STRATHCLYDE INSTITUTE OF PHARMACY & BIOMEDICAL SCIENCES

University of Strathclyde Science

Using High Resolution Mass Spectrometry to Profile Impurities in Drugs from Different Sources

By

OSAMAH M ALBASSAM

A Thesis Submitted in Fulfillment of the Requirements for the Award of Degree of Doctor of Philosophy in Strathclyde Institute of Pharmacy and Biomedical Sciences at the University of Strathclyde

2017

Declaration

‘I declare that, except where specifically indicated, all the work presented in this report is my own and I am the sole author of all parts.’

Signed: _____

Date: _____

Acknowledgments

I would like to express my gratitude to my supervisor, Dr David G. Watson, for his supported and guided and offering me this opportunity to be trained in his lab.

My thanks also extend for Dr Ian Houson for his assistance and collaboration on this project. In addition, I am thankful to Professor Sandy Gray for the wonderful support and help with NMR.

I would also like to thank Francesca Perciballi and Huaiya Yang from the CMAC to their help and collaboration.

My sincere thanks also go to all of my all my friends and colleagues in SIPBS, especially Alex and Mohammad Alwashih for their cooperation, assistance and helpful discussions.

I would like to thank my parents, I have no doubt this work cannot be done without their prayers. My great thanks to my wife, and sons and daughter “Faisal, Nawaf and Farrah” for their love, support and patience during my studies.

Finally, I wish to thank the Saudi Food and Drug Authority for sponsoring my PhD project

Abstract

Impurities in active pharmaceutical ingredients (APIs) include volatile and involatile organic compounds and inorganic anions and cations. The impurities in an API are acquired during the manufacturing process and may arise from impurities in reagents, side reactions and may, as in the case of ionic impurities associate with the molecule during purification processes. Impurity levels have to be controlled to conform to pharmacopoeial requirements; although this is less apparent in the case of organic and inorganic counterions. In the work reported in this thesis four applications of impurity profiling in APIs and API intermediates were investigated. In chapter 3 the removal of impurities in lipoic acid prior to crystallisation was investigated using high resolution mass spectrometry in combination with chemometric modelling of the data. Several impurities of lipoic acid were characterised by using MS^n and the success of the purification process in removing impurities was verified. In chapter 4 the organic impurities in chlorpheniramine maleate were investigated using high resolution mass spectrometry in combination with chemometric modelling of the data. Several new impurities of chlorpheniramine maleate were characterised by using MS^n and it was demonstrated that it was possible to distinguish between different manufacturers' batches of chlorpheniramine maleate according to their impurity profiles by using high resolution mass spectrometry data in combination with chemometric modelling. In chapter 5 the anionic impurities in different pharmaceutical bases were investigated by using capillary ion chromatography. The different pharmaceutical bases contained a wide range of trace ions as well as the main counter-anion. Aside from the anions

being present as impurities they might have an impact of processing procedures such as crystallisation. In chapter 6 a range of techniques including nuclear magnetic resonance spectroscopy, high resolution LC-MS^n, headspace gas chromatography mass spectrometry, anion and cation chromatography were applied to the analysis of 6-aminopenicillanic acid (6-APA) samples, provided by GSK, which had of low to high clarity values. It was not possible to find any direct association between clarity readings and the various impurity profiles although a number of new impurities in 6-APA were characterised.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

6-APA	6-Aminopenicillanic acid
AIF	All-ions fragmentation
APCI	Atmospheric Pressure Chemical Ionization
API	Active Pharmaceutical Ingredient
API-LC	Atmospheric Pressure Ionization Liquid Chromatography
AUROCC	Area Under the ROC Curve
BP	British Pharmacopoeia
bpt	Boiling point
CE	Capillary Electrophoresis
CID	Collision-induced dissociation
CIE	Capillary Ion Electrophoresis
CMAC	Continuous Manufacturing and Crystallisation Centre
CPAM	Chlorpheniramine
CV	Cross Validation
DAD	Diode-Array Detection
DART	Direct Analysis in Real Time
Dcrit	Critical distance
DESI	Desorption Electrospray

DHLA	Dihydrolipoic acid
DModX	The orthogonal distance of an observation to the model
DSC	Differential Scanning Calorimetry
EI	Electron Impact Ionisation
EIC	Extracted Ion Chromatography
ELSD	Evaporative Light Scattering detection
ESI	Electrospray Ionisation
ESI-MS	Electrospray Ionisation-Mass Spectrometry
FDR	False Discovery Rate
FT	Fourier Transformation
FTICR	Fourier Transform Ion Cyclotron Resonance
FT-IR	Fourier Transform Infrared Spectroscopy
GC	Gas Chromatography
GLP	Good Laboratory Practice
GMP	Good Manufacturing Practice
HCA	Hierarchical Cluster Analysis
HILIC	Hydrophilic Interaction Liquid Chromatography
HPIC	High-Pressure Ion Chromatography
HPLC	High Performance Liquid Chromatography

HRMS	High Resolution Mass Spectrometry
IC	Ion Chromatography
ICH	The International Conference on Harmonisation
IMS	Ion Mobility Spectrometry
LA	Lipoic acid
LC	Liquid Chromatography
LC/MS	Liquid Chromatography-Mass Spectrometry
LOD	Limit of Detection
m/z	Mass to Charge Ratio
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MSA	Methane sulfonic acid
MVA	Multivariate Analysis
NIR	Near Infrared
NMR	Nuclear Magnetic Resonance
NP-LC	Normal Phase Liquid Chromatography
OPLS	Orthogonal Partial Least Squares
OPLS-DA	Orthogonal Partial Least Squares Discriminant Analysis
PCA	Principal Component Analysis

PCs	Principal Components
PLS	Partial least squares projections
PRTs	Pattern recognition techniques
RMSECV	Root-Mean-Square Error of Cross Validation
ROC	Receiver Operator Characteristic
RPLC	Reversed Phase Liquid Chromatography
RF	Response Factor
RRF	Relative Response Factor
RRT	Relative Retention Time
RSD	Relative Standard Deviation
RT	Retention Time
SFC	Supercritical Fluid Chromatography
SIMCA	Soft-Independent Modelling of Class Analogy
TIC	Total Ion Chromatogram
UHPLC	Ultra-High Performance Liquid Chromatography.
UV	Ultraviolet
VIP	Variable Influence on Projection
WHO	World Health Organization
ZIC	Zwitterionic

µS	micro Siemens (conductivity unit)

Terms used interchangeably

Active pharmaceutical ingredients = drug substances

Samples = observations

Subjects = individuals

Putative biomarkers = variables= impurities

Chapter 1

General Introduction

1 Introduction

1.1 Impurities

Pharmaceutical impurities are undesirable chemicals that affect purity of active pharmaceutical ingredients (APIs) (also known as drug substances) or drug products (Ahuja, 2003). These impurities either relate to the synthesis of active pharmaceutical ingredients or may be produced during the formulation process and/or storage - depending on conditions and the period of such storage (ICH-Q3A, R2 (2002)). Pharmaceutical impurities are categorised as organic impurities, inorganic impurities and residual solvents. The acceptable impurity levels of drug substances and products are set and governed by various regulatory bodies e.g. the International Conference on Harmonisation (ICH) guidelines (Eckers et al., 2007).

Organic impurities include starting materials, by-products, intermediates, degradation products, reagents, ligands and catalysts. These impurities may have known or unknown structures, and may be volatile or non-volatile (ICH-Q3A, R2 (2002)). Starting materials and intermediates may still be present in the end product, despite the incorporation of steps to wash them away with solvents during drug manufacture (Roy, 2002).

Degradation products are common impurities in drugs that result from chemical or thermal decomposition of the active ingredients. They can also form during storage, or can be related to interaction of the API with other compounds or contaminants. Degradants can be generated physically such as dimers, trimers and polymorphs of the active compound. ICH-Q6A (R2) defines a degradation product as: ‘a molecule resulting from a chemical change in the drug molecule brought about over time and/or by the action of, e.g., light, temperature, pH, water, or by reaction with an excipient and/or the immediate container/closure system’.

In addition, excipient-related impurities and solvents are also considered to be impurities. Most excipients may be poorly characterized and could contain excessive impurity levels (Kovaleski et al., 2007). Such excipients may contain high levels of

impurities that might become concentrated in the finished drug product (Eckers et al., 2007).

Inorganic impurities are usually known and identified as reagents, ligands and catalysts, heavy metals, inorganic salts and other materials, e.g. filter aids and charcoal. They can also be present as a consequence of the manufacturing processes (Roy, 2002). Even though the chances of encountering reagents, ligands or catalyst impurities are less common, they should also be evaluated during the development of the synthesis process (ICH-Q3A, R2 (2002)). Residual solvents are volatile organic chemicals used in the manufacturing of drug substances or the preparation of drug products. Usually, these solvents are not fully controlled by practical manufacturing techniques (Roy, 2002). Therefore, a comprehensive framework relating to the control of solvents in pharmaceutical materials has been published and updated since 1997 (ICH guideline Q3C, 2005). Inorganic impurities can also result from contaminants that are produced from packaging components and other drug products manufactured in the same facility (Sacre et al., 2011). Inorganic impurities (as elemental impurities) in pharmaceutical materials are controlled by a recent ICH guideline approved in 2014 (ICH Guideline Q3D).

1.2 Impurity profiling

Recently, regulatory authorities have paid particular attention to impurity profiling to improve the quality, safety and efficacy of pharmaceutical products (Görög, 2006). The impurity profile should describe the maximum possible number of the identified and unidentified impurities. It should also estimate the actual amount of different kinds of impurities present in a typical batch of APIs (Ahuja, 2005, ICH-Q3A, R2 (2002)). Impurity profiling is a general term describing the process of acquiring and evaluating data from analytical activities with the purpose of detecting, identifying and/or quantitatively determining impurities whether organic, inorganic or residual solvents in bulk APIs and drug products (Görög, 2006). The impurity profile may include full chromatograms scanned for all peaks - irrespective of the peak area -

related to specified and unspecified impurities, along with enantiomers, conjugates and concomitant species (Satinder Ahuja, 2005).

The process of profiling organic impurities begins with chromatographic methods such as thin-layer chromatography (TLC) and/or high-performance liquid chromatography (HPLC). This is then followed by impurity identification via linking the retention times with those of accessible potential impurity standards (Görög, 2006). HPLC is regularly used to monitor the impurities in drug substances and drug products. Since HPLC retention times can change, there is still doubt about new impurities appearing as a peak with a new retention time. Therefore, characterization of the impurities using hyphenated methods such as liquid chromatography/ mass spectrometry (LC/MS) is preferred, especially when an impurity standard is not available (Edgar C. Nicolas 1998).

1.3 Impurity fingerprints

Impurity fingerprints, the unique impurity profile of samples, are developed using various chromatographic systems such as TLC, gas chromatography (GC) and HPLC. The use of hyphenated chromatographic approaches such as GC/MS, LC/MS and capillary electrophoresis (CE)/MS may demonstrate significantly improved performances with regard to selectivity, precision, separation abilities and retention time stability. An LC/MS fingerprint is also a powerful tool for differentiating the same product from different manufacturers (Xiaohui et al., 2006).

During the manufacturing process, it is crucial to have the capability to accurately measure an API and drug product's impurity profile. Providing a specific, accurate, reliable assay for raw materials and drug product tests is also now one of the requirements of many of the world's drug regulatory authorities (Plumb et al., 2008). Many articles have been published recently describing methods and approaches for the isolation and identification of process-related impurities and degradation products of pharmaceutical substances using MS, nuclear magnetic resonance (NMR), HPLC, Fourier transform ion cyclotron resonance (FTICR) MS, and tandem

mass spectrometry (MS/MS). These approaches can either be stand alone or be combined with other techniques (Roy, 2002). In impurity profiling studies several analytical techniques are applied to detect trace organic impurities in pharmaceutical formulations, however, it is widely agreed that HPLC combined with diode-array detection (DAD) or MS are the most commonly used techniques (Eckers et al., 2007, Sacre et al., 2011).

Many of the approaches that have been reported for impurity fingerprinting use off-line spectrometric techniques such as NMR, near infrared spectroscopy (NIR) and Fourier transform infrared spectroscopy (FT-IR). These techniques were widely tested in the field of quality control in several studies; however, the major deficiency was the complexity of the interpretation of the data derived from these techniques especially they were when used to detect very low quantities of impurities. For this reason, chromatographic techniques are the most preferable methods for describing impurity fingerprints (Sacre et al., 2011).

1.4 Counterfeit medicines

Counterfeit medicines can be a real threat to public health. Health problems arising from counterfeit medicines are a result of the variability of product quality in terms of qualitative, quantitative and impurity profiles (Sacre et al., 2011). A counterfeit medicine is defined by the World Health Organization (WHO, 1999) as 'one which is deliberately and fraudulently mislabelled with respect to identity and/or source'. Brand and generic products can be copied to produce counterfeit products that contain correct or incorrect ingredients, missing active ingredients and/or unequal active ingredients and which may be contained in fake packaging.

The counterfeit medicine's risk is commonly based on the lack of manufacturing process control and inexpert workers, which result in an unstable or unsuitable amount of active substance or incompatible excipients. These drugs usually contain a high concentration of impurities that can be detected even when the expected amount of API is present. These impurities come from low-quality starting materials,

cross-contamination, inadequate manufacturing processes and/or storage conditions (Sacre et al., 2011, Fernandez et al., 2006). In counterfeit product manufacturing, materials from unknown sources are used without controlled processes or adherence to Good Manufacturing Practice (GMP) guidelines. As a result, the purity of APIs and excipient materials may not meet the pharmacopoeial requirements and may change under ambient storage conditions (Sacre et al., 2011). Counterfeit products may also contain different active materials to simulate the organoleptic characteristics and/or the symptom-relief effects of the original product. They can be waste powder materials from different pharmaceutical production lines (Fernandez et al., 2006). All these various inconsistencies in the manufacturing practice make counterfeit products life-threatening (Sacre et al., 2011).

1.5 Techniques used for impurity profiling

1.5.1 Reversed phase liquid chromatography

Reversed phase liquid chromatography (RP-LC) is the first choice for a broad range of samples because it is typically more suitable and rugged than other LC techniques and gives satisfactory separation results (Yoshida, 2004). In the stationary phase, the layer of silica is bonded to and covered by hydrocarbon chains that provide a hydrophobic surface, e.g. C8, C18. One of the most important features that enhances the capacity factor of the system (an indicator of its resolving ability) is the increase in surface area provided by the porous silica particle that allows for longer contact between the stationary phase and solute molecules (Moldoveanu and David, 2013).
In addition, selectivity can be enhanced by the appropriate selection and concentration of the organic component in the mobile phase. The selection of mobile phase is crucial in the case of impurity profiling of an API. Degradation products may be eluting at the same retention time as the drug substance, therefore, the mobile phase system may consist of up to three or four of organic solvents in optimised proportions in order to achieve the desired chromatographic resolution. Capacity

factor, according to Dolan's rule, can be increased by two or three fold when the concentration of acetonitrile and methanol is reduced by 10%, respectively (Watson, 2012).

The main drawback of RP-LC is the lack of suitable retention of polar compounds. In this case, normal phase liquid chromatography (NP-LC) is usually employed to separate polar compounds. It is a useful separation mode because it offers an alternative selectivity to RPLC. However, it is also difficult to dissolve hydrophilic compounds in the organic mobile phases used in NP-LC (Yoshida, 2004). This challenge can be overcome through the use of hydrophilic interaction liquid chromatography (HILIC) discussed in the next section.

1.5.2 Hydrophilic interaction liquid chromatography

Hydrophilic interaction liquid chromatography (HILIC) offers an alternative retention mechanism with a high capability to separate small polar compounds using normal phase surface chemistry (Buszewski and Noga, 2012, Wang et al., 2005). HILIC is superior to regular NP-LC and RP-LC because it offers the ability to analyse mixtures in complex systems that elute close to the void volume in RP chromatography (Buszewski and Noga, 2012). This technique allows for the effective separation and characterization of polar samples that cannot be analysed by RP-LC (Yoshida, 2004). The partially aqueous mobile phases used in HILIC solubilise polar compounds, which resolves the problem of poor solubility that often occurs when NP-LC is used (Buszewski and Noga, 2012). The theoretical explanation of molecular retention in HILIC is demonstrated in several publications. The principal mechanism is the partitioning of the analytes between the stationary phase which is solvated with a partially immobilized layer of water and the hydrophobic mobile phase. Furthermore, hydrogen bonds and electrostatic interactions properties can be shared in the HILIC retention (Chauve et al., 2010). HILIC is a modified form of NP-LC that has more complex separation mechanisms and works by using a traditional polar stationary column such as amino, silica, or cyano, among others; making it appear like an NP-LC

stationary phase with a RP-LC mobile phase with low water content. It is also used to separate charged substances in ion chromatography (IC) (Buszewski and Noga, 2012).

Thus in HILIC separation, polar compounds are retained on the stationary phase with an organic mobile phase containing a low concentration of water. Different interactive forces govern selectivity in HILIC mode including hydrogen bonding, dipole–dipole interactions and ion-exchange and ionic repulsion interactions. The degree to which the analyte is retained depends on the number of polar groups, conformation and solubility of the sample in the mobile phase and its acidity or basicity (Yoshida, 2004). HILIC can be applied with a gradient mobile phase, starting with a low polarity organic solvent, with more polar compounds being eluted by increasing the polar (aqueous) proportion in the mobile phase. At the start of the gradient, the highly organic mobile phase provides good sensitivity and allows enhanced on-column retention for polar ionic compounds. HILIC does not require the use of ion pair reagents and it is suitable for hyphenating with MS, particularly in the electrospray ionization (ESI) mode.

HILIC is considered to be the technique of choice for separating uncharged hydrophilic and amphiphilic molecules that have high polarity: compounds which typically have poor retention in RP-LC and lack the charges required for separation by ion-exchange chromatography. In addition to the efficient separation of carbohydrates, peptides and polar pharmaceuticals, HILIC is also used to solve many problems previously found in separation-based analysis, including the isolation of small organic acids, basic drugs, and neutral substances (Buszewski and Noga, 2012).

HILIC can complement the analysis of weakly polar, or non-polar, compounds with RP-LC. The significant difference in the selectivity between HILIC and RP-LC may create the best results when the methods are combined (Yoshida, 2004). RP-LC and HILIC are widely used techniques as orthogonal methods for liquid-based LC separations. Orthogonality is a procedure for covering separations of complex mixtures in a two-dimensional (2D) space. The benefits of the RP-LC HILIC orthogonal combination technique can be employed by injecting the same sample in both RP-LC

and HILIC modes. The collected data from the two techniques offer more comprehensive information on the entire sample than just RPLC or HILIC alone. This information may present a successful identification by comparing the data with an available database or the fingerprint analysis of a similar sample (Buszewski and Noga, 2012). Multidimensional chromatography using orthogonal methods such as HILIC and RP-LC can create powerful separations. HILIC is optimal for orthogonal HPLC method development because its selectivity can be enhanced through the different retention mechanisms it employs in comparison to RP-LC (Buszewski and Noga, 2012, Wang et al., 2005).

1.5.3 Supercritical fluid chromatography

Supercritical fluid chromatography (SFC) is an NP method and typically uses carbon dioxide in the mobile phase. It provides high selectivity separations of isomers, enantiomers and product-related compounds. In contrast to RP and NP methods, the mobile phase in SFC can be utilized with any stationary phase, i.e., polar and non-polar. Although SFC is applied broadly as a support method in the field of drug discovery, it has not been used extensively in laboratories conducting analytical testing to regulated requirements (such as GLP, or GMP) mainly because of noise/sensitivity and reproducibility concerns. The relative changes in the noise of the gradient baseline and slope are a major drawback of SFC. Accordingly, this large difference in noise levels produces an additional degree of difficulty in method development and complex experimental conditions. Moreover, method development for large compounds is still at the investigational stage (Alexander et al., 2012).

1.5.4 Ultraviolet spectrometry

Ultraviolet (UV) detection may be applied with gradient HPLC as a common method for screening impurities in APIs. Unfortunately, this method can only identify a few peaks since impurity reference standards may not be present. Moreover, some impurities do not have chromophores and thus cannot absorb the UV wavelengths

this method relies on for detection. Impurities are usually assigned and traced by the retention time; however, the peaks cannot be accurately measured when they are overlapped with other peaks due to a shift in the retention time or if there are other closely eluting peaks nearby (Edgar C. Nicolas 1998). HPLC with UV detection is commonly used for detecting only known impurities (Plumb et al., 2008). Although HPLC with a UV diode array is a great tool to screen impurities in drug materials, this tool may need support from other techniques such as MS to achieve analysis conformity with peak levels less than 0.05% (Edgar C. Nicolas 1998).

1.5.5 Near Infrared and Raman Spectroscopy

NIR is widely used in the area of identification of raw materials to control excipients, API and drug products such as the identification of different types of sugar and celluloses (Roggo et al., 2007). Kreft et al. applied NIR spectroscopy with Soft-Independent Modelling of Class Analogy (SIMCA) to the determination of excipients to differentiate between povidone types (Kreft et al., 1999) . NIR and Raman spectroscopy have been developed for characterizing miokamycin polymorphism (Blanco et al., 2000).

NIR and Raman spectroscopy are fast and non-destructive techniques and do not require any sample preparation. They are extensively applied in real-time monitoring for in-process testing and assessment of finished products (De Beer et al., 2011). Many publications have reported the application of Raman and NIR spectroscopy with chemometrics in the field of in-process control of pharmaceutical products. For example, Wikstrom et al. controlled the transformation of theophylline anhydrous to theophylline monohydrate in wet granulation by Raman spectroscopy for in-line monitoring with Principal Component Analysis (PCA) (Wikstrom et al., 2005). El-Hagrasy et al. applied NIR with pattern recognition models (SIMCA) for samples having different processing conditions such as humidity, blender speed and component concentration to detect the blend homogeneity (El-Hagrasy et al., 2006). De Beer et al. applied Raman spectroscopy in combination with SIMCA to detect the in-line endpoint and real-time blending homogeneity of a multicomponent powder

(De Beer et al., 2008). Romero-Torres et al. and De Beer et al. developed Raman spectroscopy for in-line freeze-drying process monitoring of the lyophilisation of a simple mannitol solution (De Beer et al., 2008, Romero-Torres et al., 2007). However, the ability of these techniques to detect impurities at levels < 1% w/w is limited.

1.5.6 Differential scanning calorimetry

Differential scanning calorimetry (DSC) has been used in the pharmaceutical industry for several applications such as characterisation of compounds in terms of polymorphism, monitoring stability (Giron and GoMbronn, 1995), and purity determination of organic compounds (Mathkar et al., 2009).

Calorimetric purity can be attained by the principle of depression in the melting point of an analyte that is related to the effect of eutectic impurities. DSC can provide a rapid analysis for the purity of a material (within an hour) and only requires small quantities of the sample (1-2 mg). The DSC curve can be used for determining sample purity without any need to use any reference standard for either the sample or its impurities. Nevertheless, this method is not sufficient to use without a chromatographic technique for detecting different types and levels of impurities. In addition, validation of this method has some complexity and takes a long time (Giron and GoMbronn, 1995). The DSC purity method estimates the purity of the sample by (%mole/mole) according to Van't Hoff law, but this supposes that there is no difference in molecule weight between the analyte and its impurities. For this reason, DSC purity may only be valid for small molecular weight compounds (Drozdzewska et al., 2007).

Mathkar et al. applied DSC for detecting the purity of 16 organic reference standards. The results were compared with other data that were obtained from different analytical techniques such HPLC. The conclusion was the purity achieved by DSC is comparable to LC results for organic compounds with purity above or equal to 98% (Mathkar et al., 2009). Drozdzewska et al. employed DSC for measuring 34 certified reference standards of polycyclic aromatic hydrocarbons (PAHs). The results were shown to be consistent with GC and LC results (Drozdzewska et al., 2007). In this

study, the DSC results were compared with reference standard certificates to calculate the method bias.

1.5.7 Mass spectrometry

A gradient LC coupled with UV, DAD and MS detection is a common method for identifying impurities and degradation products. The response of a UV detector will measure impurities down to approximately ≥0.05% of the UV absorbance of the main band (normally the active drug ingredient), however the LC-UV detection method may give biased results since only impurities with UV absorptive properties can be measured. The use of an MS detector can increase the capability to recognize different forms of impurities (Freed et al., 2004).

MS hyphenated with LC has proved to be the most useful technique for obtaining the maximum data for an impurity profile. GC/MS has recently been used in fingerprinting, although it has not been extensively investigated (Sacre et al., 2011). MS coupled with HPLC is currently used as a general method for identifying or elucidating the structure of drug impurities since it does not depend on the impurities being volatile. For these reasons, HPLC-MS is applied as the most widely used tool for impurity detection rather than GC/MS (Görög, 2006).

