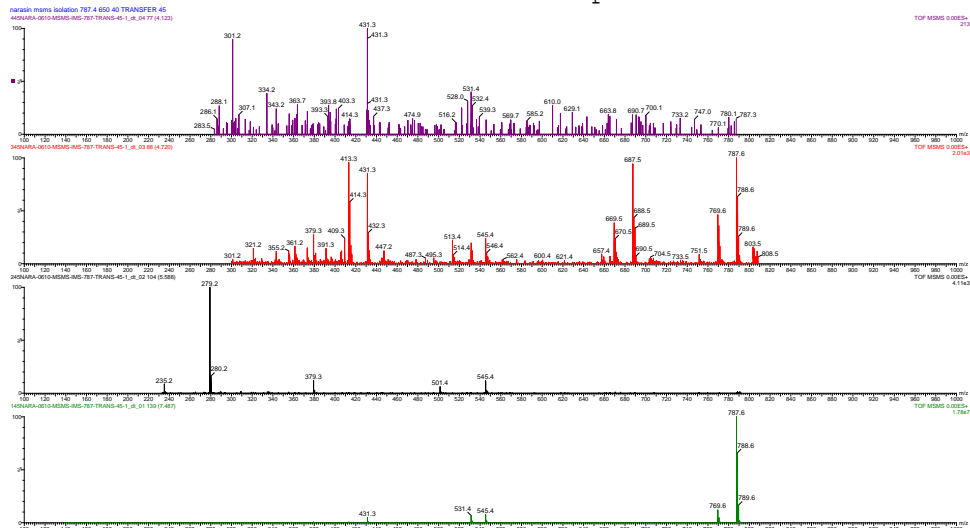
Narasin 787 trans 40 eV



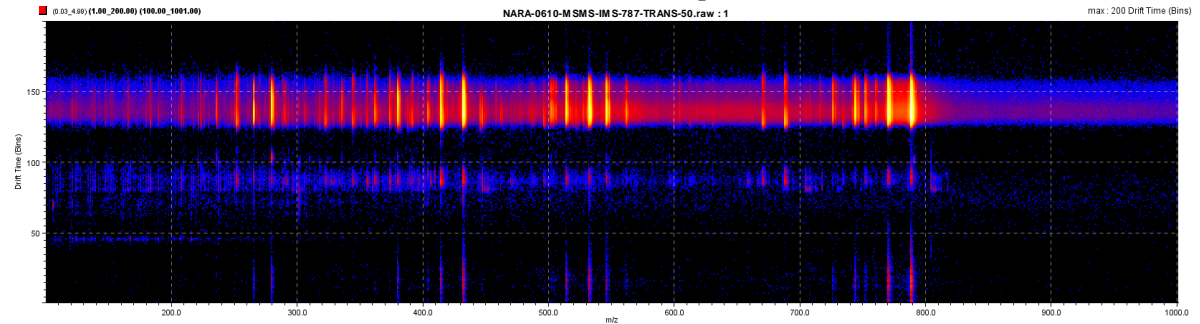
Narasin 787 trans 40 eV drift time vs spec