Atmospheric pressure ionisation (API) LC/MS is a selective and sensitive method that offers qualitative and quantitative analyses of impurities and degradation products in drug substances and drug products. API-LC/MS replaced direct probe electron impact and chemical ionisation MS for compound elucidation (Freed et al., 2004). Nowadays, electron impact ionization, EI-MS, is normally used in GC/MS for identifying volatile or semi-volatile unknown compounds making use of EI-MS spectral databases where the spectra do not depend on the MS instrument (Li et al., 2009). Using MS for accurate quantification of unknown impurities could be inconvenient because some structures may not be detected with a specific ionization mode. Nevertheless, LC/MS provides high sensitivity combined with selectivity when detecting individual molecular and fragment ions. For these reasons, LC/MS has been

identified as an essential part of determining fully characterized impurity profiles (Baertschi, 2006).

MS detectors are commonly considered to be universal, but the abundance (the strength of the ion signal) and it's relation to the weight of the ion being measured is highly dependent on the ionization mode (e.g., positive or negative ionization, electrospray, atmospheric pressure chemical ionization (APCI) or electron ionization) and the response of the compound to ionization (Baertschi, 2006).

1.5.8 Tandem Mass Spectrometry

Liquid chromatography hyphenated with tandem mass spectrometry (LC-MS/MS) has the ability to track unidentified and slight impurity peaks as well as major known impurities. In addition, LC-MS^n is used in many fingerprinting protocols because of the advantage that it provides high-confidence structures from unknown compounds (Li et al., 2009). In some cases, it may be important to compare LC-MS/MS fingerprinting to confirm data from a UV diode array detector (Nicolas and Scholz, 1998). LC-MS/MS offer faster identifier of impurities, especially those with different isomers. For example, a high-resolution MS may not be able to provide significant confirmation, whereas a unit resolution ion trap MS can produce a high level of confidence in a structure based on LC/MS^n molecular fingerprinting. With LC/MS^n fingerprinting, the need to have a reference material of the impurity may be unnecessary for identification. LC/MS^n molecular fingerprinting needs to be a close, but does not require an identical, match with an accessible database to predict a rational structure when combined with other information such as reaction mechanisms or degradation pathways (Li et al., 2009).

MS/MS fingerprinting may propose high confidence structures of unknown impurity. This technique can be time-saving and may provide impurity information relating to changes in a manufacturing process. It can also confirm more peak identities based on molecular weights and set MS/MS fragments over chromatographic fingerprints (Li et al., 2009). Furthermore, MS/MS fingerprinting can distinguish between two

constitutional isomer impurities by roughly detecting the impurities' structures (Nicolas and Scholz, 1998).

Li et al. demonstrated LC/MS^n analysis (n is typically 1 to 3) as an important strategy for the rapid elucidation of the structures of degradants and degradation pathways. All data provided from LC/MS^n such as UV profiles, molecular weights, and multiple fragmentation patterns of unknown products were initially analysed. This unique data, which were provided from LC/MS^n for each chemical structure, were collected to form a fingerprint for the drug product. The database generated from LC/MS^n molecular fingerprinting was transferred between different types of LC/MS instruments. Ion trap mass spectrometers are considered to be one of the most suitable instruments for LC/MS^n molecular fingerprinting and the easiest when transferring an applicable database when using the same mass spectrometer in another facility (Li et al., 2009).

Fernandez et al. reported Direct Analysis in Real Time (DART) as a useful technique for the rapid screening of counterfeit drugs. They predicted that both DART and desorption electrospray (DESI) will have a great influence in many scientific fields, from drug quality control, screening and discovery to biological applications. DART can be accompanied by MS/MS experiments to yield more valid identifications of unknowns because of the accurate mass and isotopic cluster intensity measurements, particularly in the case of when chromatographic separation is absent (Fernandez et al., 2006).

At present, impurity profiling frequently applies to investigations of counterfeit medicines; this is usually performed by using LC/MS, particularly MS/MS. The measurement of small or minor peaks of impurities in a sample can identify the route of synthesis used by a manufacturer of the API. This data can then be compared with the innovator's process to reveal any illegal product copies (Plumb et al., 2008).

1.5.9 Orbitrap mass spectrometry

Fragmentation product mass analyser technology is a powerful tool that adds extra speed, mass accuracy and sensitivity to mass spectrometric techniques. The LTQ Orbitrap mass spectrometer (Figure 1.1) provides sophisticated analytical performance that includes routine compound analysis and trace level component detection in complex mixtures (Zubarev and Makarov, 2013). The hyphenated technique with collision-induced-dissociation (CID) for MS/MS and/or MS^n is capable of identifying the accurate masses of fragments. This type of data, combined with other techniques (such as NMR), permits determination of the chemical structures of impurities (Freed et al., 2004).

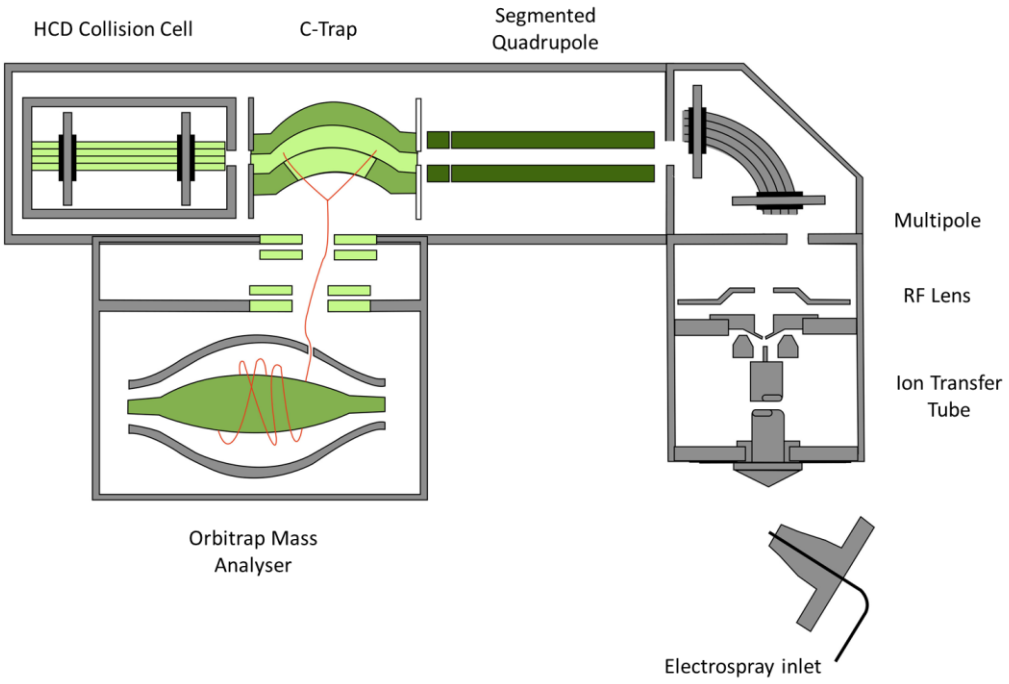


Figure 1.1 scheme of Orbitrap Mass Spectrometer (Michalski et al., 2011).
This instrument includes an S-lens, a mass selective quadrupole, and HCD collision cell directly interfaced to the C-trap.

1.5.10 Ion Chromatography - Conductivity Detector

Ion chromatography became the main technique for the separation of organic and inorganic ions (Weiss, 2004). The IC technique functions by preparing a chromatographic system where the analyte ions and the eluent have same charge, and bind to a stationary phase which has the opposite charge on the surface determined by selecting the stationary phase functional groups. The ions coming from the analytes and eluent face competition to bind on the oppositely charged stationary phase. There are two types of IC column: one is a cation exchange column which has a negative charge surface for cation separation, and anionic exchange column has a positive surface for anion separation (Conboy et al., 1990, Bhattacharyya and Rohrer, 2012).

IC hyphenated with a suppressed conductivity detector (Figure 1.2) has been used widely in pharmaceutical analysis since it offers superior sensitivity and stability compared to a normal conductivity detector. It provides more sensitivity since the eluent used can be changed to water or carbonic acid which have low conductivity. This decreases the background conductivity of the eluent which leads to signal enhancement of the components, thus offering more efficiency and robustness (Haddad et al., 2003). The baseline and drift are significantly reduced which allows detection to reach levels of low ppm relative to the drug substance (Rocheleau, 2008)(Haddad et al., 2003).

Ion chromatography with conductivity detection is often used for the separation of counterions, including inorganic ions and acidic and basic organic compounds (Rocheleau, 2008). Since the counterions play a role in enhancing the product solubility, stability and decrease the hygroscopicity and toxicity (Elder et al., 2010) .

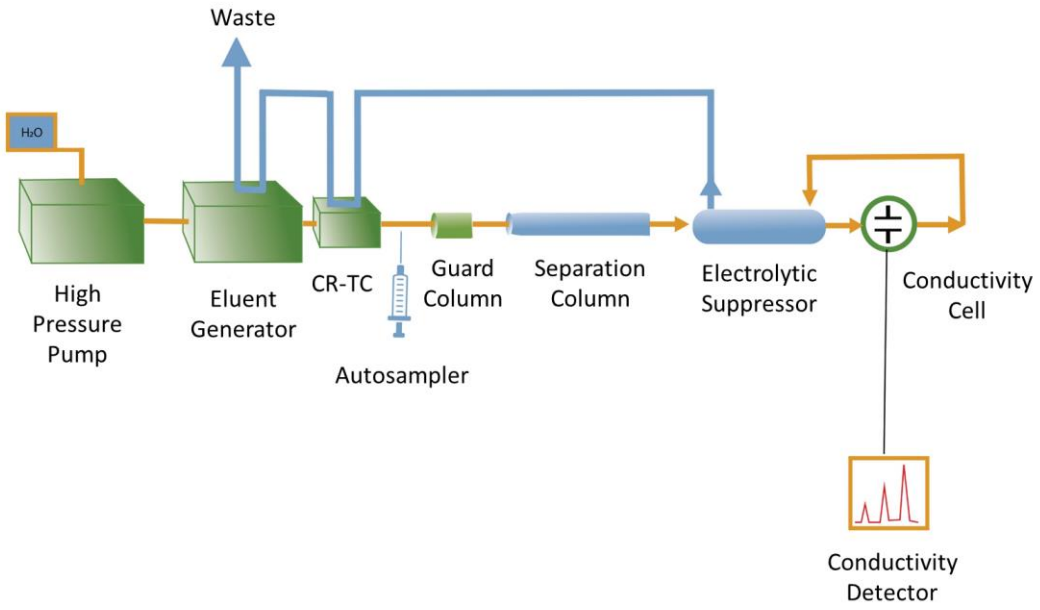


Figure 1.2 system schematic of Reagent-Free Ion Chromatography with Eluent Generation (RFIC-EG) system
This instrument consists of a high pressure pump unit, eluent generator unit (EG), trap column (CR-TC), autosampler and guard and separation columns, electrolyte suppressor and detector compartment.

1.5.11 Headspace Gas chromatography- Mass Spectrometry

Gas chromatography (GC) is the most appropriate method for the analysis of solvents. It has the capability to separate and identify component solvents at low detection limits especially when used with MS detector. Headspace -GC/MS is widely used in pharmaceutical analysis for testing residual solvents (B'Hymer, 2003, Grodowska and Parczewski, 2010). In addition, the residual solvent profile was used to recognise the samples of the same API that originate from different sources (Mulligan et al., 1996).

Headspace analysis is a vapour phase extraction which works on the basis of partitioning the sample to two phases: a volatile and non- volatile. The upper vapour-phase holds the components which are injected into the GC and are transported

through the column by the flow of an inert gaseous mobile phase such helium (Snow and Slack, 2002).

1.5.12 Ion Mobility Spectrometry

The ion mobility spectrometry (IMS) technique is a gas phase separation that analyses volatile and semi-volatile molecules depending on analytes ionized under a weak electric field at atmospheric pressure. IMS can quickly identify and quantify at nanogram concentrations with high levels of accuracy and precision. It is considered the best method for cleaning verification in pharmaceutical manufacturing and quality control. Combining IMS with MS^n presents a second analytical dimension with a high ability for detecting complex compounds by providing 3 to 5-fold more sensitivity and selectivity for polymer separation than one-stage, low-resolution MS. IMS is also robust and easier to operate than MS and offers the same level of sensitivity as a flame ionization detector in GC but better sensitivity than a UV–visible absorption detector in LC. IMS-MS also has ability to recognize size, shape, and charge for the analysis of multi-compounds with a similar molecular weight (Armenta et al., 2011).

The main disadvantage of this technique is that IMS lacks competitive ionization and this may lead to full ion suppression of compounds and charge exchange reactions. Moreover, the commercial IMS instruments have a limited efficiency of separation that use prior sample clean-up and a selective detector and/or chromatographic separation (Armenta et al., 2011).

1.5.13 Nuclear Magnetic Resonance spectroscopy

NMR spectroscopy can be a powerful detector of impurities. Degradation products can be isolated and introduced into the NMR spectroscopy through tubes, direct flow or via HPLC. Data interpretation of capillary or traditional HPLC-NMR has been difficult when characterizing low-levels of impurities including the parent molecule as a result of poor signal-to-noise from a low capacity column and/or disparity of the peak volume with that of the flow probe. This problem was resolved when capillary

coil or cryogenically cooled NMR probes were used. Although NMR applications are considered significant qualitative tools, they are not commonly used as universal or quantitative detectors. NMR has been run to confirm and/or compare the outcomes of other quantitative detectors; it was found that isolating the impurities was required to use NMR as an accurate quantitative tool (Baertschi, 2006).

In addition, NMR is a well-known technique and can provide a structure elucidation of an unknown compound. The structure identification of trace impurities may be difficult to detect in conventional NMR due to insufficient sensitivity. In this case, a high-sensitivity technique such as MS can provide valuable information (Szántay et al., 2006). The structure confirmation can be validated by using dimensional (1D) and (2D) NMR experiments (Li, Wang, et al. 2009). NMR is usually applied offline, but hyphenated techniques such as HPLC/NMR or HPLC/NMR/MS are increasing significantly (Görög, 2006).

The quantitative analysis of low levels of impurities and degradation products requires careful selection of the chromatographic methods. Furthermore, full characterization of impurity and drug related substance profiles is achieved by using various separation methods to adequately offer orthogonal assessment (Wade Demond, 2000).

NMR spectroscopy provides valuable information for impurity profiling. Impurity detection is differentiated by dividing the NMR spectrum of samples into small bins. These data can then be managed with an appropriate pattern-recognition method such as PCA. NMR also serves as a quantitative tool due to recent, major developments in sensitivity, accuracy and methodology. NMR also has an immense ability for structure elucidation of unknowns (Szántay et al., 2006).

1.6 Multivariate Analysis

LC/MS and NMR are techniques applied frequently for identifying metabolites, impurities and degradation products in the pharmaceutical field. Even with the massive benefits of NMR and LC/MS, the informed identification of degradation

products and impurities is a challenging and time-consuming process. The interpretation of data coming from full LC/MS scans requires an immense time commitment and effort aimed at interpreting the outcomes of impurity identification. In addition, these techniques generate a large amount of data that can be managed with software programmes using spectral and chromatographic data to identify degradation products and impurities. For that reason, and to hasten this process, some search programs, such as MetaboLynx^{TM} and Advanced Chemistry Development, use chromatographic integration and spectral thresholds to help detect and identify impurities, especially those hidden in the noise level (Freed et al., 2004).

The multivariate analysis contains two stages. The unsupervised technique step using pattern recognition is applied with the aim of the visualisation of the data and to detect any outliers of the observations or variables. The second step is the supervised technique for discrimination between groups, identifying the important variables that make the change in the model. The predictive ability can be assured by model validation.

1.6.1 Pre-treatment of data

Multivariate analysis approaches depend on the pre-processing of data. It is a prerequisite to use mean-centering to improve the model interpretability and scaling to reduce the effect of the abundant variables which may drive the separation and/or data transformation (e.g., log transformation) to get a normal distribution.

Mean–centering is achieved by subtracting the average intensity of each variable from the data which leads to conversion of all intensities to fluctuate around zero. The scaling method can be performed by dividing each variable by a scaling factor. The standard deviation is used as a scaling factor of unit variance and the square root of the standard deviation for Pareto scaling (van den Berg et al., 2006).

Pareto scaling can provide large dynamic ranges to ensure that the variations among peaks are treated depending on the relative alteration rather than the influence of

the abundance of peaks. In addition, Pareto scaling reduces the risk of losing abundances which are close to the limit if detection (LOD). Therefore, the inflated noise will reduce and will not have an influence on the model. This cannot be acquired by the unit variance scaling or mean-centering only (Wheelock and Wheelock, 2013).

Log transformation is often applied. It is usually used to correct the random and skewed distributions to normal and systematic ones. It converts the data to additive relations from multiplicative relations (van den Berg et al., 2006). Log transformation is useful in case some variables have an extreme value. In addition, it is a good tool that assists in removing outliers (Eriksson et al., 2013e).

1.6.2 Principal components analysis

Principal components analysis (PCA) allows for the fast classification of samples consistent with their impurity profile. PCA is an investigative statistical method for data analysis and processing in many applications in pharmaceutical analysis. It is a common technique for expressing patterns in multi-dimensional data that permit new variables (X) to be characterized in an orthogonal manner.

Assessing the Ultra-high performance liquid chromatography (UHPLC)/MS exact mass data with PCA provides easy detection and identification of impurities. In addition, this technique can identify the variances between different batches of pharmaceutical products including samples produced by one manufacturer. This approach allows for monitoring of new samples in terms of the impurity profile because known and unknown impurities are compared using the impurity profile of other samples (Plumb, 2008). In addition, multivariate data analysis software such as SIMCA classifies each cluster of samples separately using a limited number of principal components acquired from PCA (Sacre et al., 2011).

PCA is an unsupervised model used to orthogonally transform huge sets of interrelated data to a low dimensionality data set that retains the variation in the original data set. PCA consists of principal components (PCs) usually 2 – 5 which are a set of values of linear uncorrelated variables. PCA provides a visualisation of the

relation between observations and variables, in addition, it identifies trends and potential outliers (Jolliffe, 2002, Eriksson et al., 2013d).

1.6.3 Detection of outliers

Theoretically, outliers are observations that are at an extreme from or do not fit with the remainder of the data. Those observations have a severe and disproportionate outcome on the PCA model. It is easy to detect the strong and moderate outliers by applying Hotellings' T2 plot and DModX plot, respectively. The value of removing and diminishing the effect of the outliers is to make the analysis of the data set more robust (Jolliffe, 2002, Eriksson et al., 2013d). Hotellings' T^2 is used as a measure of the distance of the observations to the origin of the model plane. Observations that have T^2 range > 95% of confidence limit are considered as suspect, and > 99% they are considered as serious outliers which may drive the model in detrimental way (Figure 1.3A).

The orthogonal distance of an observation to the model (DModX) is a second type of the control limit that can be calculated for each observation. It is proportional to residual standard deviation (RSD). DmodX is used to detect the moderate outliers that are larger than the critical distance (Dcrit) at the 0.05 level, and these are considered to not fit the model well (Figure 1.3). The observations which exceed the double distance of the Dcrit(0.05) are considered as strong outliers (Eriksson et al., 2013d) (Figure 3B).

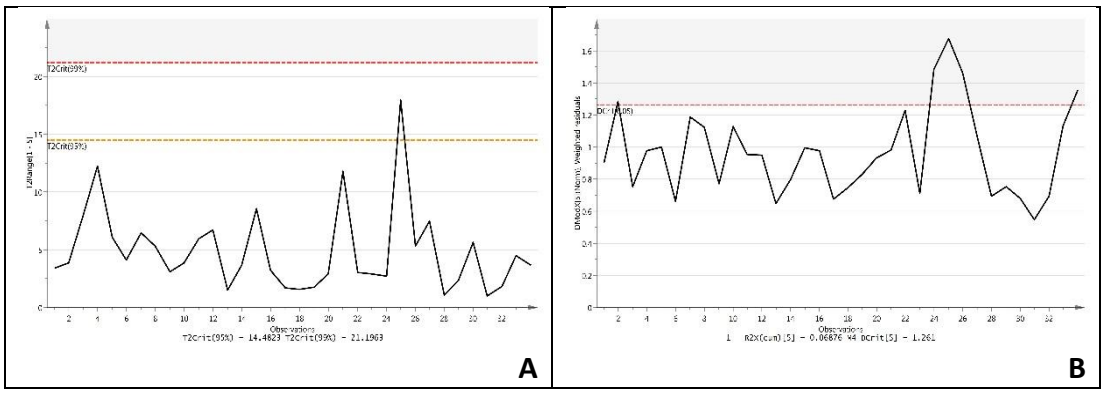


Figure 1.3 (A) Hotellings T2 plot, (B) Distance to model (DModX).
DModX on x-axis versus Hotelling's T2 on Y-axis. Hotelling's T2 has two limits on the y-axis; the warning limit (T2 Crit(95%)) represented by yellow dotted line and action limit (T2 Crit(99%)) represented by red dotted line. DModX uses critical distance DCrit at level 0.05 represented by red dotted line on the Y-axis. Observations on x-axis considered strong outliers if; above the action limit or above the warning limit and above the DModX critical limit.

1.6.4 Hierarchical cluster analysis (HCA)

Cluster analysis can be represented by a tree diagram or dendrogram. It aims to group samples based on their similarity to each other. It is a common method applied to summarise data structure (Figure 1.4). It arranges observations using the algorithm to provide a tree structure depending on similarity or dissimilarity between clusters based on the distribution of the variables (de Carvalho Rocha and Nogueira, 2011). The distances are calculated according to the Ward's linkage principle which is less susceptible to noise and outliers (Ward Jr, 1963, Eriksson et al., 2013a).

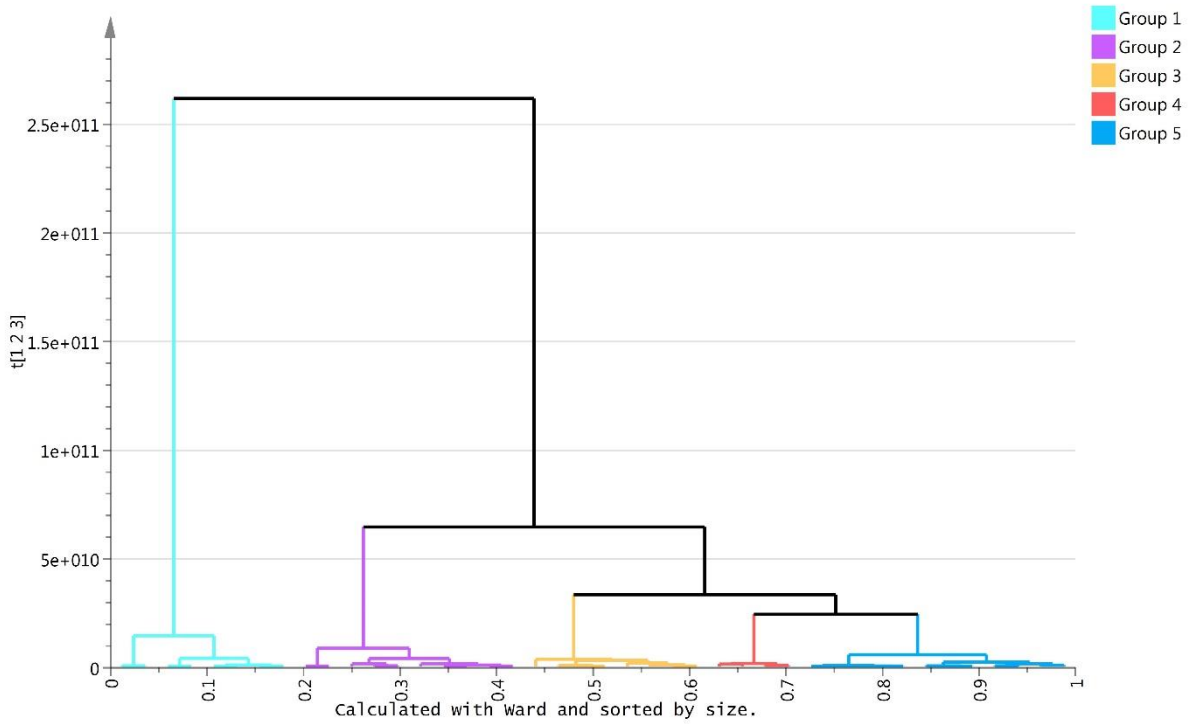


Figure 1.4 Hierarchical Clustering Analysis (HCA) plot.
The dendrogram shows observations clustered into five groups. X-axis represents observations and y-axis shows variability index.

1.6.5 Orthogonal projections to latent structures-discriminant analysis (OPLS-DA)

The aim of the statistical analysis is to establish a strong model that is capable of discriminating between the classes and providing a logical interpretation of the variances for the observations and variables. OPLS is an extended version of partial least squares projections (PLS) regression where the maximum variance can occur between the predictor variables and independent variables. OPLS has the advance over PLS since it can separate systematic variability in the predictors (x-variables) that are correlated to the responses (y-variables) which produce the predictive components. In addition to, systematic variability in the predictors (x-variables) that are uncorrelated to the responses (y-variables) which filtered out as orthogonal components (Trygg and Wold, 2002, Wheelock and Wheelock, 2013) as shown in the figure 1.5.

OPLS can be applied as a discriminate analysis (OPLS-DA). It provides a suitable method with a high degree of orthogonal variations of the predictors (Eriksson et al., 2013c). It is

straightforward to interpret the model and the analysis of qualitative data structures (Bylesjö et al., 2006).

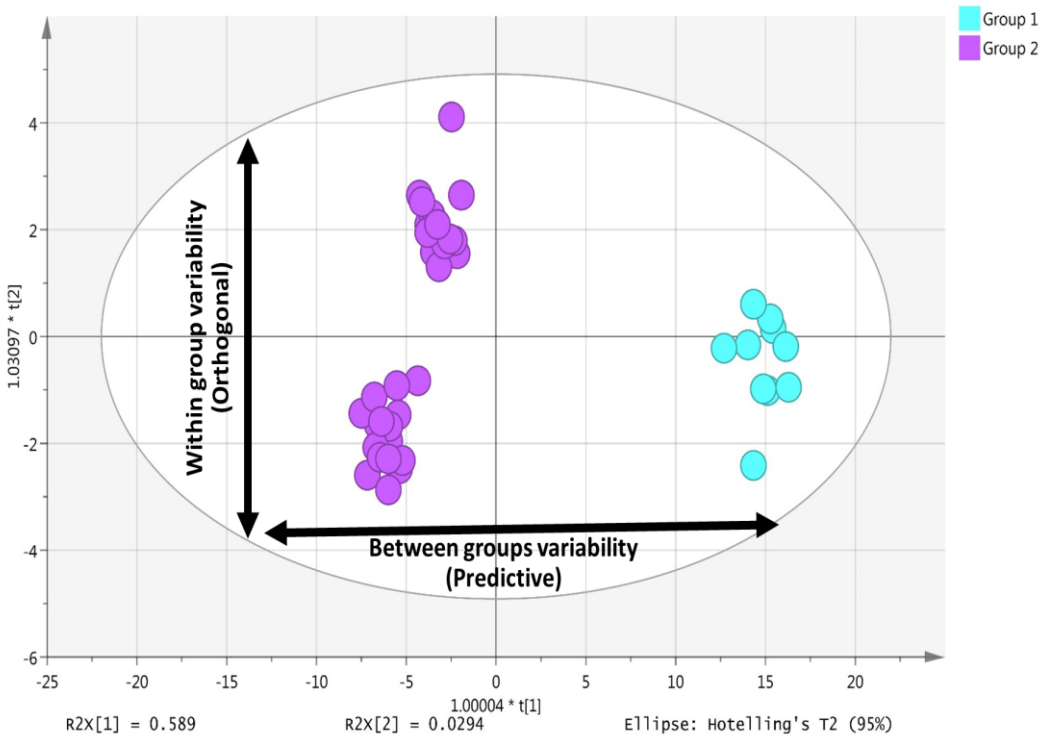


Figure 1.5 Orthogonal Partial Least Square Discriminant Analysis (OPLS-DA) score plot.
The OPLS-DA score plot shows two groups. The groups clearly separated horizontally (predictive), within group variability (orthogonal). The model explains 58.9% of the predictive variation (between groups variability) and 2.94% of the orthogonal variation (within group variability).

1.6.6 Model validation

A good model should have model coefficients with little bias which produce the correct analysis and clearly show the important variables. To achieve that, cross-validation - which measures the internal validation - should be considered. Cross-validation is used to measure the significance of the PCA and OPLS mode. The cross-validation parameters R^2 and Q^2 are used to understand the difference between the model fit and the predictive ability of the model, respectively. A goodness of fit, which is measured by parameter R^2X, is expressed by the explained variation. R^2X give a value of the closeness of the responses of the data to the training (calibration) data. R^2 (cum) is a sum of predictive and orthogonal variation in X that is explained by

model. The goodness of prediction can be estimated by parameter Q^2X which is expressed on the predicted variation. It is a predictive power of the model which measures the accuracy of the model in predicting the class of new observations using the model. The arbitrarily maximum value of R^2X and Q^2X is 1.

Although cross- validation tests the predictive power of a model, the statistical significance of the model predicatively is not a statement. This can be achieved by the response permutation test in SIMCA where the significance of R^2Y and Q^2Y values can be assessed. In the permutation test, the OPLS produces parallel models and randomly reorders the response (Y) data. At that point, the real R^2Y and Q^2Y from cross-validation are evaluated based on the distribution of a different order of Y data (Eriksson et al., 2013d).

The receiver operator characteristic (ROC) is used to assess the performance of the classes in the model by using two factors; the true positive (*sensitivity*) and false positive (1-*specificity*). The ROC curve plot combines both the sensitivity and 1-specificity of a test. The maximum sensitivity and *specificity* occur when the values are close to 1. Therefore, the false positive (1-*specificity*) should be close to zero (Figure 1.6). The area under the ROC curve (AUROC) is used as quality measurement, a high value near or equal to 1 indicates a perfect separation between the classes and a value close to 0.5 reveals that no separation is obtained (Westerhuis et al., 2008).

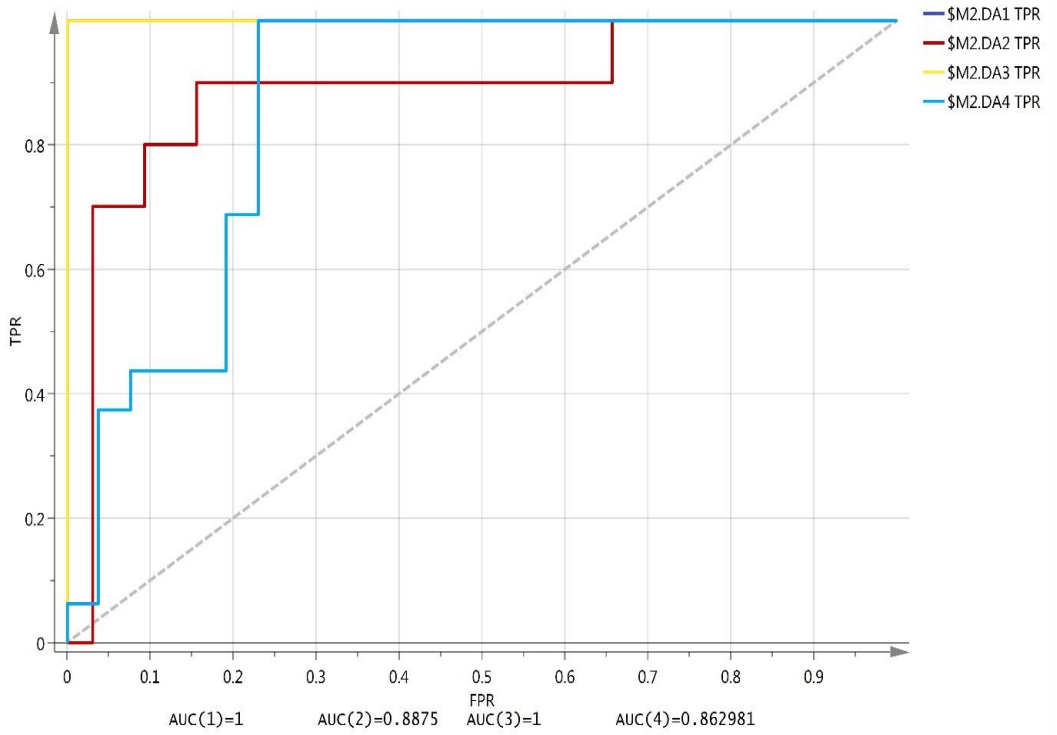


Figure 1.6 Area Under the Receiver Operating Characteristics Curve (AUROCC).
The ROC curve shows false positive rate (FPR = 1 - Specificity)) in the x-axis versus sensitivity (true positive rate (TPR)) on the y-axis. The quantitative measure of the area under the ROC (AUC) of the four group are showing and parameter ranges between 0.5 (bad classification) and 1.0 (perfect classification).

1.6.7 The S-plot

The S-plot offers a visualization tool that supports the interpretation of two classes in a supervised OPLS-DA model of the multivariate analysis data. The S-plot is, a scatter plot that has the shape of the letter S, showing a combination between the contribution (covariance) and correlation (reliability) structure between the variables and observations. The S-plot is applicable for Pareto scaling or with centering without a scaling model. It is composed of two vectors a predictive component, the x-axis (p1) representing the covariance and the y-axis (p(corr)) displaying the correlation (Wiklund et al., 2008).

The importance of the S-plot is to show the intensity (magnitude) as well as the reliability (i.e. high effect and lower uncertainty of putative biomarker). In spectrometric data, the peak intensity plays a role into distinguish the spurious correlation (false) peaks, especially low-intensity peaks which are close to the noise

level. The reliability shows the x-variables which have a strong correlation with the scores (observations), and they therefore improve the confidence of discriminating the variables which have an influence on the group classifications (Eriksson et al., 2013b).

1.6.8 Variable influence on projection (VIP)

VIP is a parameter that detects and measures the importance of the individual X-variables which have an influence on the OPLS-DA model with many components. The variable is considered as most relevant when it has a VIP value above 1 and irrelevant when the variable has a VIP value less than 0.5. In OPLS model, total VIP (VIP*tot*) is an accumulation of both the predictive (VIP*pred*) and orthogonal (VIP*orth*) components. The three VIP vectors are given to provide a strong lead for assessment of variables and interpretations of the model (Galindo-Prieto et al., 2015).

1.6.9 Jack-knifing uncertainties

For OPLS-VIP, jack-knifing uncertainty is determined based on the cross-validation procedure. It is used for measuring the standard error and 95% confidence intervals of variables to expose the precision of the estimate of the correct variables that are selected: it is important for filtering out unreliable variables (Afanador et al., 2013).

1.7 Conclusion

Forthcoming techniques in pharmaceutical impurity profiling will include advances in both technique-oriented and chemistry-guided approaches. The quick development of analytical technology will lead to fast profiling and significantly reduce the risk of missing potential impurities thanks to developments in orthogonal separations. The impurity profiling of pharmaceutical materials is important in qualifying and quantifying impurity levels, as well as identifying the sources of APIs. The development of analytical technologies with orthogonal separations or detection

mechanisms (such as LC/MS^n, NMR, HILIC and IC) will lead to rapid impurity profiling techniques and significantly reduce the risk of missing potential impurities. The chemistry-guided approaches will involve rapid structure elucidation technologies and optimizing statistical and computational tools (such as PCA and HCA) will result in more accurate elucidation of impurity structures and advanced data screening (Baertschi, 2006).

Chapter 2

Materials and Methods

2 Materials and Methods

2.1 Chemicals and Solvents

For liquid chromatography, HPLC-grade acetonitrile and methanol were purchased from Fisher Scientific (Loughborough, UK) and HPLC water was purified by a Milli-Q purification system from Millipore (Watford, UK). Analytical grade ammonium acetate, formic acid, HPLC-grade acetonitrile and methanol were purchased from Fisher Scientific (Loughborough, UK).

All IC reagents were of analytical or reagent grade. Samples and standard materials were diluted using IC water, which was purchased from Sigma-Aldrich (Dorset, UK). A Seven anionic mixture standard containing; fluoride, chloride, nitrite, bromate, nitrate, phosphate and sulphate was purchased from Thermo Scientific (Hemel Hempstead, UK). Ammonium citrate, ammonium lactate, ammonium bromide, ammonium formate, butyric acid, sodium chloride, ammonium chloride, potassium hydrogen phosphate, calcium chloride and ammonium acetate were purchased from Sigma Aldrich (Dorset, UK).

2.2 Sample Preparation

The samples were diluted as required for the respective analyses.

2.3 Liquid chromatography- Mass spectrometry

The liquid chromatography - high resolution mass spectrometry (LC-HRMS) analysis was performed by using an Accela HPLC system interfaced with an Orbitrap Exactive mass spectrometer from Thermo Fisher (Hemel Hempsted, UK). The LC/MS was calibrated according to the manufacturer's instructions and set at 30000 resolution. Samples were analysed in both negative and positive ion modes with a spray voltage of 4.5 kV in positive mode and -4.0kV in negative ion mode. The ion transfer capillary

temperature was 275˚C. 50 and 17 arbitrary units were used as flow rates for sheath and auxiliary gases, respectively. The full mass range was m/z 70-1200.

2.4 LC - MS^n

LC/MS^n was performed by a Surveyor HPLC system coupled with a Finnigan LTQ Orbitrap from Thermo Fisher (Hemel Hempstead, UK). The instrument was calibrated according of the manufacturer's instructions and set at 30000 resolution. Samples were analysed in positive and negative ion mode with a spray voltage of (+/-) 4.5 kV. The ion transfer capillary temperature was 275˚C. 45 and 15 arbitrary units were used as flow rate for sheath and auxiliary gases respectively.

2.5 LC/MS- MSMS software

Xcalibur software version 2.1.0 (Thermo Fisher Scientific) was employed for data acquisition and processing.

2.6 LC-HRMS data processing

All raw data files (Thermo-Xcalibur format) were extracted to obtain an electrospray ionisation (ESI) positive and negative dataset. Typical data processing involving filtering, feature detection, alignment, normalization and visualization and was obtained by using the MZmine 2.20 software (Pluskal et al., 2010). A maximum retention time (RT) shift of 2 min and mass tolerances not exceeding 5 ppm were adopted. The peak detection peak was done by applying the intensity of no less 1.0E5 and no larger than 3 signal to noise with a maximum charge of 2. The data were also de-isotoped. The multivariate statistical analysis of datasets was performed by using SIMCA-P version 14.1 software, Umetrics (Umea, Sweden). Pareto scaling was applied for each unsupervised and supervised model. Principal component analysis (PCA) was applied for unsupervised data visualisation for all materials depending on the distribution of data based on the direction of highest variation. An

orthogonal partial least squares discriminant analysis (OPLS-DA) is used as a supervised model to obtain perfect discrimination between selected classes.

2.7 Capillary-High pressure ion chromatography system

All measurements were conducted using the Reagent-Free Capillary HPIC system, an ICS-5000+ instrument (Thermo Fisher Scientific, Hemel Hempstead UK). Anions were separated as the system was equipped with a potassium hydroxide electrolytic eluent generation module, an anion capillary eluent suppressor (ACES), an AS-AP autosampler and a conductivity detector.

Cations were separated as the system was equipped with a methane sulphonic acid (MSA) electrolytic eluent generation module, a cation capillary eluent suppressor (CCES), an AS-AP autosampler and a conductivity detector. All eluents were generated as potassium hydroxide and methane sulphonic acid using an ultrapure water (ICW 3000) purification system with a resistivity of 18.2 MΩ.cm ((ICW 3000), Millipore, Watford, UK) as a diluent. A Dionex IonPac AS11-HC 4 µm (250 × 0.4 mm) capillary column was used in the anionic separations. The optimized gradient program was: 1mM KOH (0 to 10 min), 1 to 30 mM KOH (10 to 30 min), 30 to 60 mM (30 to 40 min) and 60 mM KOH (40 to 50 min); IC flow rate was 10 µL min^{-1}, the injection volume was 1 µL, maximum back pressure was 3200 psi and the column temperature was maintained at 15°C. Cations were separated with a Thermo Scientific Dionex IonPac™ CS19 capillary column 4 µm (0.4 mm × 250 mm). The system was employed using a Thermo Scientific Dionex IonPac™ CG19 guard column. The optimized gradient program was: 8mM MSA (0 to 20 min), IC flow rate was 10 µL min^{-1}, the injection volume was 1 µL, maximum back pressure was 3200 psi and the column temperature was maintained at 35°C. The data were collected and analysed using Chromeleon 6.8 software (Thermo Fisher Scientific).

2.8 Headspace GC-MS

The HS-GC-MS analysis was performed by using a Combipal Headspace sampler with a TRACE GC 2000 gas chromatograph interfaced to a MD 800 mass spectrometer quadrupole (Thermo, Hemel, UK). The GC housed a fused-silica column BP-5MS (30 m×0.25 mm, *d*f – 0.25 µm) (Crawford Scientific, Strathaven, Scotland), the column was programmed from 40 to 100°C at 5°C/min. The helium carrier gas was set at 150 kPa. The sample was dissolved in water and 1 ml of solution was put into a 20 ml headspace vial (Chromcol, obtained for Speck-Burke, Alva, Scotland) closed with a crimped cap with a rubber septum. The sample was shaken at 70°C for 5 minutes and 0.5 ml of the headspace was injected into the GC with the split ratio set at 10:1. The injector port was maintained at 150°C. The temperature of interface was 270°C and source was operated at 200°C, respectively. Ionisation was carried out in electron impact mode at 70 eV.

2.9 NMR

A Bruker AVANCE III 600 NMR (Rhinstetten, Germany) spectrometer operated at a proton resonance frequency of 600.13 MHz was equipped with a TBI [1H, 13C, 31P–15N]-z triple resonance probe head fitted with an actively shielded gradient coil for delivery of pulsed-field gradients. The high acid clarity sample was prepared by dissolving 1 g in 10mL of 1M HCl and filtered using nylon polyethylene filter with a 2µm pore size. The residue from the filter was dissolved in 2 ml (1:1) methanol/water and dried, then dissolved in 700 µl of DMSO-d_6 and transferred to a 5 mm Ø standard NMR tube.

Chapter 3

Impurity profiling of alpha lipoic acid using LC/HRMS-(MS^n) and multivariate analysis

3 Impurity profiling of alpha lipoic acid using LC/HRMS-(MS^n) and multivariate analysis

3.1 Introduction

Alpha Lipoic acid (LA) is also known as thioctic acid; 1, 2-dithiolane-3-pentanoic acid; 1, 2-dithiolane-3-valeric acid; and 6, 8-thioctic acid. It was discovered by Snell in 1937 and first isolated by Reed in 1951 (Packer et al., 1996). LA is a potent free radical scavenger and antioxidant. It is a cofactor of many enzymes which have roles in the oxidative decarboxylation of pyruvic acid and other ketoacids and in the glycine cleavage system. It is biosynthesised from cysteine and octanoic acid in tissues and is abundant in mitochondria where cell metabolism and energy production reactions occur (Packer et al., 1995, Biewenga et al., 1997).

The synthesis ways of LA were described in many publications and academic journals. Racemic LA was synthesised earlier by a different approach. The general synthesis route starts from a compound containing 6-8 carbon atoms with a hydroxyl or halogen function group and then these are replaced by various sulphur nucleophiles (Brookes, 1986).

The biological activity of LA shows that the R-enantiomer is more active than its antipode. The enantioselective synthesis LA was continuously developed to produce high overall yields, reduce the synthesis steps and for diastereoisomer separation. A method of synthesis with a low number of steps was devised from 6-bromohex-1-ene and using a sharpless asymmetric epoxidation of the intermediate allyl alcohol to control the enantioselectivity as shown in figure 3.1 (Page et al., 1990).

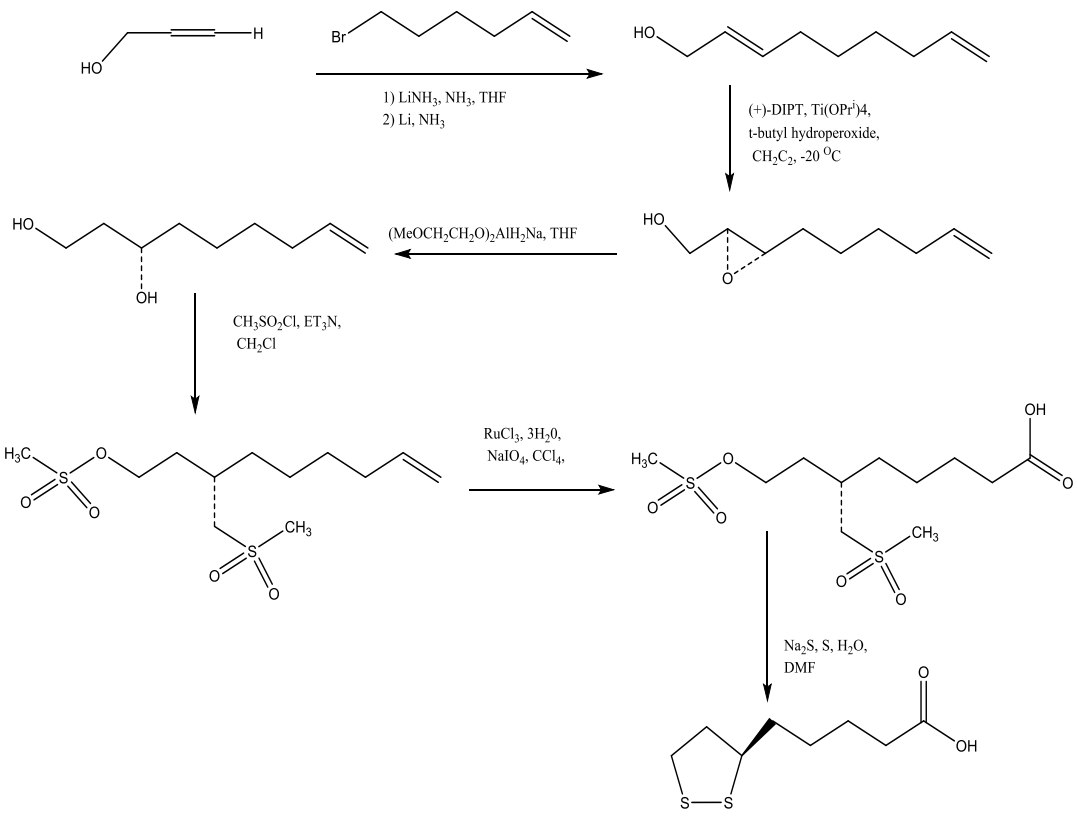


Figure 3.1 Chemical Synthesis of LA as formed by (Page et al., 1990).

LA and, its reduced form, dihydrolipoic acid (DHLA) have been widely used in the treatment of diseases such as diabetic retinopathy and neuropathy, to protect against oxidative stress, age associated cardiovascular, cognitive, and neuromuscular deficits, liver diseases, radiation injury and neurodegenerative disorders (Krishnan et al., Reljanovic et al., 1999, Marshall et al., 1982, Ramakrishnan et al., 1992).

LA can be degraded gradually at room temperature or by exposure to light and polymerisation occurs at temperatures higher than its melting point (59–62 °C) (Wagner et al., 1956). LA exists in two enantiomeric forms, R and S. The commercial form of LA contains a synthetic racemic mixture of the natural isomer (R-form) and the unnatural isomer (S-form). There is still concern that the R-form is unstable since it is gradually degraded at temperatures above 40 °C. In addition, LA synthetic mixture has certain drawbacks such as being hygroscopic. Therefore, stabilization methods for LA were widely developed and include crystallizing the R-form with materials such as Tris buffer or nicotinamide (Ames, 2006, Zhao et al., 2014).

Kataoka (1998) reported successful application of chromatographic methods in the resolution of LA from some of its analogues which include DHLA, lipoamide and methyl lipoic acid (MLA) (Kataoka, 1998). Chromatographic analysis has also been widely used in determination of free LA, DHLA and/or LA derivatives in different matrices such as biological samples (Satoh et al., 2007, Chng et al., 2010, Satoh et al., 2008), foodstuffs (Durrani et al., 2010) and dietary supplements (Santos Pereira et al., 2016).

HPLC methods coupled with mass spectrometry (MS) and tandem mass spectrometry (MS/MS) detectors are the most common methods applied for the analysis of LA (Santos Pereira et al., 2016, Durrani et al., 2010, Nikolić et al., 2014, Chen et al., 2005, Lauridsen et al., 2001) and for identifying or elucidating the structure of drug impurities. The structure of the impurity can be proposed from integrated chromatographic and spectroscopic data. One of the important aspects of generating an impurity profile is that the data acquired in the chromatogram and spectrum are highly informative for providing evidence on the sources of impurities such as starting materials or side reactions. This will ensure that appropriate steps have been taken to avoid or minimise their formation during the manufacturing process thus improving quality of the manufactured products (Görög, 2006) or improving the crystallization process (Ottens et al., 2001).

Recently, pattern recognition techniques (PRTs) such as multivariate analysis have attracted great attention as powerful statistical tools which can be used for fingerprinting chemical compounds. PRTs are broadly used for various analytical purposes such as quality control in the pharmaceutical industry (Ferreira and Tobyn, 2015), metabolomics (Worley and Powers, 2013), and targeted and non-targeted impurity profiling (Acevska et al., 2015, Zhou, 2016).

In the current study, there was a requirement for a method to able to distinguish the impurity profile of lipoic acid raw material from those of three of its purified samples in order to determine the success of the crystallisation process in removing impurities without introducing further impurities. Thus the aim of this research was to study the

impurity profiles of the four lipoic acid materials using LC/MS^n and multivariate analysis of the acquired chromatographic and mass spectrometric data, to characterise unknown impurities and assess the success of the crystallisation process in removing impurities.