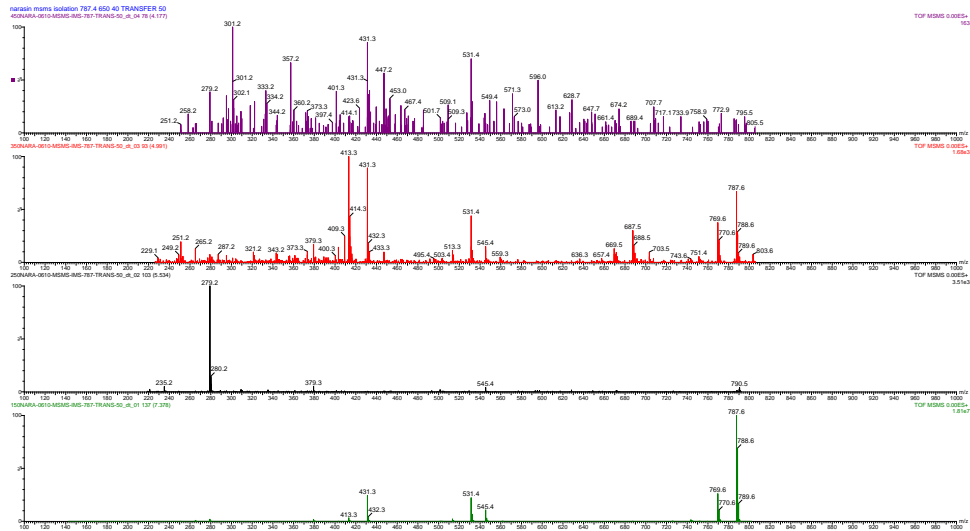
Narasin 787 trans 45 eV drift scope



Narasin 787 trans 45 eV drift time vs spec

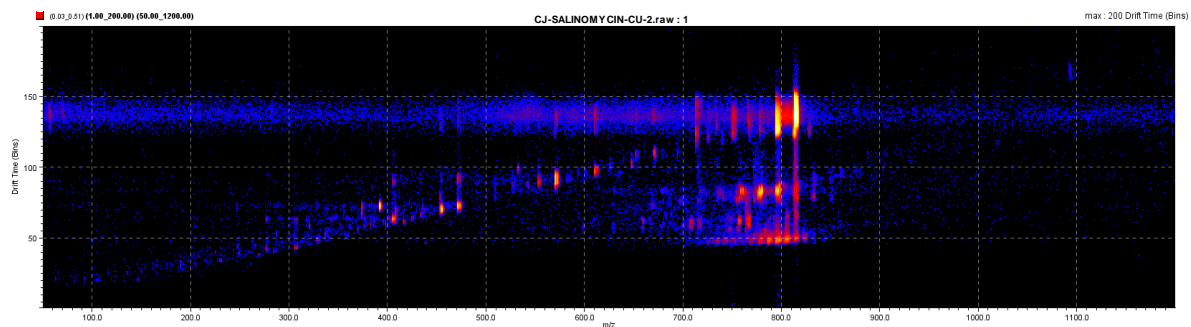


Narasin 787 trans 50 eV drift time

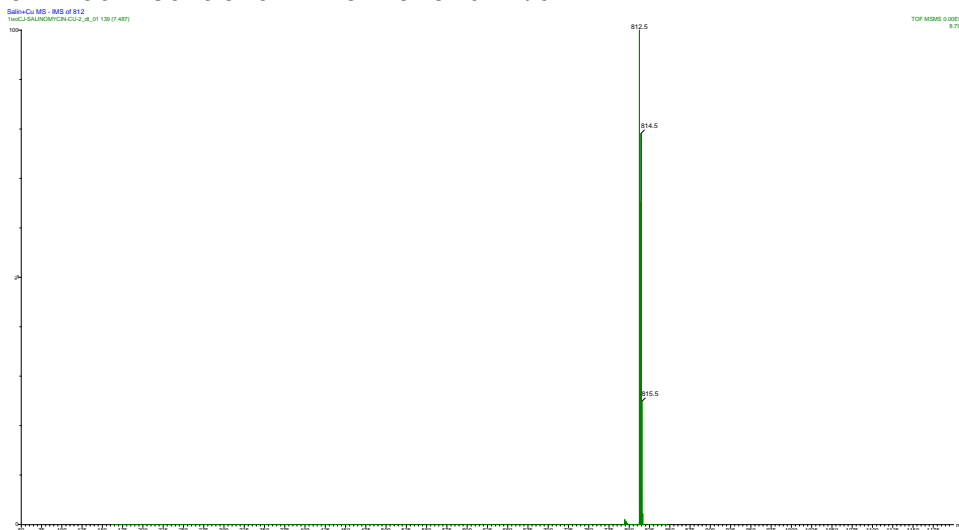


Narasin 787 trans 50 eV drift time vs spec

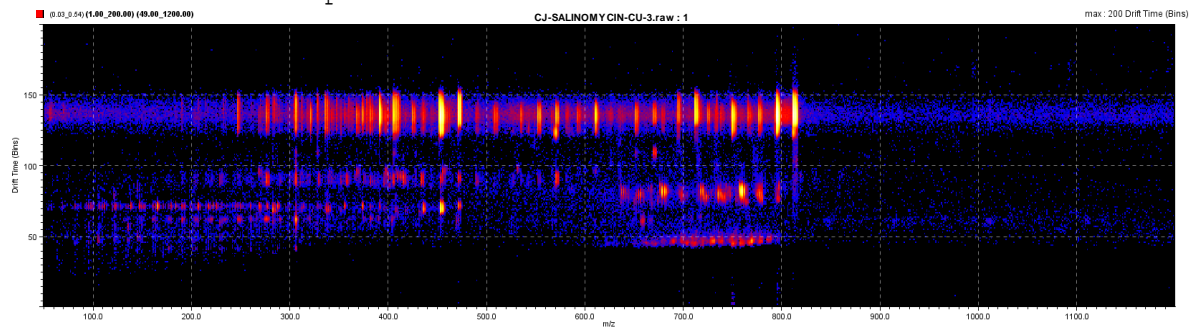
IMS-MS/MS data of salinomycin+Cu complex