3.2 Methodology

3.2.1 Chemicals and Reagents

The DL- lipoic acid was obtained from Molekula Chemicals and Biochemical (Dorset, UK). The LA raw material was purified by dissolving a specific amount of it in ethyl acetate at a ratio of 10 mg LA to 20 ml ethyl acetate in an ultrasonic bath at room temperature. After that, 80 ml hexane was added in proportion to the amount of LA and then filtered.

The three purified materials of Lipoic acid were prepared by the Continuous Manufacture and Crystallisation (CMAC group) as part of a study on the crystallisation of lipoic acid.

3.2.2 Sample preparation

The samples were prepared by dissolving 20 mg of the material in 10 mL of ethanol then adding water till 20 mL at the concentration of 1.0 mg/mL.

3.2.3 Reversed Phase liquid chromatography (RP-LC) conditions

In the RP-HPLC system, the reversed phase column used was ACE Super C18 (150 × 4.6 mm, 3 µm) from HiChrom (Reading, UK). A gradient was run between 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The gradient was started at 10% B at 1 min, and then linearly increased to 90% B in 30 min and a flow rate of 0.3 ml/min for 30 min. The full mass range was (70 - 1200) m/z. Xcalibur software version 2.1.0 (Thermo Fisher Scientific) was employed for data acquisition

and processing. Deprotonated molecular ions ($[M-H]^-$) were examined in order to detect the molecular weights of the impurities.

3.2.4 Liquid chromatography- Mass spectrometry

Refer to section 2.3.

3.2.5 Liquid chromatography- MS^n

Refer to section 2.4. Deprotonated molecular ions ($[M-H]^-$) were applied for detecting the molecular weights.

3.3 Results and discussion

3.3.1 LC-HRMS/MS analysis

Liquid chromatography – high resolution mass spectrometry (LC/HRMS) can be used in the separation and identification of impurities based on their mass to charge ratios even in the case of LC co-elution peaks. It was applied to characterize the impurity profiles of α-lipoic acid (LA) raw material and three purified samples. Earlier attempts to crystallise LA raw material with nicotinamide did not yield any co-crystals but instead formed a jelly-like material. However, upon purification, co-crystals were easily formed with nicotinamide. The failure of the co-crystallisation process was thought to arise from impurities in the LA raw material which were removed upon initial purification using the ethyl acetate hexane system. By identifying the impurities which are present in the LA raw material but absent in the purified materials, it might possible to determine which of the impurities are responsible for the failure of the co-crystallisation process. The total ion chromatograms (TICs) of the four LA materials are presented in Figure 3.2. All peaks were manually integrated and the accurate ion masses and isotope patterns were used to detect and identify the impurities based on their elemental compositions generated with a mass tolerance of less than 3 ppm between theoretical and experimental values. The electrospray ionization multistage mass spectrometry (ESI-MS^n, n = 2, 3) was applied to get on the

fragmentation patterns of LA and its impurities. This was performed to provide strong supporting evidence for the identity of the impurities and to understand how the side products might be formed from the synthesis pathway of LA.

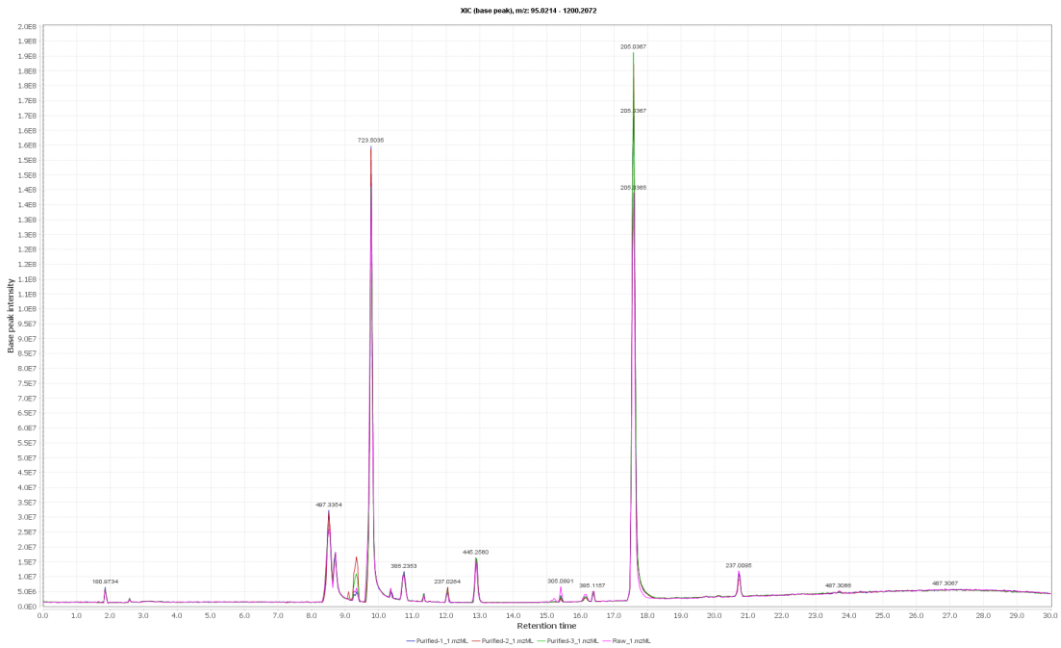


Figure 3.2 Horizontal alignment of the total ion chromatograms (TICs) of all four lipoic acid materials produced by MZmine 2.20 software.

Initially, non-targeted, full scan LC/HRMS data were generated and diagnostic ions from the spectra were obtained by all-ions fragmentation (AIF) which exposed the precursor ions to collisional energy and fragments were measured in the Orbitrap mass analyser.

The deprotonated LA which has the molecular weight, $[M-H]^-$, of m/z: 205.03514 and elemental composition ($C_8H_{13}O_2S_2^-$) was confirmed by negative ion HRMS. ESI-MS/MS results of the LA ion generated main fragment ions at m/z 171.048, 159.048, 127.0588 and 93.071 which represented the fragments $C_8H_{11}O_2S$, $C_7H_{11}O_2S$, $C_7H_{11}S$ and C_7H_9, respectively. The expected fragmentation pathway of LA generated the main fragment ion at m/z 171.04 $[M-H-H_2S]^-$ which is due to loss of the thiol group and then the alpha-cleavage could sequentially occur by losing the carboxylic acid to produce ions at m/z 159.04 and 127.05 (Figure 3.3).

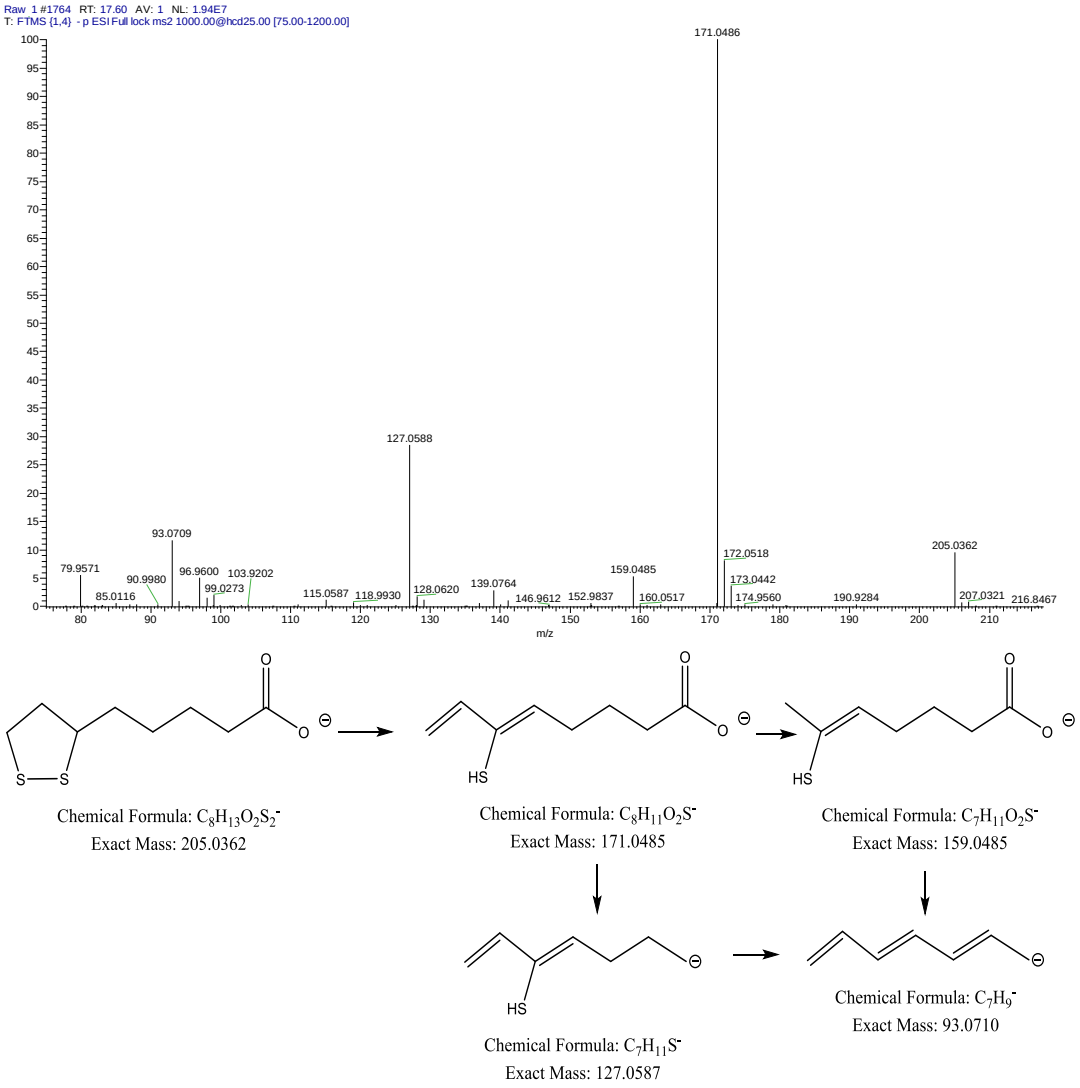


Figure 3.3 MS^2 Fragmentation mass spectrum obtained for lipoic acid and the proposed structures of the characteristic fragments.

The relative contents (%) of the detected peaks were calculated and expressed as a percentage of lipoic acid in the sample based on their peak areas. As a result, more than 25 impurities were detected most of which degradation products of LA were revealing some very complex chemistry in the side reactions of lipoic acid during its synthesis (Table 3.1).

Table 3.1. The retention time, mass to charge ratio and relative intensities (% relative to lipoic acid) of lipoic acid impurities of the raw material and the three purified samples from LC/HRMS.

RT (min)	m/z (measured)	Relative content (%)			
		Raw material	Purified sample 1	Purified sample 2	Purified sample 3
17.59	171.0488	13.8	13.95	13.6	13.87
9.34	221.0313	6.07	3.8	12.32	7.42
20.71	237.0086	7.39	5.29	4.62	5.2
12.01	237.0265	2.84	1.99	2.69	2.1
17.59	251.0421	72.14	71.13	71.98	70.65
9.26	267.037	4.04	2.51	7.83	4.76
20.71	283.0142	6.17	4.34	3.85	4.28
12.01	283.0321	1.36	1.02	1.35	0.99
15.41	305.0893	3.09	1.32	1.07	1.33
19.88	313.1485	3.21	0.58	0.25	0.43
15.22	333.0844	1.4	0.55	0.51	0.61
17.55	334.9952	0.95	0.82	0.88	0.82
17.55	339.0167	2.32	1.8	1.91	1.75
21.82	345.1205	0.93	0.15	0.08	0.11
18.89	347.1358	1.53	0.75	0.59	0.82
19.45	347.1358	0.7	0.3	0.18	0.23
17.55	350.9673	4.21	3.67	3.79	3.58
17.55	367.0118	3.91	3.29	3.21	3.39
16.36	385.1153	2.63	1.63	1.96	1.75
17.59	411.0807	19.11	19.35	20.66	20.24
17.59	527.0498	3.15	2.73	2.84	2.75
22.3	585.1523	2.55	0.94	1.27	0.95
23.69	617.1243	2.78	0.95	1.34	1.01
24.36	617.1243	0.82	0.24	0.28	0.24
17.59	687.0877	2.56	2.33	2.37	2.31
20.67	203.0208	0.83	0.6	0.51	0.6

The majority of the selected ions were investigated by MS^n analysis (n = 2, 3). This led to the elucidation of structures and degradation pathways (Table 3.2). It seems clear that the most abundant ion at m/z 251.0413, which is a formic acid adduct of LA

formed in the mass spectrometer This gives fragments at m/z 205.036 (LA) due to loss of formic acid and at m/z 171.049 due to additional loss of H_2S from LA.

Table 3.2 The accurate m/z, possible elemental composition, MS^2 and MS^3 fragments of lipoic acid impurities.

Impurity				MS^2 product ion		MS^3 product ion	
m/z (measure)	**Formula**	**RDB**	**Error (ppm)**	**m/z**	**Formula**	**m/z**	**Formula**
221.0307	$C_8H_{13}O_3S_2$	2.5	-1.715	157.087	$C_8H_{13}O_3$	139.077	$C_8H_{11}O_2$
						113.097	$C_7H_{13}O$
						83.0503	C_5H_7O
				125.043	C_7H_9S		
				79.9397	OS_2		
				139.077	$C_8H_{11}O_2$		
				203.021	$C_8H_{11}O_2S_2$		
				187.044	$C_8H_{11}O_3S$		
237.0079	$C_8H_{13}O_2S_3$	2.5	-1.878	203.021	$C_8H_{11}O_2S_2$		
				169.033	$C_8H_9O_2S$		
				125.043	C_7H_9S		
251.0413	$C_9H_{15}O_4S_2$	FORMATE ADDUCT		205.036	$C_8H_{13}O_2S_2$	171.049	$C_8H_{11}O_2S$
						127.059	$C_7H_{11}S$
				171.049	$C_8H_{11}O_2S$		
267.0363	$C_9H_{15}O_4S_2$	FORMATE ADDUCT of m/z 221		221.031	$C_8H_{13}O_3S_2$	157.087	$C_8H_{13}O_3$
						203.021	$C_8H_{11}O_2S_2$
						139.077	$C8H_{11}O_2$
						125.043	C_7H_9S
				203.02	$C_8H_{11}O_2S_2$		
				157.087	$C_8H_{13}O_3$		
				125.043	C_7H_9S		
				187.043	$C_8H_{11}O_3S$		
305.088	$C_{13}H_{21}O_4S_2$	3.5	-2.176	287.079	$C_{13}H_{19}O_3S_2$	243.089	$C_{12}H_{19}OS_2$
						189.042	$C_8H_{13}OS_2$
						209.101	$C_{12}H_{17}OS$
						173.065	$C_8H_{13}O_2S$
						139.077	$C_8H_{11}O_2$
				243.089	$C_{12}H_{19}OS_2$		
				207.052	$C_8H_{15}O_2S_2$		
				173.065	$C_8H_{13}O2S$		
				131.017	$C_5H_7O_2S$		

(Cont.)

Impurity				MS^2 product ion		MS^3 product ion	
m/z (measure)	**Formula**	**RDB**	**Error (ppm)**	**m/z**	**Formula**	**m/z**	**Formula**
333.0831	$C_{14}H_{21}O_5S_2$	4.5	-1.556	159.012	$C_6H_7O_3S$		
				125.025	$C_6H_7O_3$	81.0347	C_5H_5O
				315.074	$C_{14}H_{19}O_4S_2$		
				207.052	$C_8H_{15}O_2S_2$		
				115.023	C_5H_7OS		
				171.049	$C_8H_{11}O_2S$		
385.1144	$C_{18}H_{25}O_5S_2$	6.5	-1.009	367.104	$C_{18}H_{23}O_4S_2$	173.064	$C_8H_{13}O_2S$
				207.051	$C_8H_{15}O_2S_2$	173.064	$C_8H_{13}O_2S$
						139.077	$C_8H_{11}O_2$
				211.043	$C_{10}H_{11}O_3S$		
				177.055	$C_{10}H_9O_3$		
				173.064	$C_8H_{13}O_2S$		
411.0791	$C_{16}H_{27}O_4S_4$	3.5	-0.281	207.052	$C_8H_{15}O_2S_2$		
				171.049	$C_8H_{11}O_2S$		
585.1508	$C_{24}H_{41}O_6S_5$	4.5	-0.471	381.123	$C_{16}H_{29}O_4S_3$	363.113	$C_{16}H_{27}O_3S_3$
						347.136	$C_{16}H_{27}O_4S_2$
						329.126	$C_{16}H_{25}O_3S_2$
						295.138	$C_{16}H_{23}O_3S$
						251.148	$C_{15}H_{23}OS$
						173.064	$C_8H_{13}O_2S$
				171.048	$C_8H_{11}O_2S$		
617.1229	$C_{24}H_{41}O_6S_6$	4.5	-0.594	207.052	$C_8H_{15}O_2S_2$	173.064	$C_8H_{13}O_2S$
				411.08	$C_{16}H_{27}O_4S_4$		

The data from LC-MS/MS and accurate mass together confirm the structure of m/z 221.0307 whose elemental composition is $C_8H_{13}O_3S_2$. Collision induced dissociation (CID) of this precursor ion generated daughter ions at m/z 203.021, 187.043, 157.087, 139.076, 125.043 and 79.939 (Figure 3.4A). The product ion at m/z 139.0776 results from the neutral loss of disulphur monoxide (S_2O) from the precursor ion and/or of water from the abundant daughter ion at m/z 157.087. The ion at m/z 157.087 also generates other fragment ions at m/z 113.097 and 83.050 as shown in the spectrum (Figure 3.4B). This leads to a proposed structure of m/z 221.0307 as being a thiosulphinate of LA (Beta-Lipoic acid) as shown in Figure 3.5. In addition, the ion at m/z 267.0363 is the formic acid adduct of the precursor ion at m/z 221.0307.

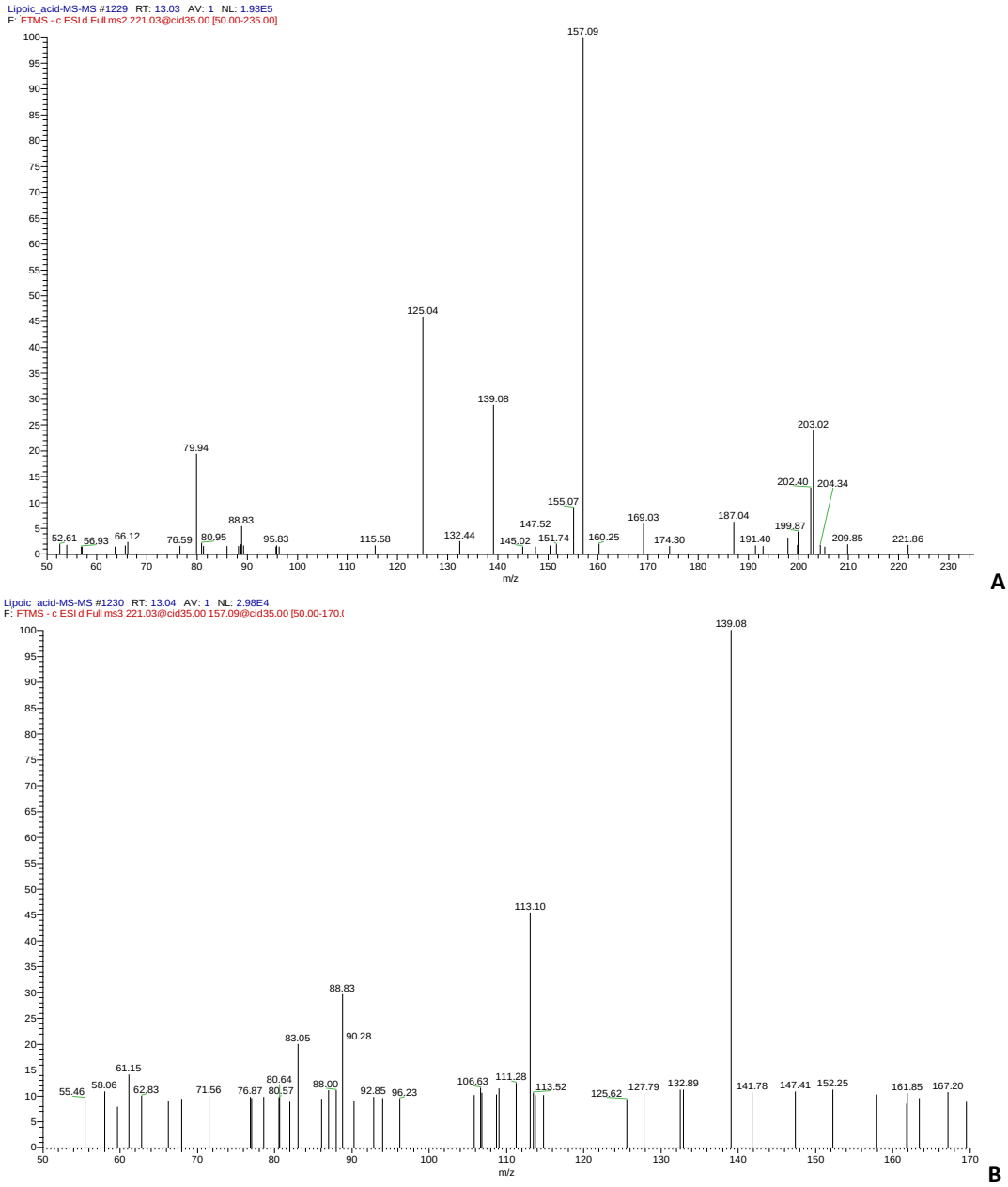


Figure 3.4 Impurity at m/z 221.0307 (A) MS^2 spectrum of m/z 221.0307 (B) MS^3 of fragment ion at m/z 157.087.

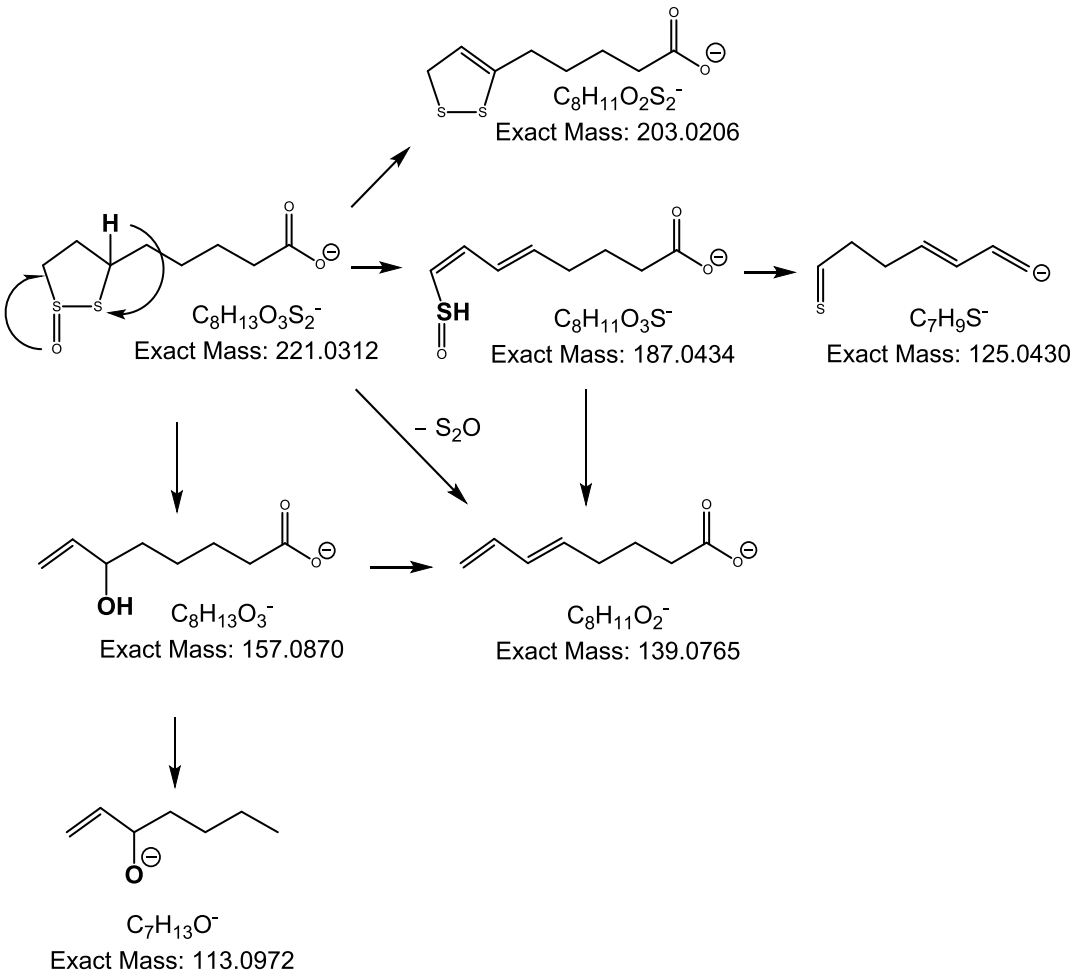


Figure 3.5 Proposed fragmentation mechanism of impurity at m/z 221.0307

The HR-MS/MS data revealed that impurity m/z 237.0079 has an elemental composition of $C_8H_{13}O_2S_3$. The CID produced fragments of the molecular ion at m/z 203.021, 169.033 and 125.043 (Figure 3.6A). This is related to the sequential loss of the sulphur atoms 1, 2 and decarboxylation, respectively (Figure 3.6B). The more abundant fragment peak at m/z 203.021, which is an unsaturated derivative of lipoic acid, which also shows up as a fragment ion of the precursor ion at m/z 221.0307. Presumably this impurity can form during synthesis since reagents such as Na_2S can insert an additional sulphur atom into the five membered ring.

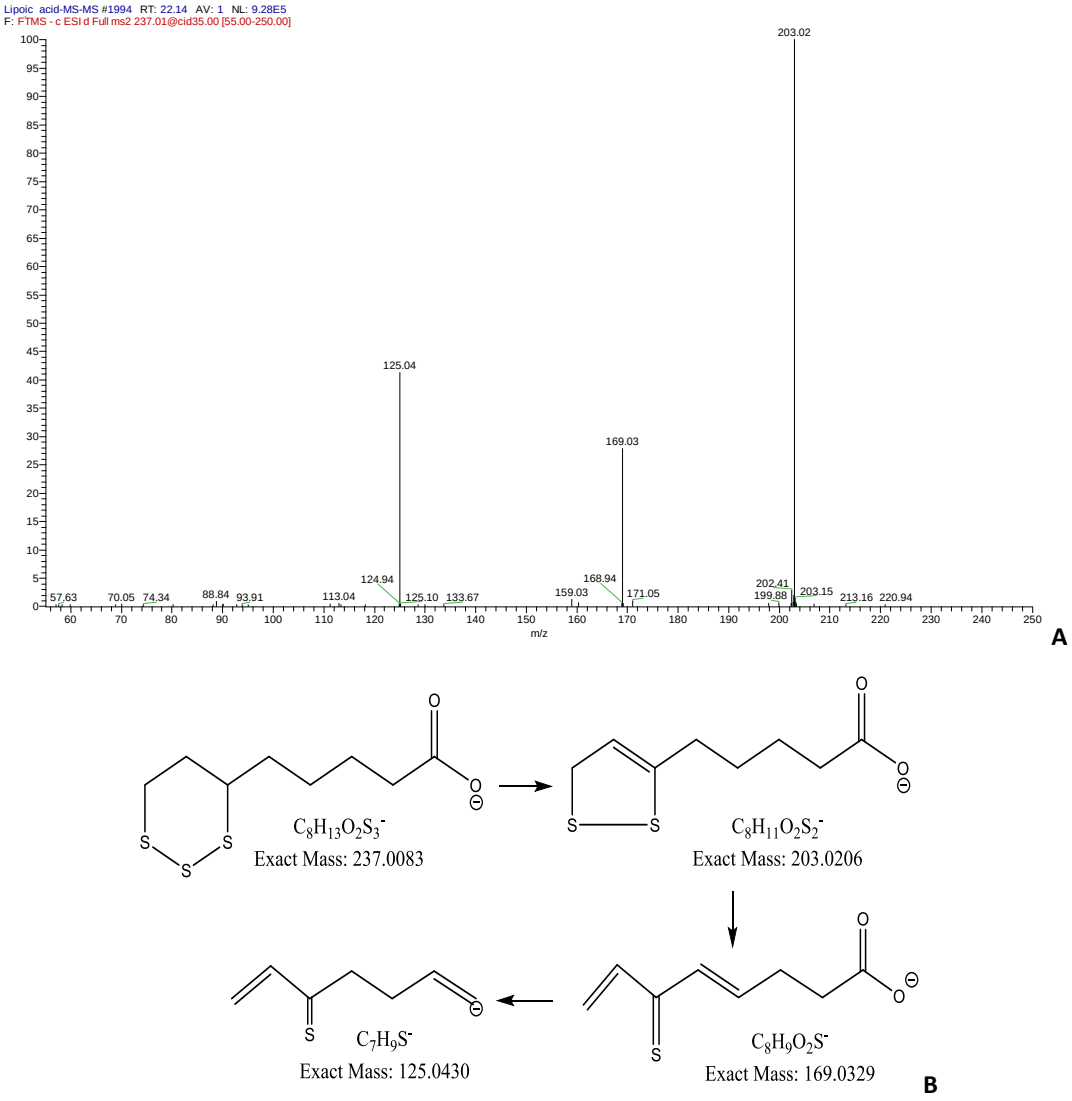


Figure 3.6 Impurity at m/z 237.0083 (A) MS^2 spectrum (B) proposed fragmentation mechanism.

The HRMS data proposed that the impurity at m/z 305.0880 has an elemental composition of $C_{13}H_{21}O_4S_2^-$. The collision activities induced product ions at m/z 287.078, 243.088, 207.052, 173.064 and 131.01 as shown in Figure 3.7 (A). The more abundant fragment peak at m/z 287.078 is related to water loss. The MS^3 of the ion at m/z 287.078 was generated to give ions at m/z 243.089, 189.041, 173.064, 209.101 and 139.086 (Figure 3.7B). This impurity could be produced by adding a ($C_5H_8O_2$) to LA via one of the sulphur atoms in the dithiolane ring (Figure 3.7C). This is not a straightforward side product like the two impurities discussed above and an assignment of its identity is very speculative.

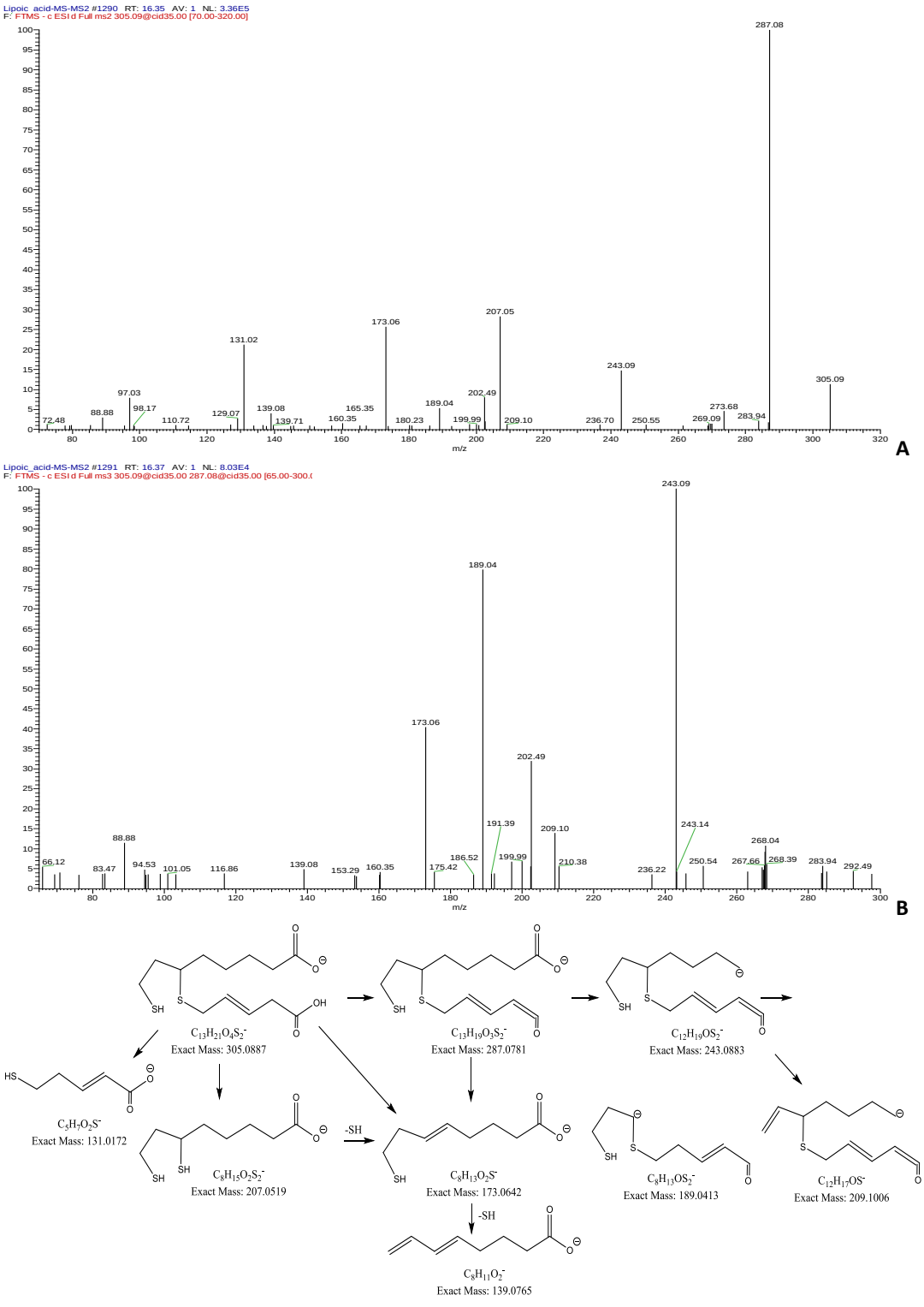
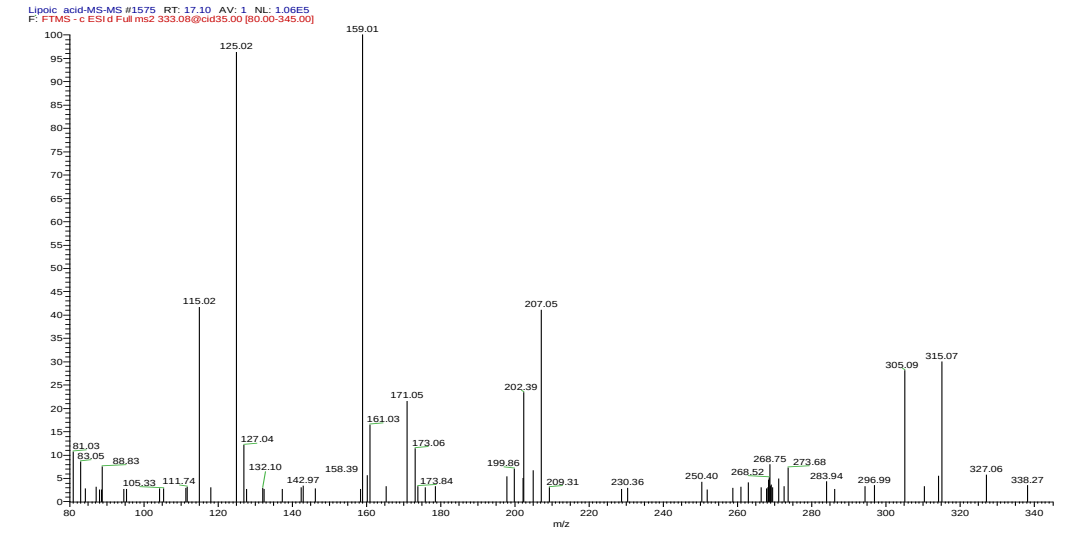
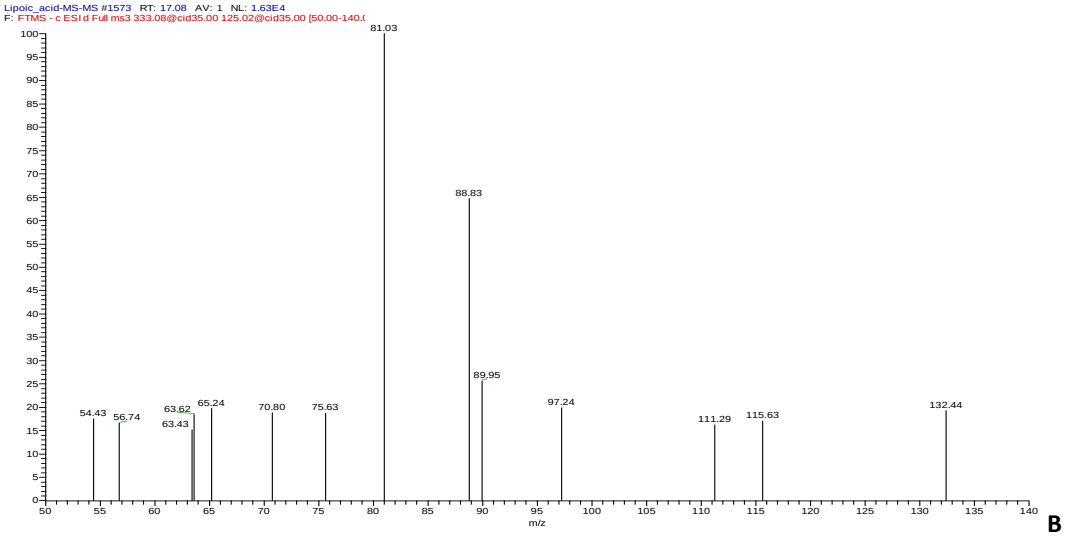


Figure 3.7 Impurity at m/z 305.0880 (A) MS^2 spectrum of m/z 305.0880 (B) MS^3 of fragment ion at m/z 287.078 (C) proposed fragmentation mechanism.

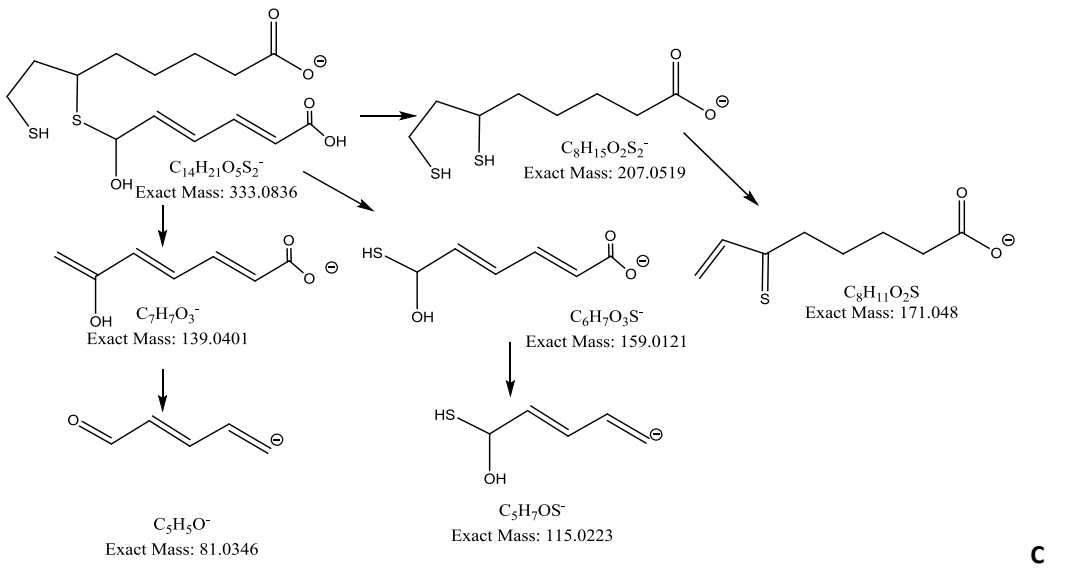
The deprotonated molecular ion at m/z 333.0831 with predicted molecular formula $C_{14}H_{21}O_5S_2$ fragmented to form product ions at m/z 125.024, 159.012, 115.022, 171.048, 207.052 and 315.073 (Figure 3.8A). The more abundant ion at m/z 125.024 has elemental composition ($C_6H_7O_3$) and generated the ion at m/z 81.034 after MS^3, resulting from loss of carboxylic acid group (Figure 3.8B). The fragment ion at m/z 207.052 has a chemical formula of $C_8H_{15}O_2S_2$ with a mass error of 0.626 ppm which corresponds to deprotonated dihydrolipoic acid (DHLA) (Figure 3.8C). The minor fragment ion at m/z 315.073 is due to water loss. Again the structure is only speculative since there is nothing in the synthetic route that would rationalise its formation.



A



B



C

Figure 3.8 Impurity at m/z 333.0831 (A) MS^2 spectrum of m/z 333.0831 (B) MS^3 of fragment ion at m/z 125.024 (C) proposed fragmentation mechanism.

The impurity at m/z 385.1144 with formula ($C_{18}H_{25}O_5S_2$) fragmented to yield product ions at m/z 207.051, 211.043, 177.055, 367.103 and 173.064 (Figure 3.9A). The most abundant fragment at m/z 207.051 is related to DHLA. The remaining part was shown in the fragmentation ions at m/z 211.043 and 177.055 with a mass error less than 0.4 ppm, which correlated to $C_{10}H_{11}O_3S$ and $C_{10}H_9O_3$, respectively. The product ions at m/z 173.064 and 139.076 were formed from MS^3 of DHLA, are a result of sequential loss of the SH_2 group (Figure 3.9B). The minor fragment ion at m/z 367.103 results from water loss (Figure 3.9C).

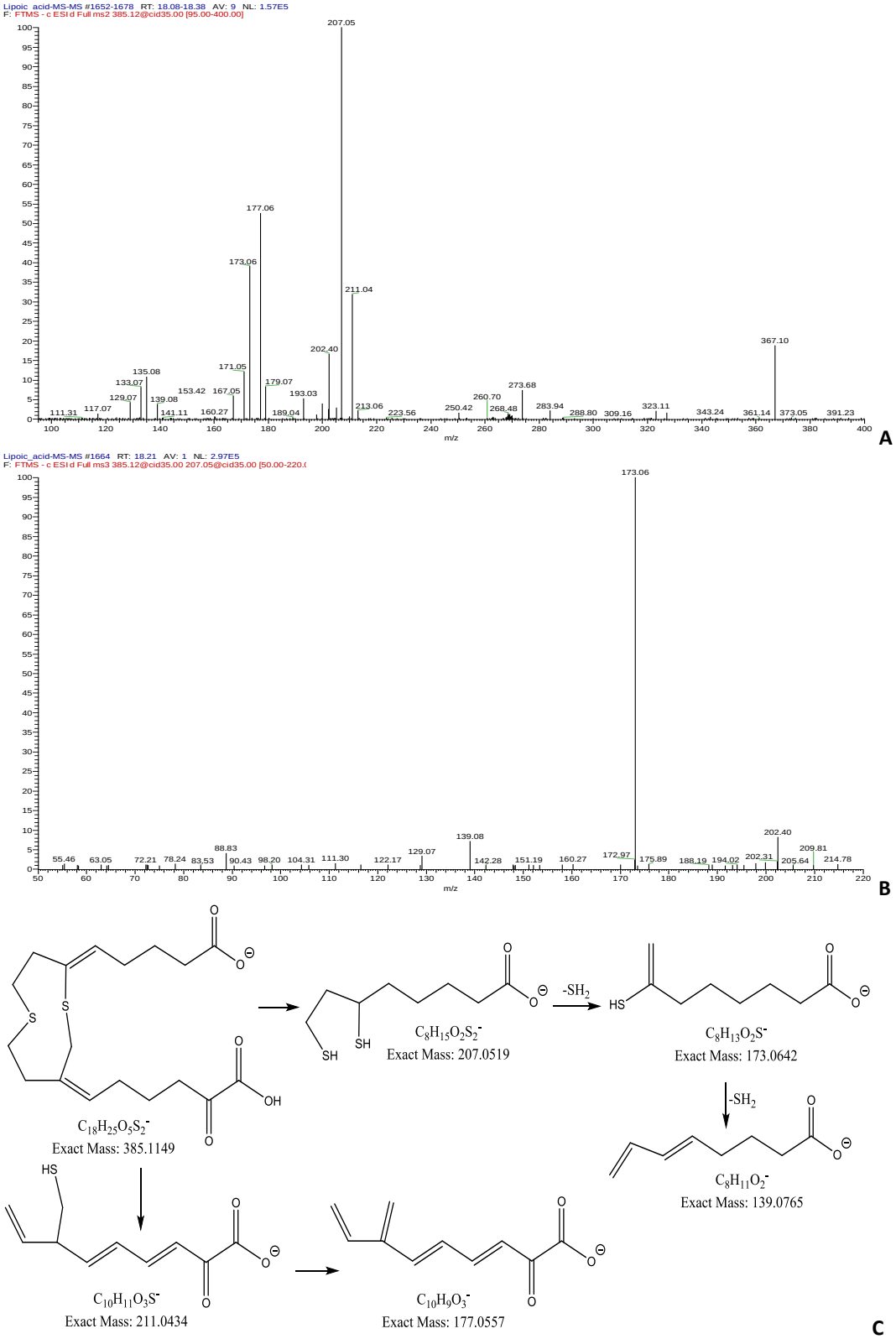


Figure 3.9 Impurity at m/z 385.1144 (A) MS^2 spectrum of m/z 385.1144 (B) MS^3 of fragment ion at m/z 207.051 (C) proposed fragmentation mechanism.

The precursor ion at 411.0791 with elemental composition ($C_{16}H_{27}O_4S_4$) is clearly shown to be a dimer of LA. The MS^2 data provided the fragment ions at m/z 207.051 and 171.048 (Figure 3.10A), of which the former is related to DHLA and the latter has elemental composition ($C_8H_{11}O_2S$), respectively (Figure 3.10B). However, on inspection of the chromatographic data this compound has the same peak shape and retention time as lipoic acid and is thus formed in the mass spectrometer (Fig. 3.11).

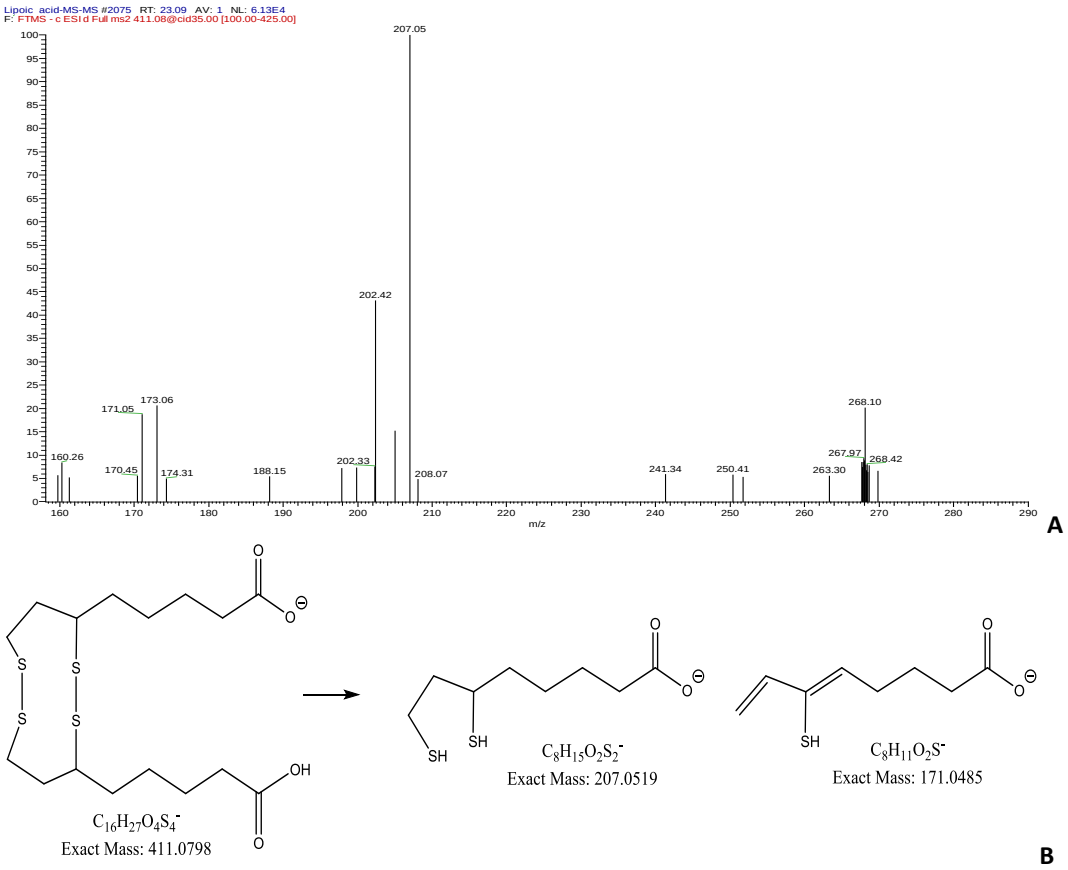


Figure 3.10 Impurity at m/z 411.0791 (A) MS^2 spectrum (B) proposed fragmentation mechanism.