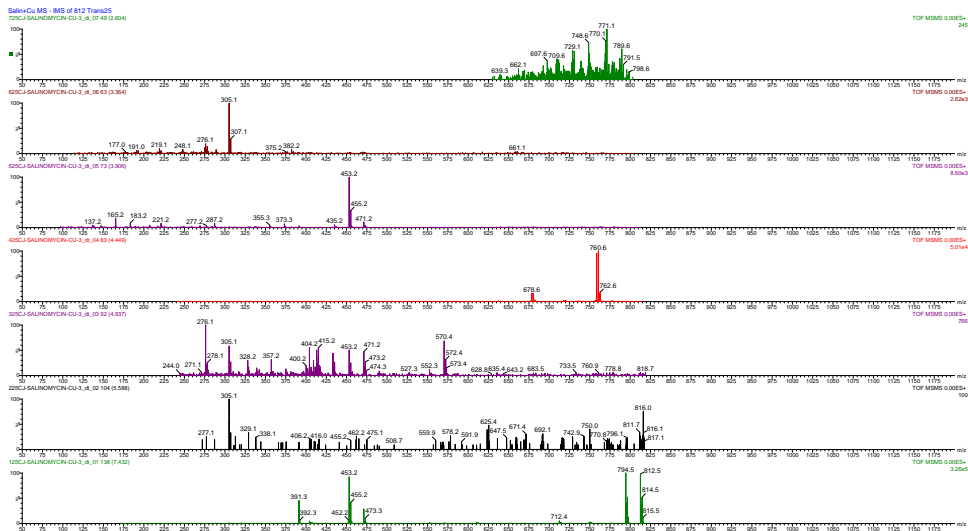
SAL CU Isolation IMS MSMS 812.5



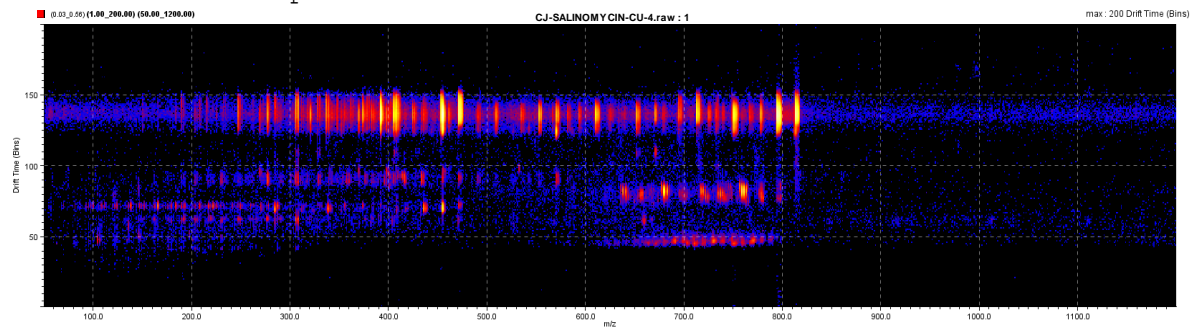
SAL CU Isolation spec



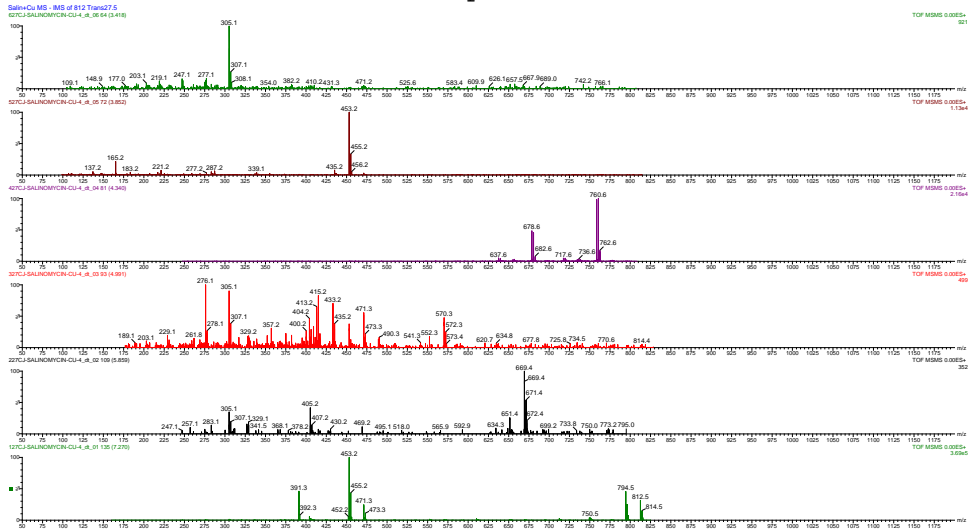
Cu Trans 25 eV drift scope



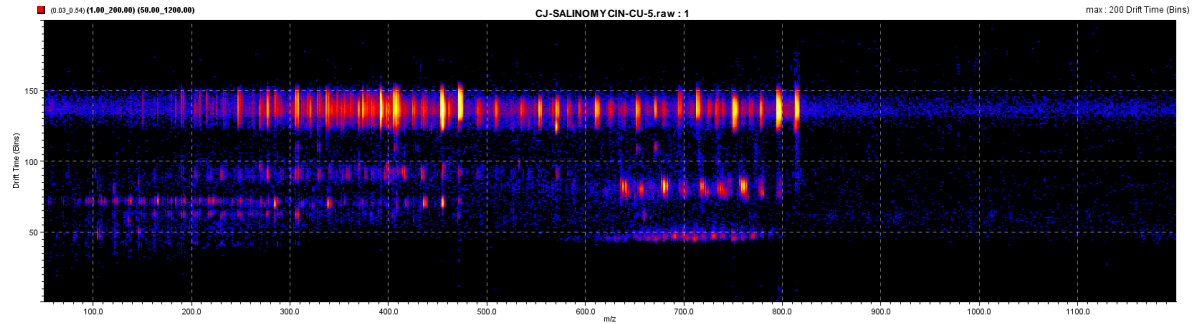