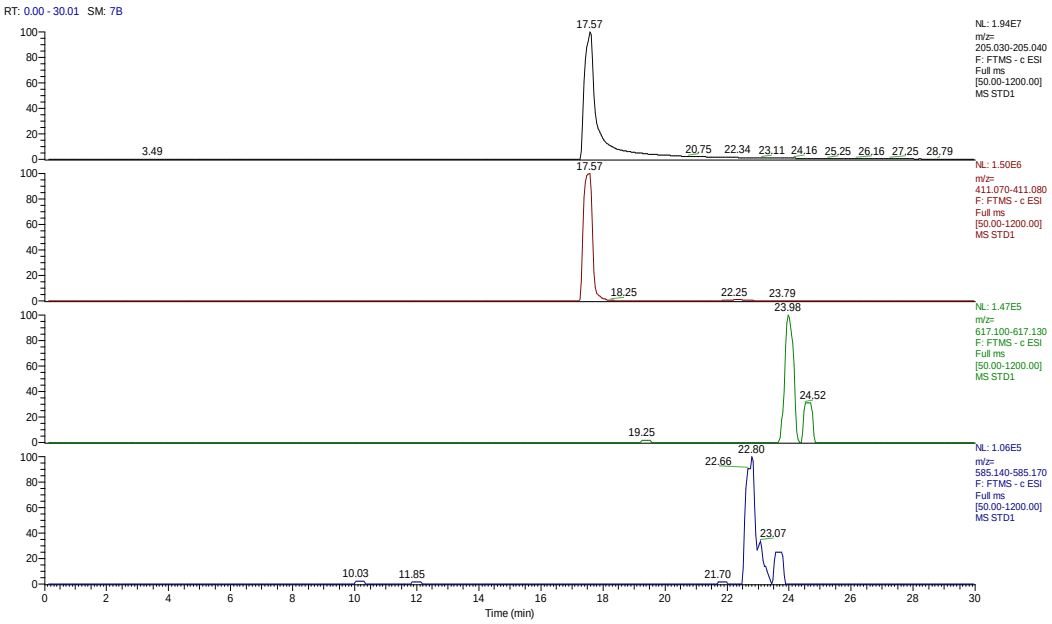


Figure 3.11 EICs for LA, LA dimer formed in the mass spectrometer and chemically formed trimers.

The trimer of lipoic acid at m/z 585.1508 is chemically formed rather than being formed in the MS since it runs at a different retention time from the lipoic acid (figure 3.11).

The deprotonated molecular at m/z 585.1508 with elemental composition ($C_{24}H_{41}O_6S_5$) fragmented into product ions at m/z 381.122 and 171.04 (Figure 3.12A). The more abundant ion is m/z 381.122 which has an elemental composition ($C_{16}H_{29}O_4S_3$) and generated the ions at m/z 363.113, 347.136, 329.125, 295.137, 251.148 and 173.064 after MS^3 (Figure 3.12B). Figure 3.12C presents the proposed structure and fragmentation pathway of the impurity at m/z 585.1508 which represents a very complete elucidation of the structure.

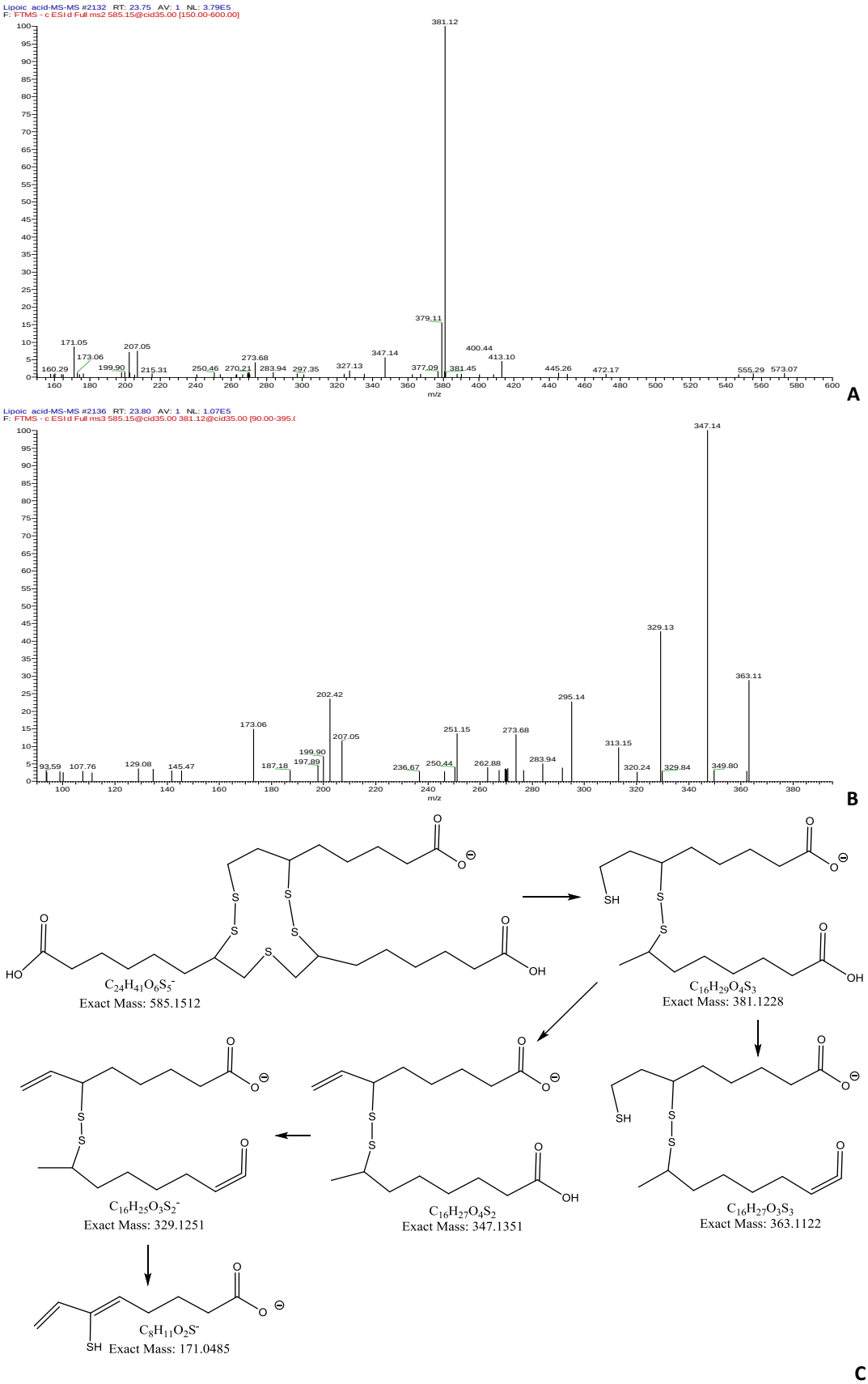


Figure 3.12 Impurity at m/z 585.1508 (A) MS^2 spectrum of m/z 585.1508 (B) MS^3 of fragment ion at m/z 381.122 (C) proposed fragmentation mechanism.

The impurity $[M-H]^-$ at m/z 617.1229 and formula ($C_{24}H_{41}O_6S_6$) is again a chemically formed trimer since it has a later retention time than lipoic acid. It produced fragment ions at m/z 207.051 and 411.080 implying that it was a trimer structure of LA (Figure 3.13A) containing a ring with three sulphur atoms. The fragment ions are correlated to the DHLA and dimer LA with elemental compositions ($C_8H_{15}O_2S_2$) and ($C_{16}H_{27}O_4S_4$), respectively. The product ion at 173.064 (Figure 3.13B) formed from MS^3 of DHLA is a result of loss of the SH_2 group. Figure 3.13 C presents the proposed structure and fragmentation pathway of the impurity at m/z 617.1229.

The above information regarding the structure elucidation for the potential impurities was collected to propose the pathways through lipoic acid impurities can be formed.

Some of the impurities are easy to understand being due to combinations of lipoic acid to from trimers of oxidation of the sulphur atoms. However, there are several impurities where the structures are purely speculative. Since the sulphur atoms in the molecule are reactive is would be possible for impurities to arise via reaction with intermediates in the synthetic process. Despite not knowing the exact identity of the impurities the impurity profiles can be used to monitor the success of the purification process. Lipoic acid is a fascinating molecule with regard to its degradation profile.

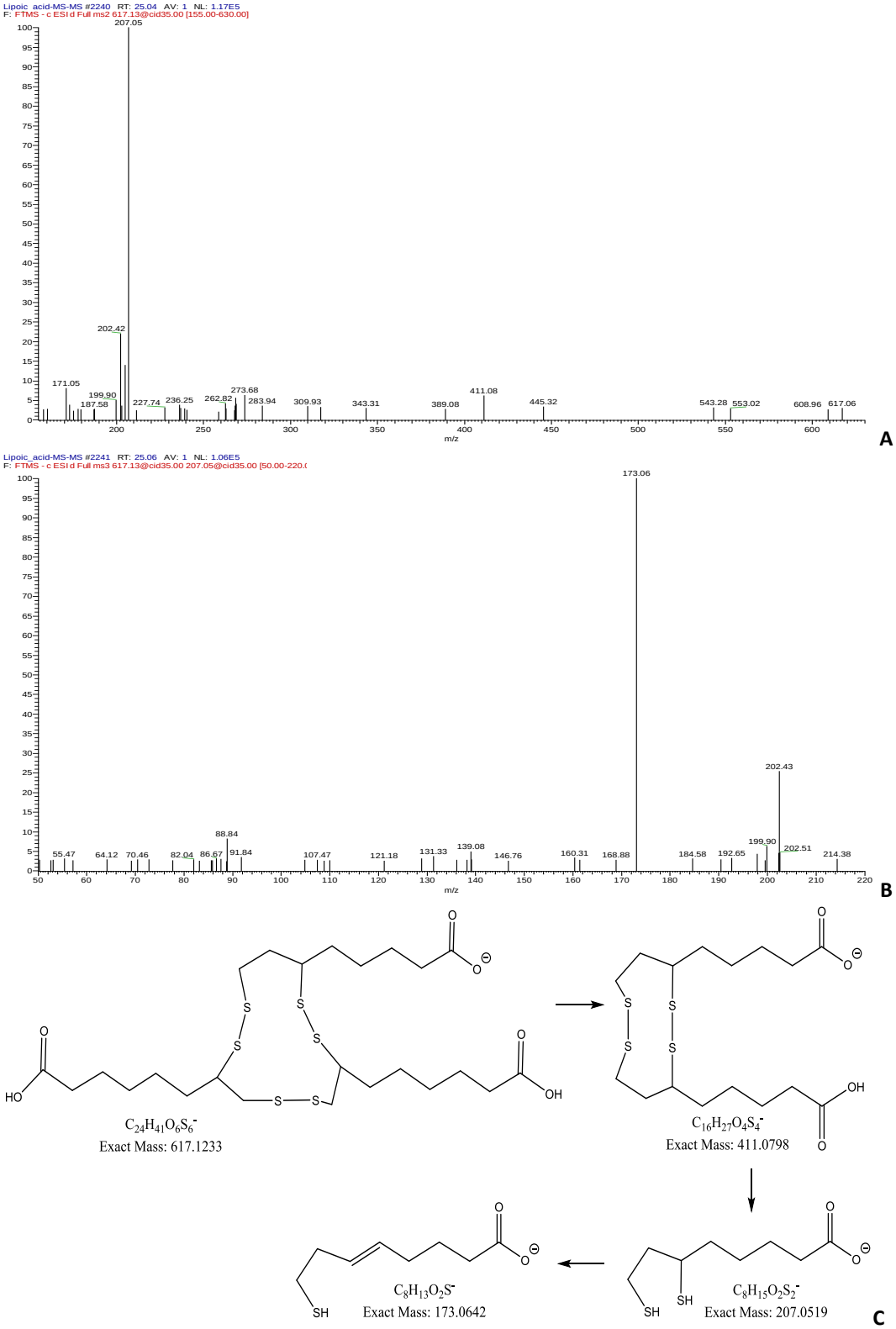


Figure 3.13 Impurity at m/z 617.1229 (A) MS^2 spectrum of m/z 617.1229 (B) MS^3 of fragment ion at m/z 207.051 (C) proposed fragmentation mechanism.

3.3.2 Multivariate Analysis of the impurity profiles

Before the data visualization, MZmine was first employed to process the LC-HRMS data through a series of steps involving peak detection, mass detection, chromatogram deconvolution, deisotoping, normalisation, alignment and gap filling. At the end of the processing step, a total of 59 different components were extracted from more than 450 detected peaks. SIMCA-P+ 14 software (Soft Independent Modelling of Class Analogy) was used as a statistical classifying method to observe both within and between class variances of the impurity profiles.

3.3.2.1 Principal Components Analysis (PCA)

PCA is a procedure that applies an orthogonal translation to statistically convert the correlated variables of the observations into principal components (PCs), which are a set of values of linear uncorrelated variables. A scores plot is used to display the pattern of observations (material) due to the differences and similarities expressed by their variables (impurities). The scores plot obtained by PCA shows distinct clusters for the LA raw material and the other three purified materials (Figure 3.15). The first principal component (PC1), which is a direction of maximum variance of the whole data set, described 49.3% of the variation in the data. PC2, which is the second orthogonal linear direction of maximum variance in the data which is unexplained by PC1, described 19.4% of the total variation. In total, the three PCs described 84.4% of the whole variation in the data (PC1: 49.3%, PC2:19.4% and PC3: 15.7%).

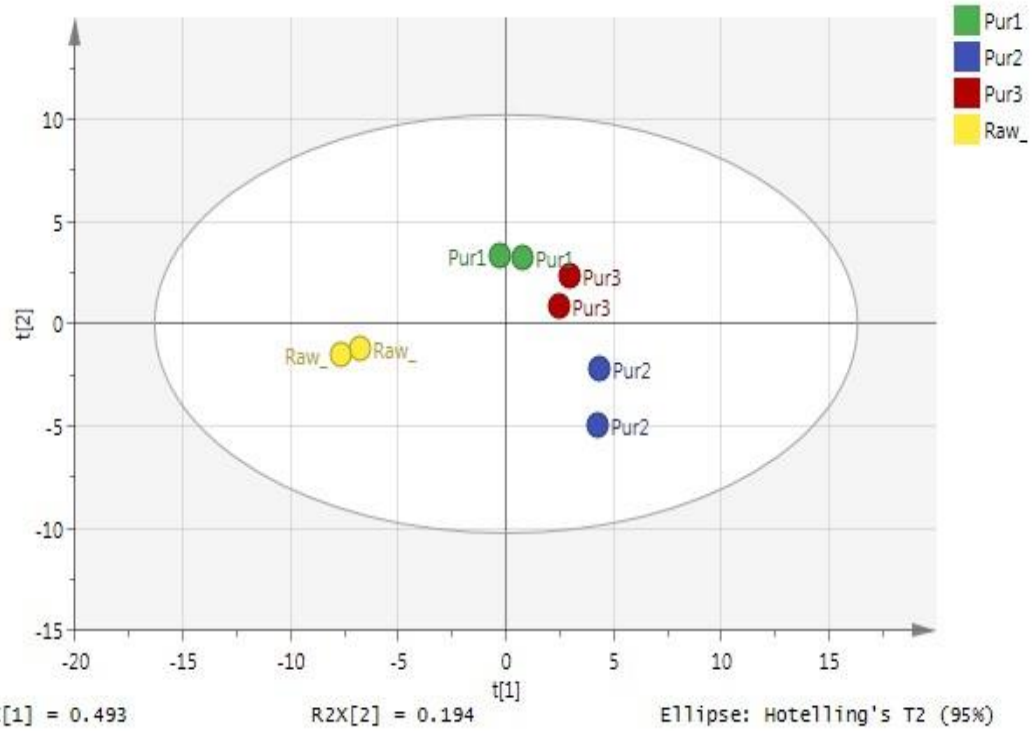


Figure 3.14 PCA scores scatter plot t1 vs. t2 for lipoic acid materials (PC1 versus PC2). (Yellow; raw material, Green; Purified material-1, Red; purified material-2 and Blue; purified material-3) obtained by using Pareto scale. Three PCs described 84.4% of the whole variation data (PC1: 49.3%, PC2:19.4% and PC3: 15.7%).

3.3.2.2 Hierarchical clustering analysis

A hierarchical clustering analysis (HCA) uses the total variation in the dataset to provide a natural clustering of observations without prior knowledge (Beebe et al., 1998). The HCA shows that the four materials cluster into two main groups with less marked differences between subgroups. Inter individual variability (i.e. the outcome of variable intensity happening in the different individuals at a specific material) has a main role in the classification pattern of group 1, which belongs to LA raw material, and group 2, which is related to three LA purified materials (Figure 3.16). The plot classified the samples into two distinct hierarchical groups with a similarity value of 120. The similarity value (y-axis) is expressed as the measure of the degree of closeness between the individual data points or groups.

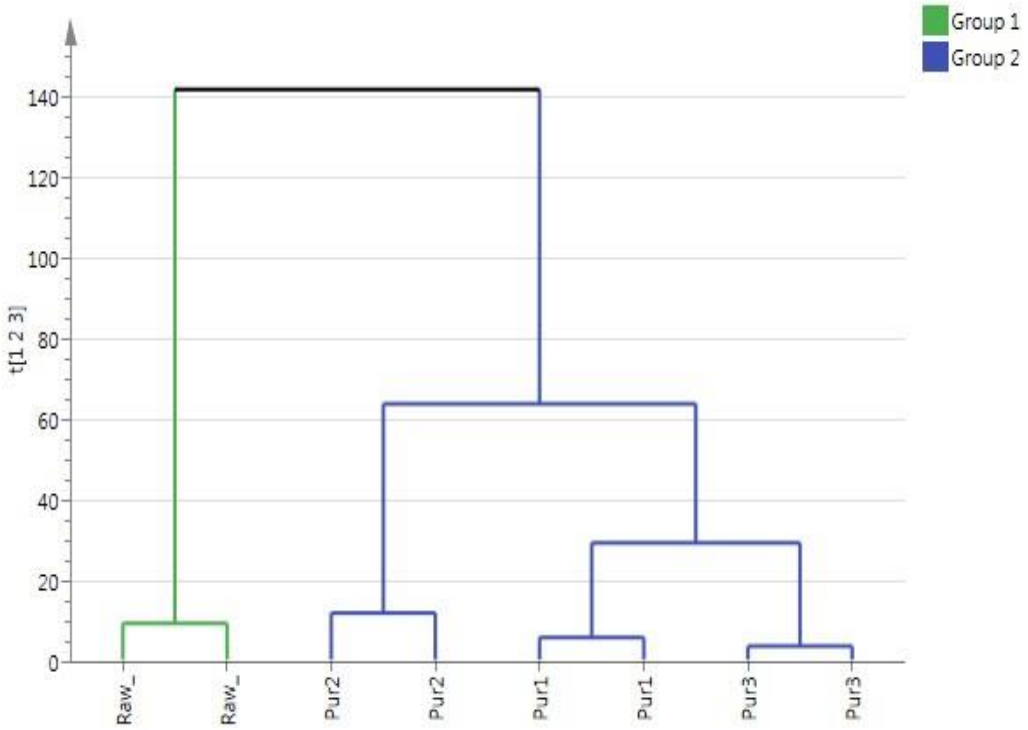


Figure 3.15 Hierarchical Clustering Analysis (HCA) plot for the four materials of lipoic acid. The plot classified the samples into two distinct hierarchical groups by using similarity value at 120 and Ward clustering method (vertical axes); (group 1 = green), (group 2 = blue).

3.3.2.3 OPLS-DA

The PCA scores plot (Figure 3.15) shows a clear unsupervised clustering pattern. The separation between groups based on a type of intervention can be perceived by characterizing the multivariate model using supervised OPLS-DA analysis. The model shows a high value of R^2(cum) = 0.994 (i.e. $R^2 > 0.5$) which means the OPLS-DA model has a high degree of fit and therefore most of the variance related to the response variable (Y) can be explained. In addition, the high value of Q^2(cum) = 0.976 (i.e. $0.9 \leq Q2 \leq 1$) means that the model has a high degree of consistency between the original and predicted data.

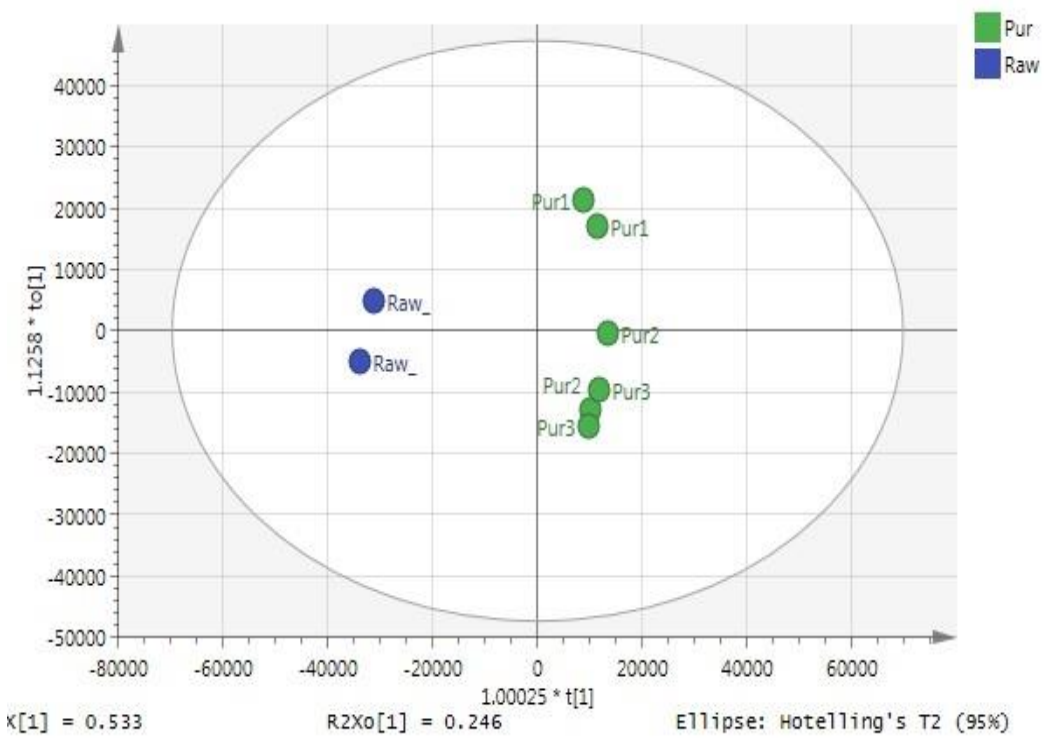
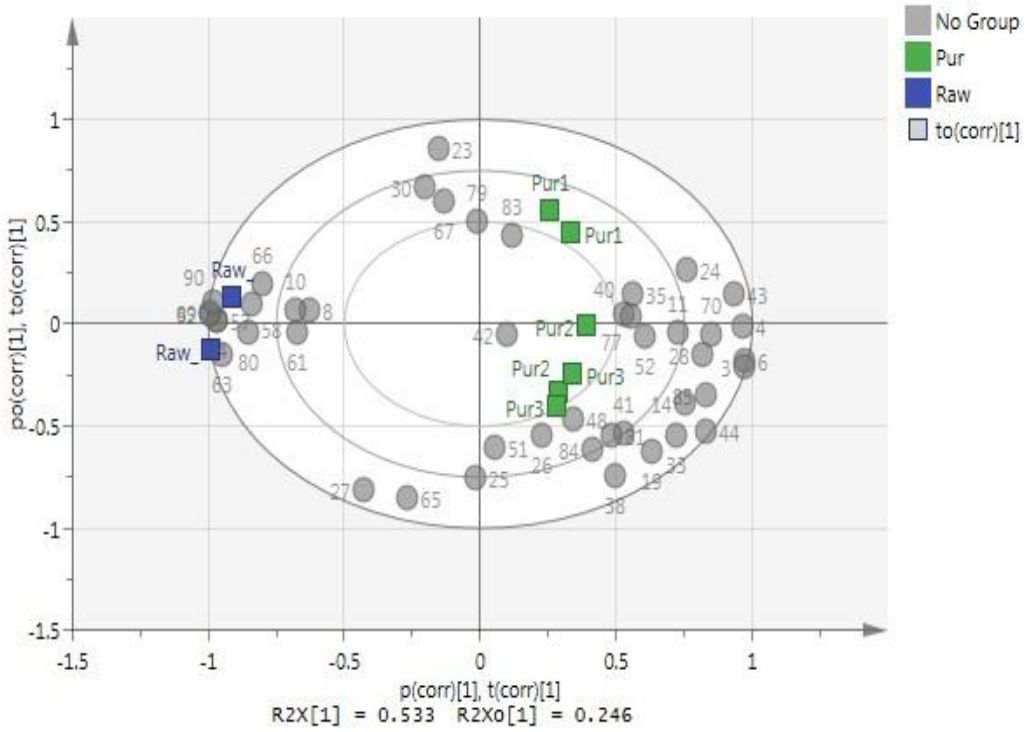
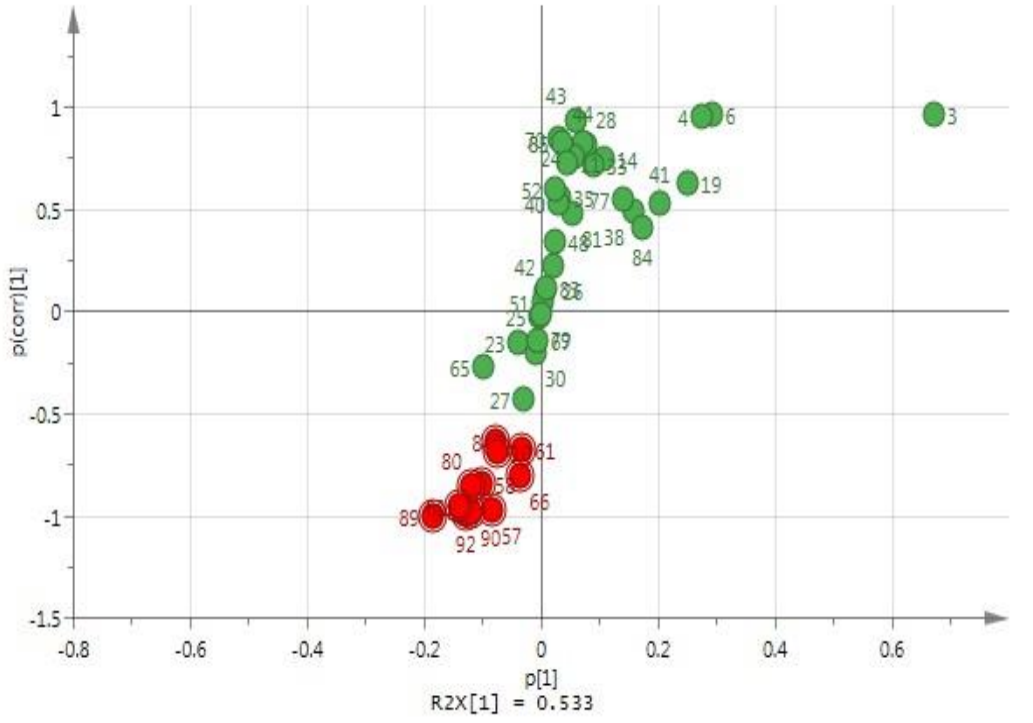


Figure 3.16 OPLS-DA score plot shows the raw material (Blue) vs three LA purified materials (Green).

The OPLS-DA scores plot (Figure 3.16) shows that there are clear differences which can be observed between the raw and purified materials. It means that there is a significant difference between these two groups according to their variables (i.e. impurity profiles). For investigating the impurities which bring about the difference between the two groups, Biplot and S-plot were applied to illustrate the major contributing variables.



A



B

Figure 3.17 (A) Biplot; scores (materials) and loadings (impurity) with low discrimination power on PC1 vs. PC2. (B) S-plots at the first component.The red marked points have high influence on the raw material.

A biplot (Figure 3.17A) shows which of the variables are more highly correlated to a given group of observations than to the others; this can also be confirmed by an S-plot. The highlighted variables in the S-plot (red points in figure 3.17B) represent the most relevant impurities which contribute towards differentiation of the raw material from the other purified materials. These potential impurities were investigated to confirm differences in their real intensity levels by considering the p-values (< 0.05) and 95% CIs (Table 3.3). As a result, the significant selected variables clearly showed higher intensity levels in the raw material than in the purified materials group (Figure 3.18). This might explain the change in the material's performance during the crystallization process. The variables include some of impurities characterised or partly characterised as discussed above but there are also some additional ones which were not considered above.

Table 3.3 Mass to charge ratio, retention time (min) and P-value of impurities have contributed in the change of the raw material profile.

m/z	RT (min)	P value	95% CI				
272.985	20.73	0.02207	(	0.12	,	0.47	)
305.089	15.42	6.74E-05	(	0.76	,	1.20	) Fig. 3.7
313.148	19.87	8.86E-08	(	1.13	,	1.72	)
333.084	15.23	3.78E-05	(	0.50	,	0.78	) Fig. 3.8
345.12	21.81	1.08E-07	(	0.75	,	1.20	)
347.136	18.9	0.008684	(	0.50	,	1.06	)
349.125	20.68	2.07E-05	(	0.73	,	1.09	)
585.152	22.3	0.005702	(	0.57	,	1.36	) Fig. 3.12
617.124	23.7	0.00014	(	0.76	,	1.47	) Fig. 3.13

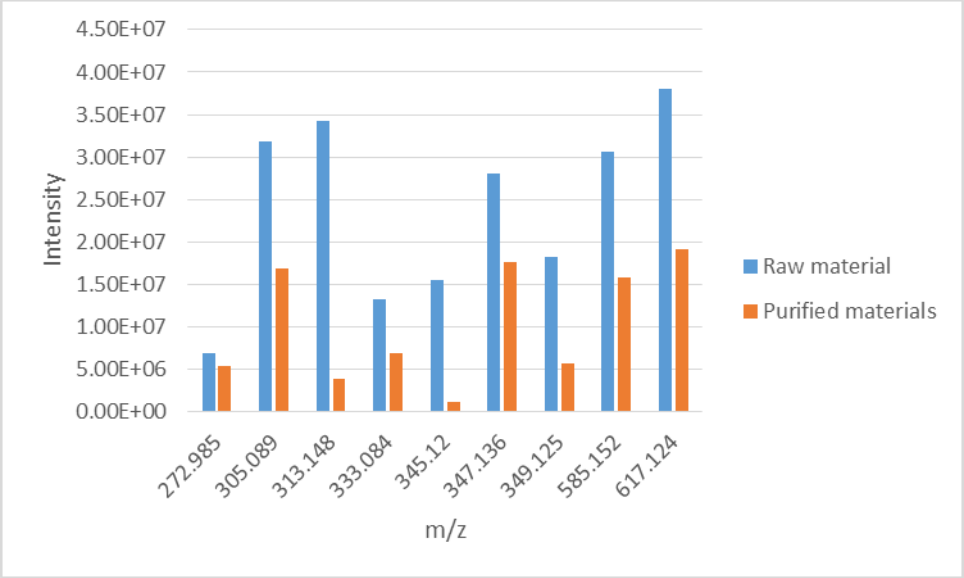


Figure 3.18 Measurement the intensity level of impurities which are selected by S-plot. The impurities show superior high intensity in the raw material group rather than the purified materials group. The values for impurities in the purified materials were the mean of three samples.

Figure 3.18 shows that there was clearly a reduction in the intensity of some of the impurity peaks in the purified samples although they were not completely removed. The supervised classification model was validated by using permutation test as a cross-validation method to optimise the model fitting and to assess the performance parameters. As shown in the permutation test (Figure 3.19), the y-axis gives the value of R^2 = (0.0, 0.56) and Q^2 = (0.0, -0.856) which is related to the goodness of fit and prediction, respectively. The x-axis characterises the correlation between the permuted y-vectors (on the left) and the original y-vector (on the right). The model could be considered statistically significant since the regression line of the Q^2 points intersects the y-axis below zero. In addition, the Receiver Operator Characteristic (ROC) curve shows the robustness of the OPLS-DA model obtained from the cross-validated predicted Y (predicted class). It shows that (AUC = 1.0) which means the prediction model of OPLS-DA has great sensitivity and specificity (i.e. AUC > 0.5 corresponds to a good predictive model) (Figure 3.20).

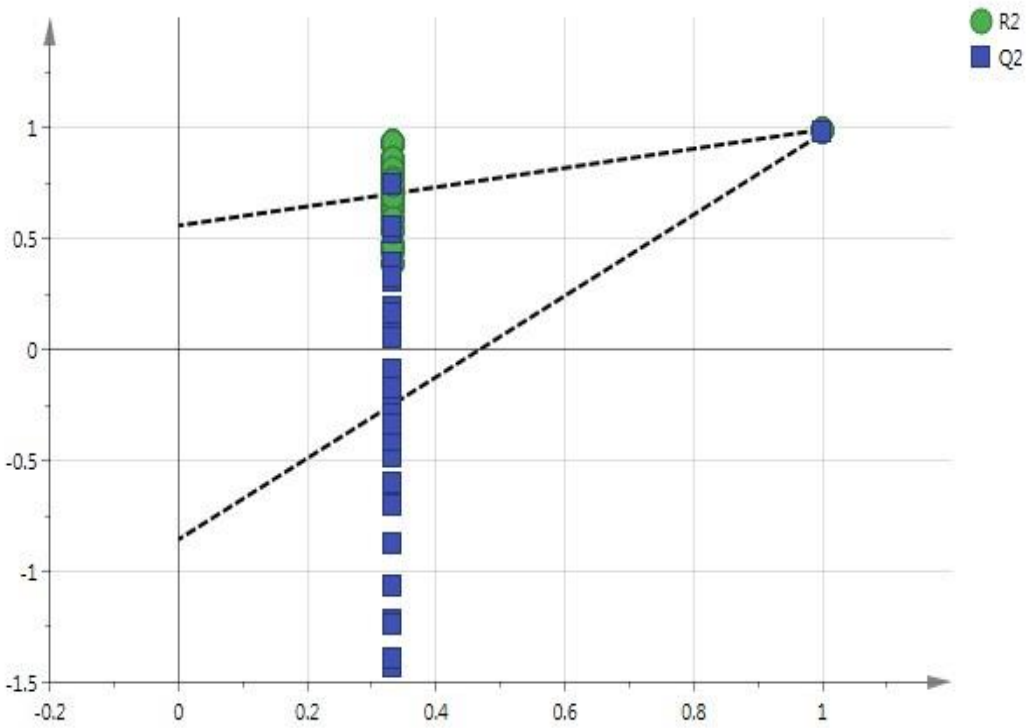


Figure 3.19 Validation plot obtained from the permutation test.
Validation plot obtained from the permutation test. R^2 = (0.0, 0.56), Q^2 = (0.0, -0.856) and the permutation number (n = 999).

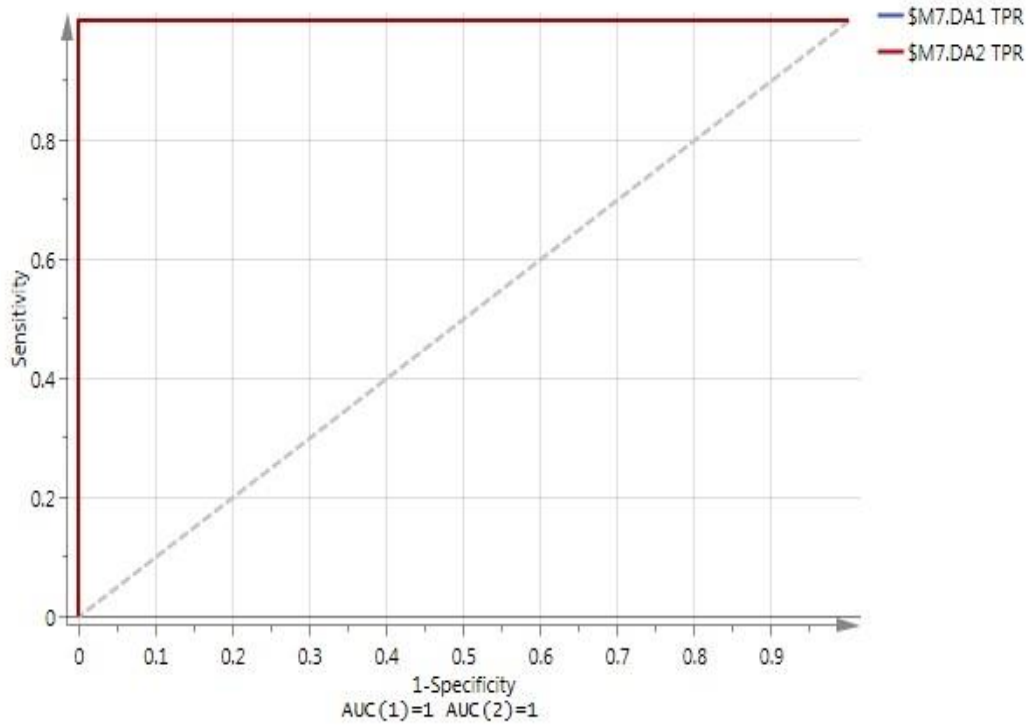


Figure 3.20 ROC curve calculated from the cross-validated predicted Y-values of the OPLS-DA model.

3.4 Conclusion

This work shows that the impurity profile of the LA raw material is very similar to, but quite distinguishable from, those of the three purified materials. The observed inter material variability in the impurity profiles might be an effect of the purification process. However, is very clear that the raw material is well separated from all the three purified materials, implying that the purification process causes significant reduction in the levels of certain impurities, particularly trimers of LA (with 5 or 6 S atoms) and LA with extra fatty acid chains with 5 or 6 C atoms. Thus these impurities probably interfere with the crystallisation of LA and their removal through purification improves the crystallisation process. The approach to assessing the success of a purification protocol might be applicable to other raw materials.

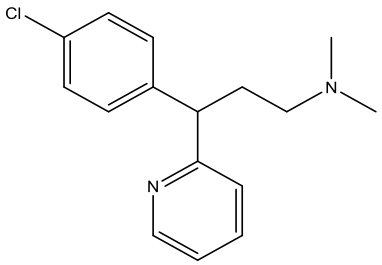
Chapter 4

Impurity profiling of chlorpheniramine maleate by high resolution mass spectrometry using orthogonal HPLC methods

4 Impurity profiling of chlorpheniramine maleate by high resolution mass spectrometry using orthogonal HPLC methods

4.1 Introduction

Chlorpheniramine malate is a first generation histamine H1 (antihistaminic) receptor antagonist. It has been widely used as an over-the-counter medication in relieving symptoms of the common cold and allergic symptoms such as hay fever and eye inflammation (Nia, Fuhr et al. 2010). Chlorpheniramine (USAN, former BAN) or Chlorphenamine (INN), (3-(4-Chlorophenyl)-N,N-dimethyl-3-(pyridin-2-yl)propan-1-amine) (Figure 4.1) and chlorpheniramine its molecular weight is 274.788 g/mol and the molecular formula is $C_{16}H_{19}N_2Cl$.



Exact Mass: 274.1237
m/z: 274.1237 (100.0%), 276.1207 (32.0%), 275.1270 (17.3%), 277.1241 (5.5%), 276.1304 (1.4%)

Figure 4.1 Structure, exact mass of chlorpheniramine and isotopic abundances.

Organic impurities from the manufacturing process may include starting materials, by-products, intermediates, degradation products, reagents, ligands and catalysts or may arise during storage under unappropriate conditions (ICH-Q3A R2 (2002)).

Both reversed phase high performance chromatography (RP-HPLC) and hydrophilic interaction chromatography (HILIC) techniques can be used in-line with high resolution mass spectrometry for screening drug impurities. RP-HPLC has high specificity that is capable of detecting and separating known, unknown, and potential impurities. HILIC provides a further dimension that offers an alternative retention mechanism and selectivity which makes it a useful technique for orthogonal method development (Wang, Li et al. 2005). Orthogonal methods can afford a comprehensive

investigation of related substances giving a high degree of confidence that all impurities are covered (Argentine, Owens et al. 2007).

Principal component analysis (PCA) can be generated to provide data visualization of LC/MS data in order to offer valuable information for clustering and classification (Marhuenda-Egea, et al. 2013). Orthogonal partial least squares-discriminant analysis (OPLS-DA) is used as a supervised data modelling technique in order to develop the discrimination. It is directed by a dependent variable(s) that specifies the grouping of the samples and to give valuable information about the material characterisation (Krakowska et al. 2016). The method has been employed in drug impurity profiling in order to differentiate among the several synthetic routes and manufacturers (Yuan et al. 2012).

Several methods have been used for the synthesis of chlorpheniramine. It can be synthesised by starting with 2-chloropyridine and 4-chlorobenzyl cyanide in the presence of $NaNH_2$. The reaction forms an intermediate compound of 4-chlorphenyl(2-pyridyl)acetonitrile which is then alkylated with 2-dimethylaminoethylchloride to produce γ-(4-chlorphenyl)-γ-cyano-N,N-dimethyl-2-pyridine-propanamine. The last step is a decarboxylation which follows from hydrolysis with H_2SO_4 leading to chlorpheniramine (Figure 4.2) (Domenick, et al. 1951). The another route is starting from alkylation of pyridine with 4-chlorobenzylchloride to form 2-(4-chlorobenzyl)pyridine which reacts with 2-dimethylaminoethylchloride in the presence of $NaNH_2$ to produce chlorpheniramine (Figure 4.3) (Domenick, Erwin et al. 1954).

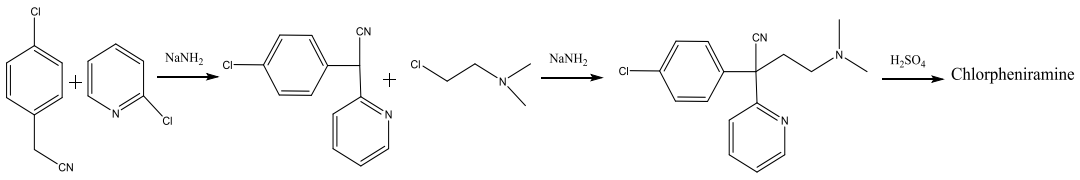


Figure 4.2 Synthetic scheme for chlorpheniramine (route 1) (from Nichols, 2006, p. 225).

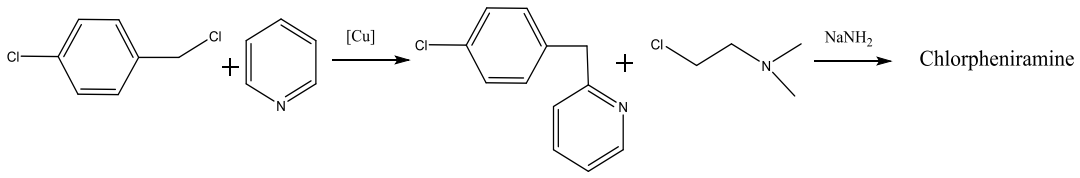


Figure 4.3 Synthetic scheme for chlorpheniramine (route 2) (from Nichols, 2006, p. 225).

A third route for the synthesis of chlorpheniramine starts from the reaction between pyridine-2-aldehyde and 4-chlorophenylmagnesium bromide resulting in the preparation of a tertiary alcohol. After reduction of the alcohol, dimethylamine ethyl chloride is used for alkylation in presence of sodamide to produce the chlorpheniramine (Figure 4.4) (Sriram and Yogeeswari 2010).

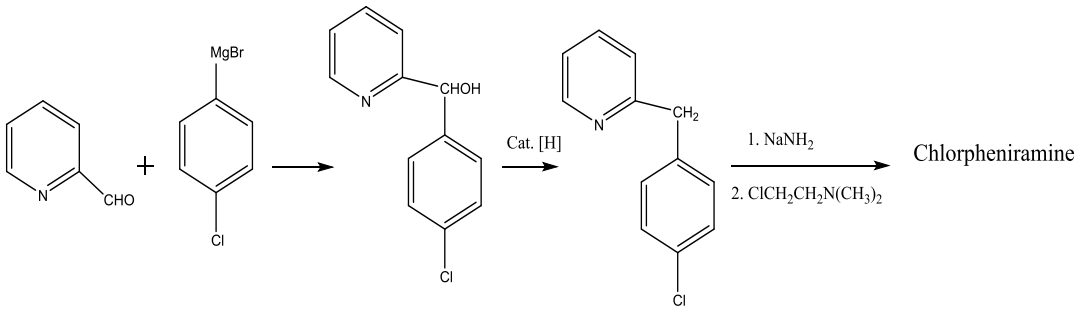
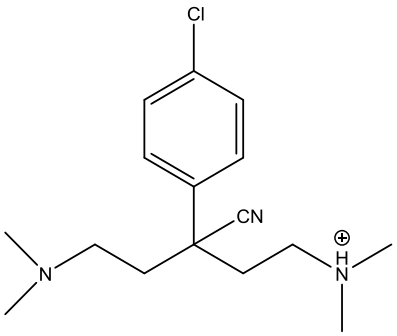


Figure 4.4 Synthetic scheme for chlorpheniramine (route 3) (Sriram and Yogeeswari 2010).

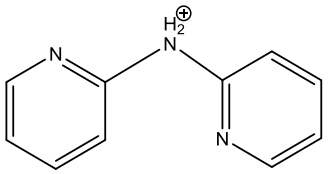
The impurity pattern of samples can give an indication in of the synthetic pathway used to make the drug.

Figure 4.5 shows the four impurities listed in the European Pharmacopoeia for chlorpheniramine in their protonated form. Impurities A, B and D would be expected for arise in the course of the synthetic route (1) shown in figure 4.2. Impurity C could arise from an impurity in the dimethylamine ethyl chloride and could arise from any of the routes.



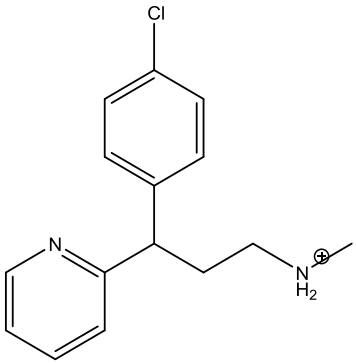
Chemical Formula: $C_{16}H_{25}ClN_3^+$
Exact Mass: 294.1732

Impurity A



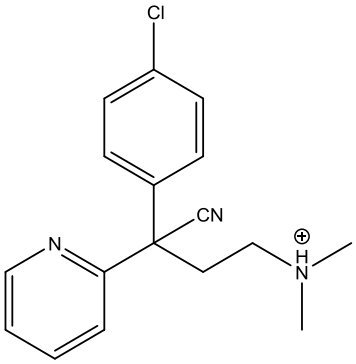
$C_{10}H_{10}N_3^+$
Exact Mass: 172.0869

Impurity B



$C_{15}H_{18}ClN_2^+$
Exact Mass: 261.1153

Impurity C



$C_{17}H_{19}ClN_3^+$
Exact Mass: 300.1262

Impurity D

Figure 4.5 Structure of protonated molecular ions of impurities A, B, C and D according to chlorpheniramine maleate monograph (Europe 2017).

In this work, four chlorpheniramine maleate drug substances were investigated for the purpose of impurity profiling using both RP-HPLC and HILIC techniques hyphenated to HRMS and MS^2. The data were investigated to give information on the proposed structure of impurities and degradation pathways of the samples.

Drug impurity analysis achieved by LC/MS can offer semi-quantitative results about the concentration of impurities by comparing the peak areas of the impurities with the peak area of the API isotope peak. In addition, the synthetic routes were proposed based on the classification and discrimination information from SIMCA.

4.2 Methodology

4.2.1 Chemicals and Reagents

The chlorpheniramine drug substances were obtained from Sigma-Aldrich (UK) and FLUKA (UK) (Table 4.1).

Table 4.1 Samples of chlorpheniramine maleate.

Sample	Abbr.	Specification
Chlorpheniramine Maleate	CPAM1	>98.5%
(±)-Chlorpheniramine Maleate	CPAM2	≥99% (perchloric acid titration)
S-(+)-Chlorpheniramine Maleate	CPAM3	Purity (TLC) > 98 %
(±)-Chlorpheniramine Maleate	CPAM4	Pharmaceutical secondary standard; traceable to USP, PhEur and BP

4.2.2 Sample preparation

Samples were prepared at a concentration of 0.5 mg/mL in methanol.

4.2.3 Hydrophilic interaction chromatography (HILIC) conditions

In the HILIC system, the silica gel column used was Kromasil 60-5-SIL column (150 × 4.6 mm, 5 µm) from Hichrom (Reading UK). Isocratic separation was carried out with mobile phase (0.01 M aqueous ammonium acetate) /ACN in a 10:90 (v/v) ratio and a flow rate of 0.4 ml/min for 30 min.

4.2.4 Reversed Phase liquid chromatography (RP-LC) conditions

In the RP-HPLC system, the reversed phase column used was ACE C18-AR column (150 × 4.6 mm, 5 µm) from HiChrom Reading UK. Isocratic separation was carried out with mobile phase A (0.1% v/v formic acid in water) and B (0.1% v/v formic acid in ACN) in a 40:60 (v/v) ratio and a flow rate of 0.4 ml/min for 30 min.

4.2.5 Liquid chromatography- Mass spectrometry

Refer to section 2.3.

4.2.6 Liquid chromatography- MS^n mass spectrometry

Refer to section 2.4. Protonated molecular ions ($[M+H]^+$) were applied for detecting the molecular weights.

4.3 Results and discussion

4.3.1 Orthogonal methods

In this study, four chlorpheniramine standard materials from different manufactures were investigated (Table 4.1). Orthogonal methods used were RP-LC and HILIC systems each hyphenated to a high resolution mass spectrometer with high sensitivity for the purpose of comprehensive characterisation of chlorpheniramine impurity profile. The chromatograms for both HILIC and RP-HPLC systems were generated using the method conditions described in section 4.2.3 and 4.2.4. The overlayed total ion current (TIC) traces for the samples is shown in (Figure 4.6) for RP-LC and (Figure 4.7) for HILIC. It is clear for the HILIC trace that the HILIC column has a poorer capacity than the RP column since the peak of the overloaded chlorpheniramine is very broad.