Cu Trans 25 eV spec



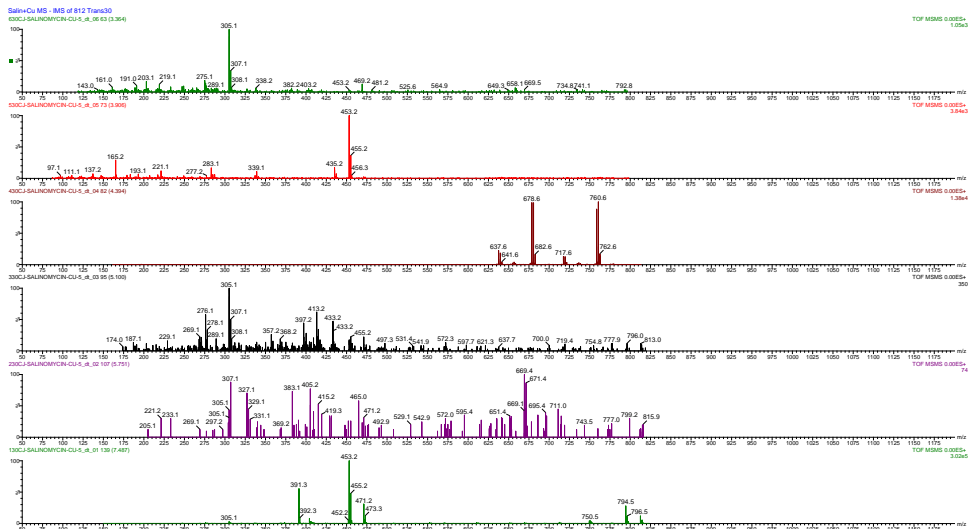
Cu trans 27.5 eV drift scope



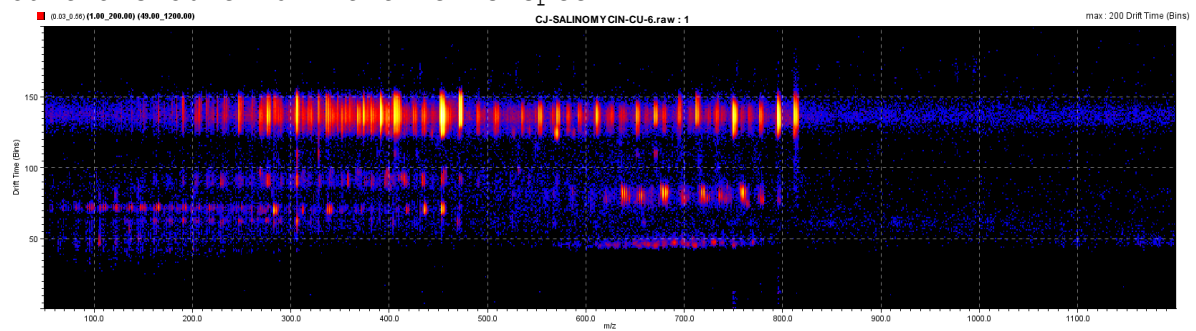
Cu trans 27.5 eV drift vs spec



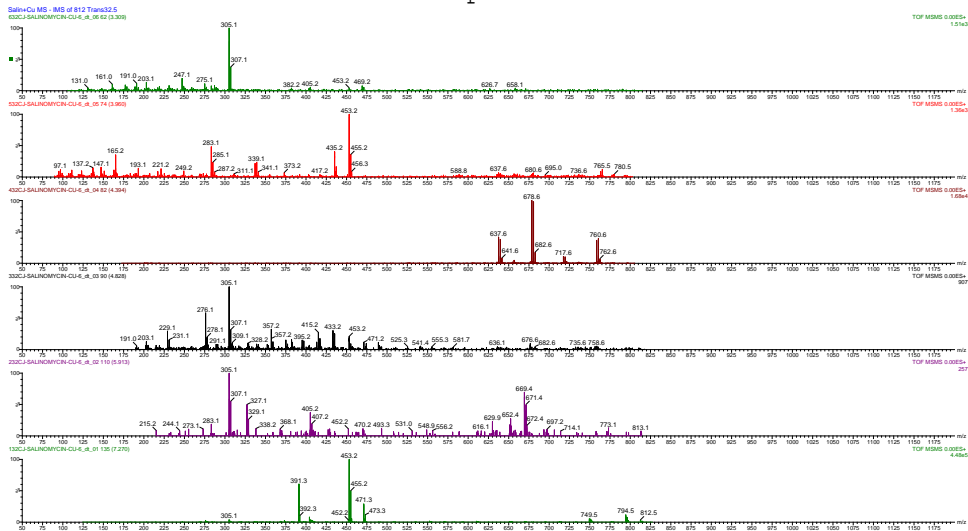
Cu trans 30 eV drift scope



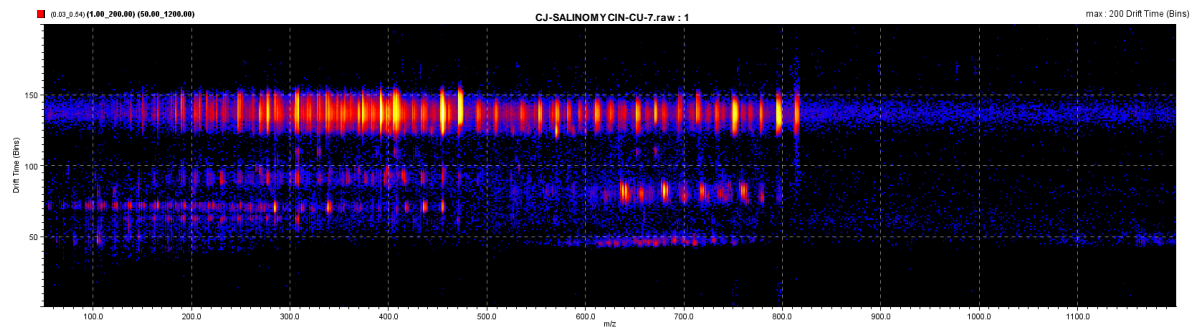
Cu trans 30 eV drift time vs spec



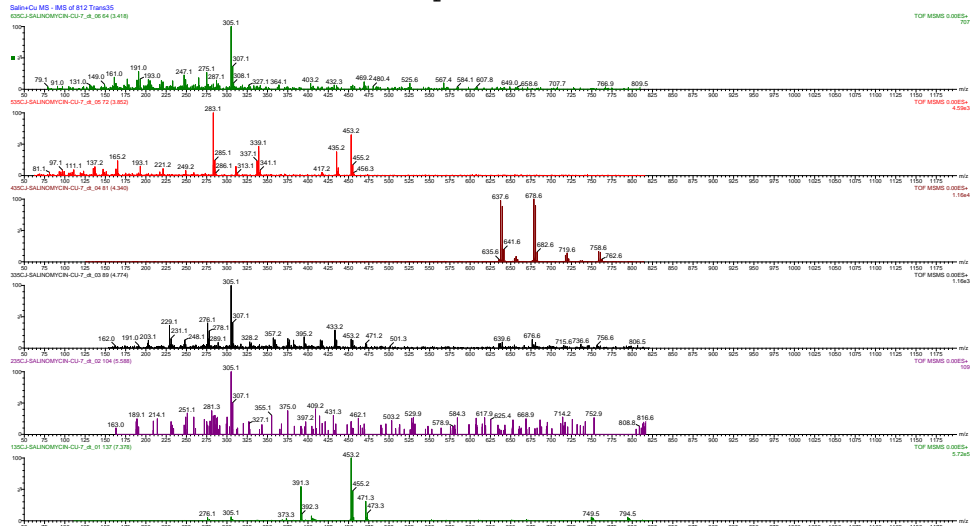
Cu trans 32.5 eV drift scope



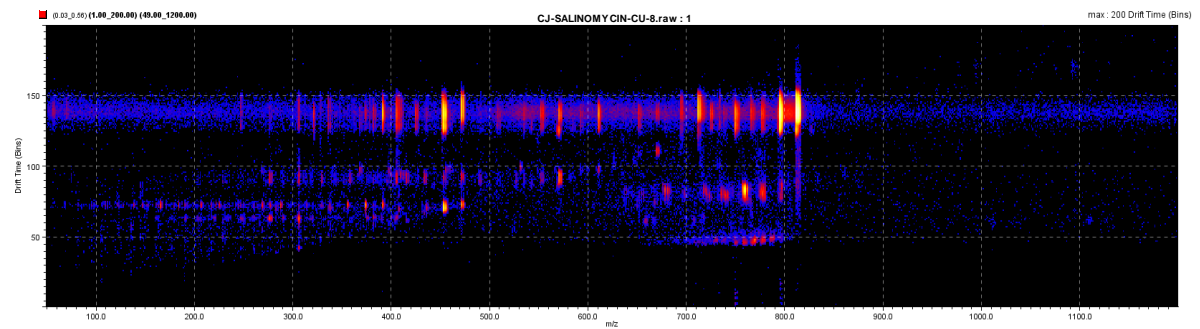
Cu trans 32.5 eV drift time vs spec



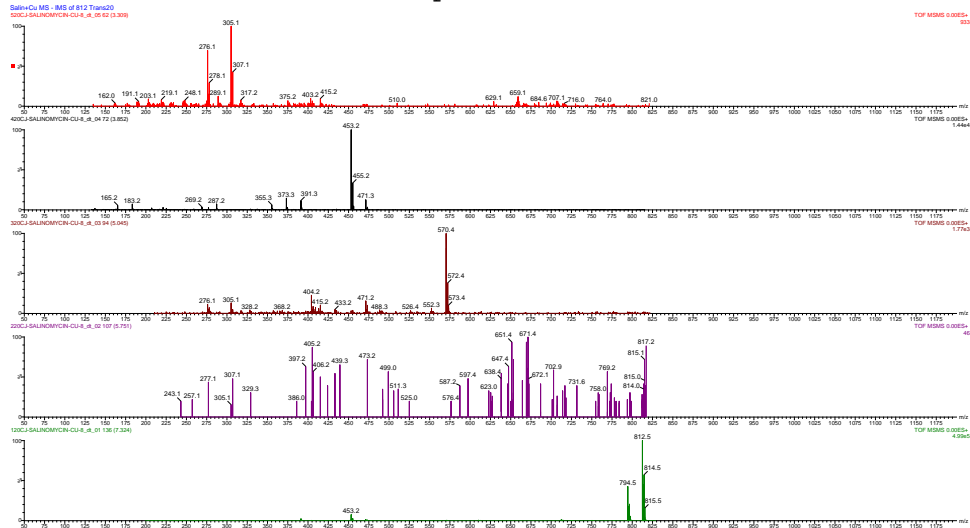
Cu trans 35 eV drift scope



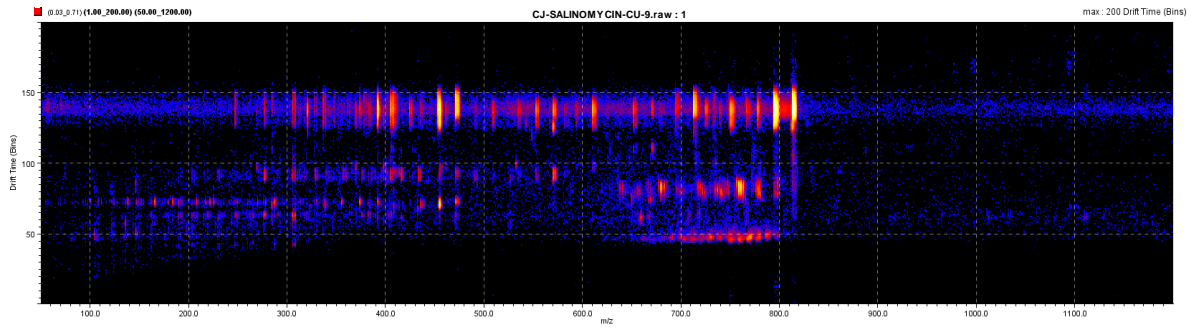
Cu trans 35 eV drift time vs spec



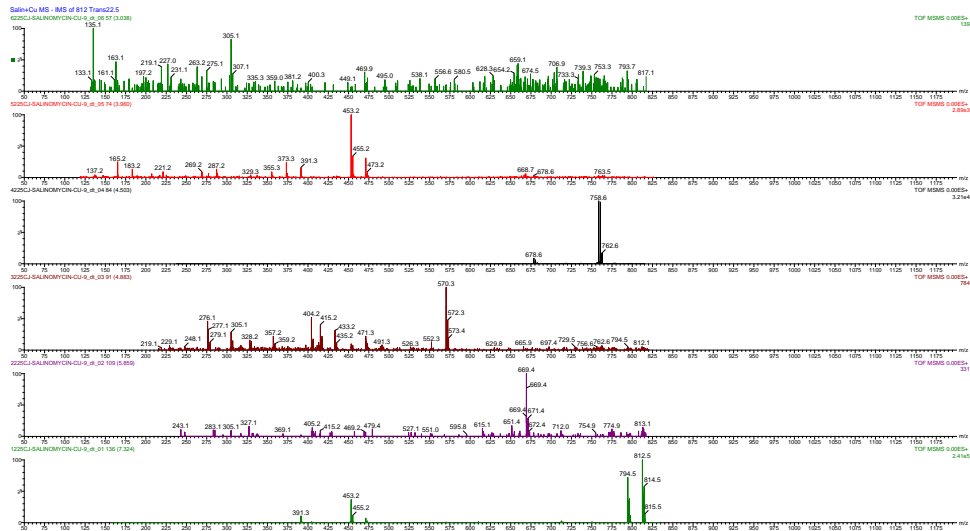
Cu trans 20 eV drift scope



Cu trans 20 eV drift time vs spec



Cu trans 22.5 eV drift scope



Cu trans 22.5 eV drift time vs spec

4. Gaussian

All gaussian data are saved electronically as .log files, hence will not be included here but will be demonstrated if required.