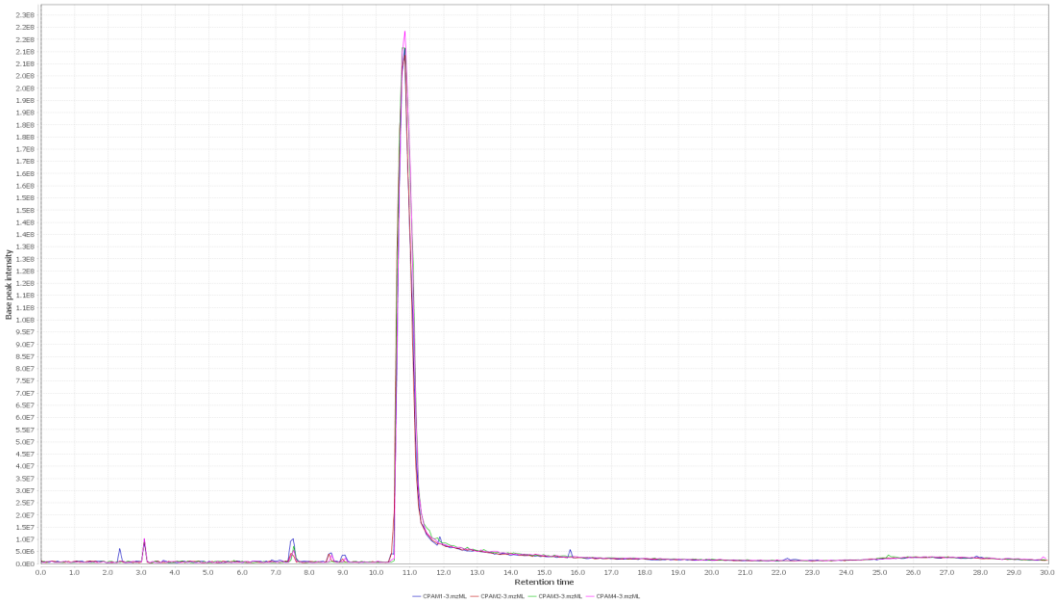


Figure 4.6 Overlapped RP-LC total ion current chromatograms of chlorpheniramine samples 1, 2, 3 & 4 characteristics by MZmine 2.20.

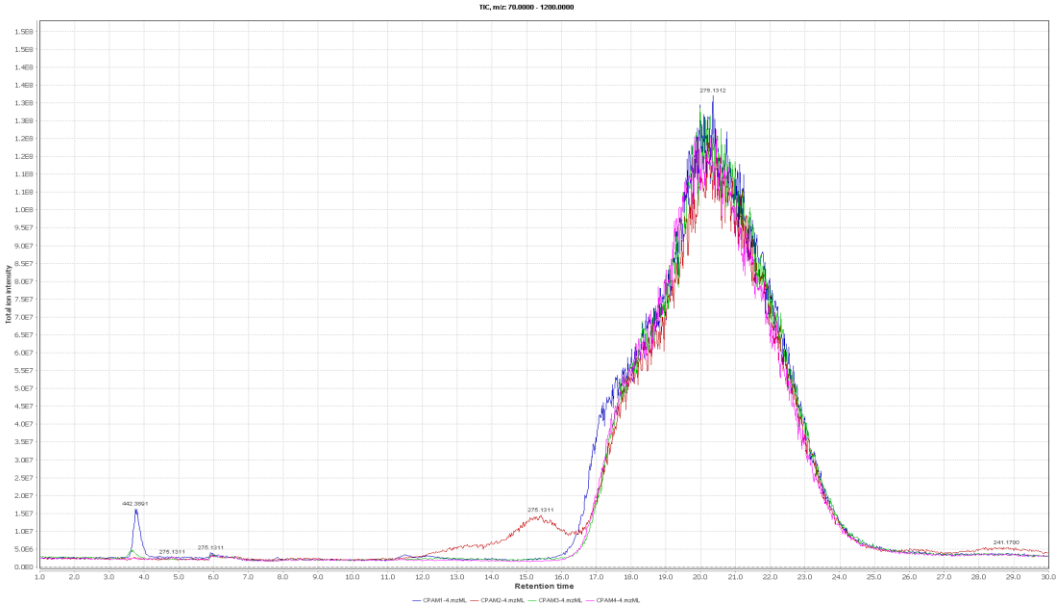


Figure 4.7 Overlapped HILIC total ion current chromatograms of chlorpheniramine samples 1, 2, 3 & 4 characteristics by MZmine 2.20.

The LC/MS data were captured in Xcalibur data format. These data acquired from Xcalibur software were used to obtain the empirical formulae. The data for exact masses of the molecular ions with an error margin of less than 3ppm combined with information from nitrogen rule, Relative double bonds (RDB), and isotopic patterns, were used for interpretation. In addition, adduct formation was investigated for the purpose of confirming the impurity molecular ion in some cases.

The chromatogram and mass spectrometry data were analysed for each sample. Moreover, relative semi-quantitative measurement was applied to estimate the amount of impurity by comparison of its peak to the main isotope peak of API (the isotope peak for the API was on scale). The LC/MS acquired data for the impurities in each sample are listed in the Table 4.2 for the HILIC data and Table 4.3 for the RP-HPLC data. Impurities, B and D can be observed in varying amounts in the samples. Figures 4.8 and 4.9 show extracted ion traces for these compounds in each of the samples. Impurity C although picked up by the data extraction software has the same broad shape as chlorpheniramine and thus is probably due to a fragment of the drug formed in the mass spectrometer. From figures 4.8 and 4.9 it can be seen that CPAM1

and CPAM4 are synthesised via route 1 since only this route can produce impurities B and D. This looks less likely for samples 2 and 3.

Figures S4.1 and S4.2, appendix, show the confirmatory MS^2 spectra for impurities C and D.

Table 4.2 List of impurity of chlorpheniramine samples from HILIC analysis including mass to charge ratio, RT, proposed formula and the relative quantity percentage (%).

ID	m/z	RT	Formula	Relative content %			
				CPAM1	CPAM2	CPAM3	CPAM4
Imp B	172.0869	5.26	$C_{10}H_{10}N_3$	0.0157	0.0001	0	0.0006
UK1	198.1040	8.26	$C_{11}H_{17}NCl$	0.0107	0.3266	0.0004	0.0001
UK 2	223.0996	12.05	$C_{12}H_{16}N_2Cl$	0.0767	0.0081	0.0104	0.0097
UK 3	241.1697	28.5	$C_{16}H_{21}N_2$	0.4792	1.5804	0.439	0.2656
Imp C	261.1154	18.42	$C_{15}H_{18}N_2Cl$	0.0399	0.2052	0.0379	0.005
UK 4	281.0841	3.69	$C_{17}H_{14}N_2Cl$	0.0003	0.0017	0	0.0104
UK 5	289.1463	12.33	$C_{17}H_{22}N_2Cl$	0.0177	0.1311	0.0298	0.012
UK 6	291.1256	12.82	$C_{16}H_{20}ON_2Cl$	0.0591	0.612	0.0447	0.0008
Imp D	300.1244	11.56	$C_{17}H_{19}N_3Cl$	0.1429	0.0007	0	0.0175
UK 7	309.0927	26.46	$C_{16}H_{19}N_2Cl_2$	0.0186	0.1994	0	0.0371
UK 8	297.0789	4.54	$C_{17}H_{14}ON_2Cl$	0.0089	0.0011	0.0000	0.0124

Table 4.3 List of impurity of chlorpheniramine samples from RP-LC analysis including mass to charge ratio, RT, proposed formula and the relative quantity percentage (%).

ID	m/z	RT	Formula	Relative content %			
				CPAM1	CPAM2	CPAM3	CPAM4
	99.0077	3.08	$C_4H_3O_3$	0.3694	0.4045	0.4171	0.4253
	117.0183	3.08	$C_4H_5O_4$	0.9462	1.021	1.0556	1.0962
	121.0886	7.52	$C_8H_{11}N$	0.2588	0.0532	0.0825	0.073
	166.598	9.05		0.6904	0.1115	0	0.1259
	196.1119	7.52	$C_{14}H_{14}N$	0.4081	0.1302	0.1432	0.1258
UK 1	198.1042	11.89	$C_{11}H_{17}NCl$	1.3074	0.0049	0	0.0394
	241.1101	8.64	$C_{12}H_{18}ON_2Cl$	1.0308	0.5599	0.0008	0.4077
UK 3	241.1698	7.52	$C_{16}H_{21}N_2$	2.1612	0.7709	0.8949	0.75
Imp C	261.1152	10.53	$C_{15}H_{18}N_2Cl$	0.4299	0.0454	0	0
UK 5	289.1463	11.82	$C_{17}H_{22}N_2Cl$	0.2848	0.0513	0	0.038
UK 6	291.1255	11.3	$C_{16}H_{20}ON_2Cl$	0.6579	0.0366	0.0011	0.046
	304.2996	23.53	$C_{21}H_{38}N$	0.1714	0.1464	0.1326	0.0975
UK 7	309.0923	15.3	$C_{16}H_{19}N_2Cl_2$	0.07	0.343	0	0.0698

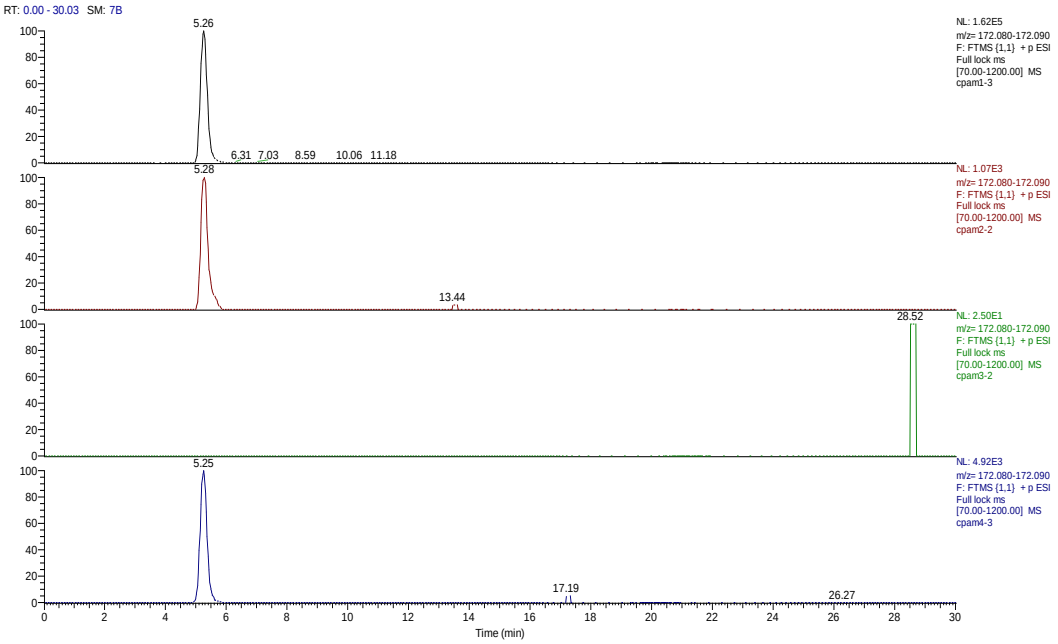


Figure 4.8 EICs for impurity B in CPAM 1-4

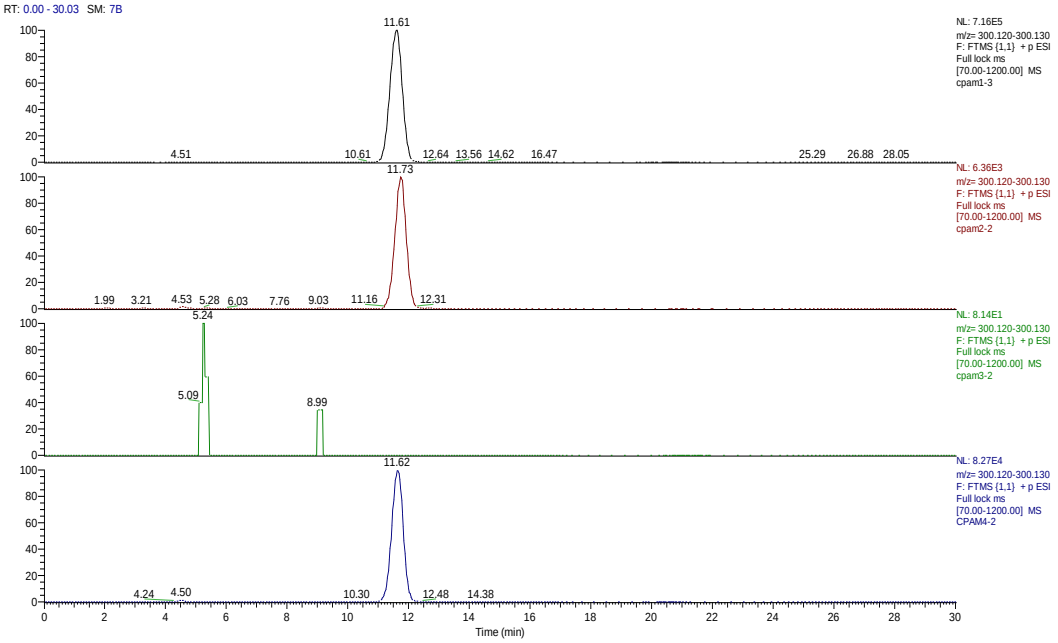


Figure 4.9 EICs for impurity D in CPAM 1-4

Table 4.4 summarises the MS^2 data for this compound.

Table 4.4 The accurate m/z, possible elemental composition and MS^2 fragments of chlorpheniramine impurities.

ID	MS				MS^2				
	m/z	Formula	RDB	ppm	m/z	Fragments	RDB	ppm	Note
Imp B	172.086	$C_{10}H_{10}N_3$	7.5	0.21					Impurity B (BP)*
UK 1	198.105	$C_{11}H_{17}NCl$	3.5	1.04	125.015	C_7H_6Cl	4.5	-0.04	Absence of pyridine ring
					153.047	$C_9H_{10}Cl$	4.5	-0.03	
UK 2	223.099	$C_{12}H_{16}N_2Cl$	5.5	0.08	207.068	$C_{11}H_{12}N_2Cl$	6.5	-0.79	Alpha-cleavage
					178.042	$C_{10}H_9NCl$	6.5	0.21	
	230.073	$C_{14}H_{13}NCl$	8.5	0.29	202.042	$C_{12}H_9NCl$	8.5	-0.61	
					167.073	$C_{12}H_9N$	9	-0.42	
					180.081	$C_{13}H_{10}N$	9.5	-0.37	
UK 3	241.17	$C_{16}H_{21}N_2$	7.5	-0.4	196.112	$C_{14}H_{14}N$	8.5	-1.41	Pheniramine
Imp C	261.115	$C_{15}H_{18}N_2Cl$	7.5	0.56	230.073	$C_{14}H_{13}NCl$	8.5	-1.67	Impurity C (BP)*
UK 4	281.084	$C_{17}H_{14}N_2Cl$	11.5	-0.29	169.076	$C_{11}H_9N_2$	8.5	0.15	
UK 5	289.146	$C_{17}H_{22}N_2Cl$	7.5	-0.53					$[M+H+CH_3]^+$
UK6	291.126	$C_{16}H_{20}ON_2Cl$	7.5	0.7	202.042	$C_{12}H_9NCl$	8.5	-1.65	
					246.068	$C_{14}H_{13}ONCl$	8.5	-0.6	
					167.073	$C_{12}H_9N$	9	-1.08	
Imp D	300.126	$C_{17}H_{19}N_3Cl$	9.5	0.46	255.068	$C_{15}H_{12}N_2Cl$	11	-2.01	Impurity D (BP)*
UK 7	309.092	$C_{16}H_{19}N_2Cl_2$	7.5	-0.29	264.034	$C_{14}H_{12}NCl_2$	8.5	-0.61	$[M+H+Cl]^+$
					238.019	$C_{12}H_{10}NCl_2$	7.5	-0.09	
UK 8	297.079	$C_{17}H_{14}ON_2Cl$	11.5	-0.027	279.068	$C_{17}H_{12}N_2Cl$	12.5	-0.14	

*BP= British pharmacopoeia

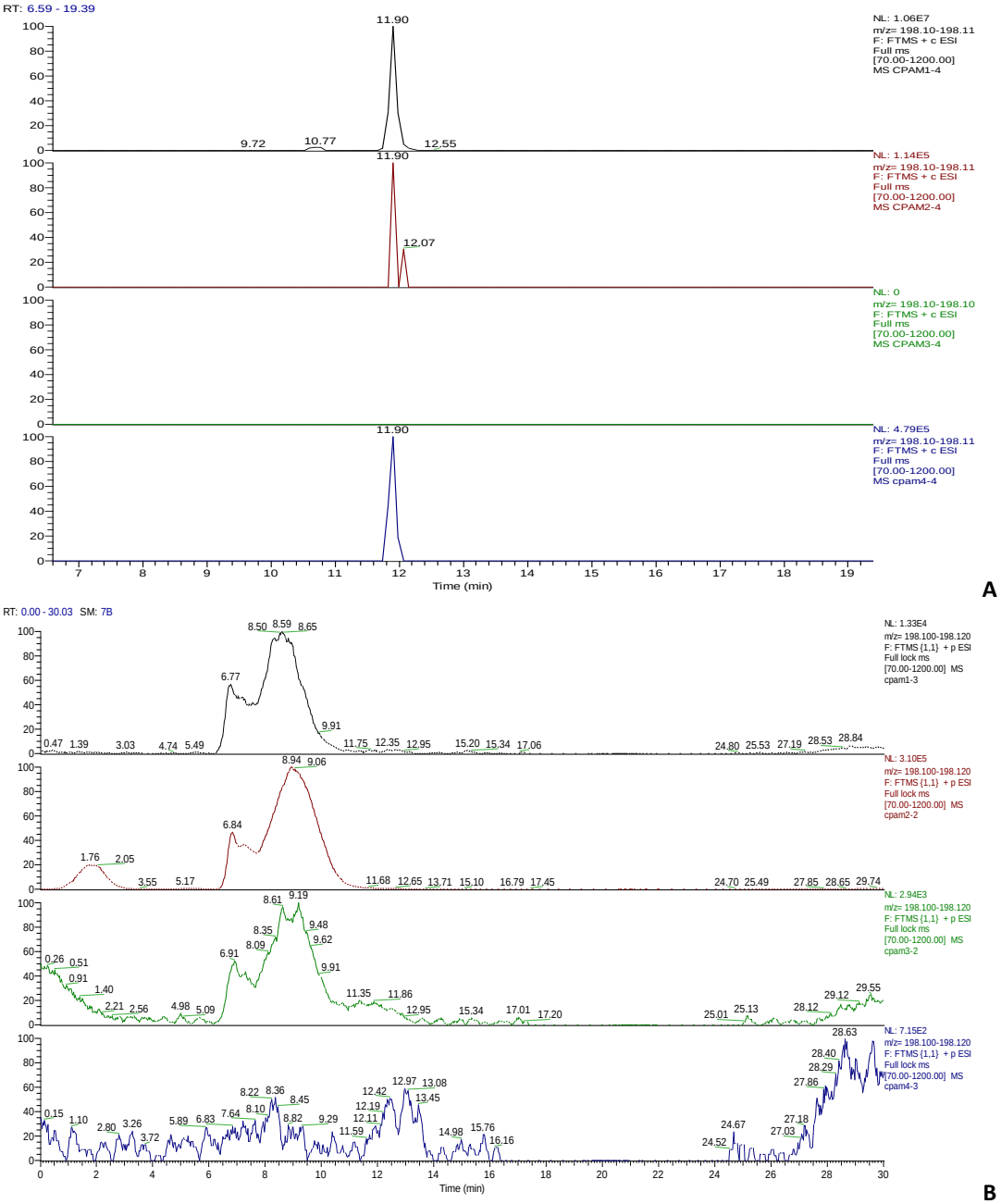


Figure 4.10 EIC trace for unknown impurity 1 in chlorpheniramine from (A) RP-LC (B) HILIC analysis.

An unknown impurity with the elemental composition $C_{11}H_{17}NCl$ can be observed in three samples. This impurity gives a poor peak shape but is well resolved from chlorpheniramine and thus is not an associated fragment.

The LC-MS^2 and accurate mass data reveal the structure of m/z 198.1039 whose elemental composition is $C_{11}H_{17}NCl$. Collision induced dissociation (CID) of this precursor ion generated daughter ions at m/z 153.046 and 125.015. The product ion

at m/z 153.046 results from the neutral loss of (C_2H_7N) from the precursor ion related to fragment ion coming from alpha-cleavage. The ion at m/z 125.015 generates from the loss of dimethylaminomethyl group ($C_4H_{11}N$) from an aliphatic chain. Figure 4.10 shows an extracted ion chromatogram for unknown impurity 1 in chlorpheniramine. This leads to a proposed structure for the impurity with m/z 198.1039 as being due to an impurity where the pyridine ring is absent from chlorpheniramine as shown in Figure 4.11. Formation of this impurity would be possible via any of the three synthetic routes shown in figures 4.2 - 4.4.

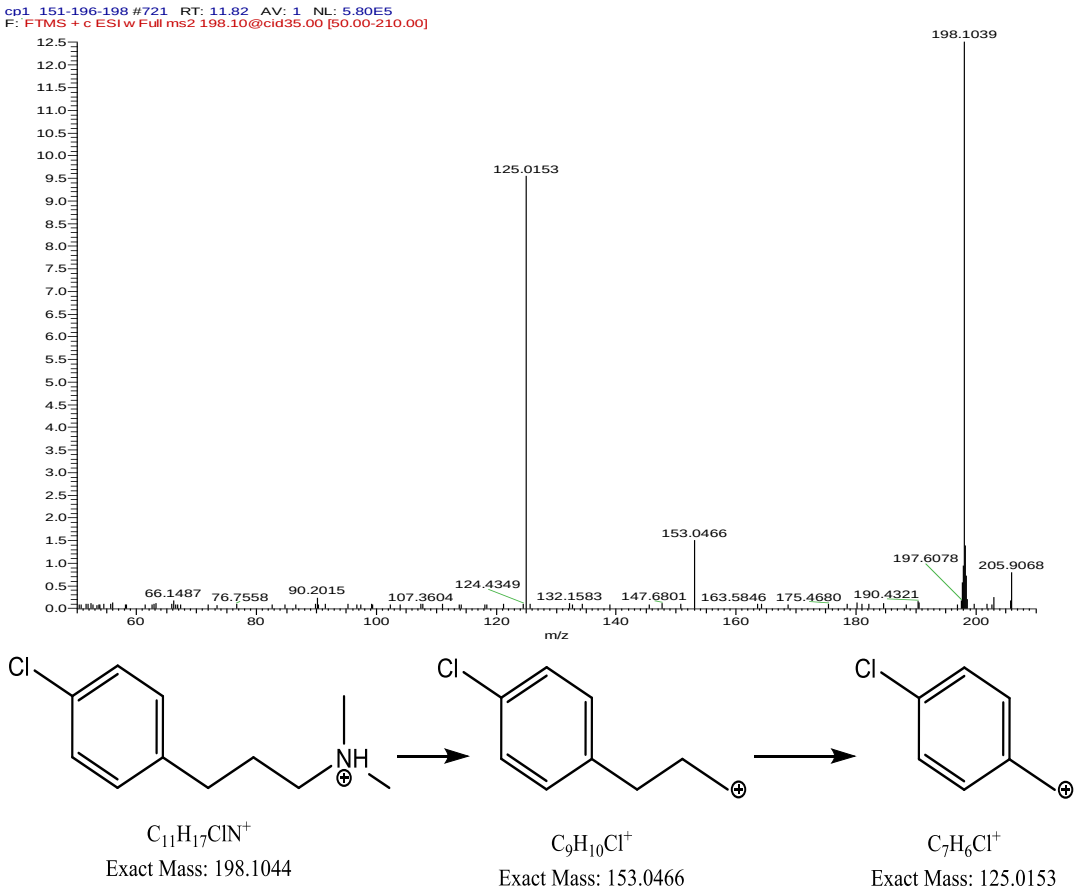


Figure 4.11 MS^2 spectra and proposed structure of impurity at ($[M+H]^+$, m/z 198.1044) and its fragments.

The MS data confirmed that an impurity with m/z 223.0994 had an elemental composition of $C_{12}H_{16}N_2Cl$. The CID from the impurity with m/z 223.0994 produced

fragments of the molecular ion at m/z 207.068 and 178.042. These fragments are related to the loss of the methyl and dimethylamine groups, respectively. Figure 4.12 shows the extracted ion chromatography of unknown impurity 2 in chlorpheniramine. The structure of unknown impurity 2 is consistent with the absence of a pyridine ring in impurity D (Figure 4.13). This impurity could only arise via the synthetic route (1) shown in figure 4.2 and suggests that CPAM3 is also synthesised via this route.

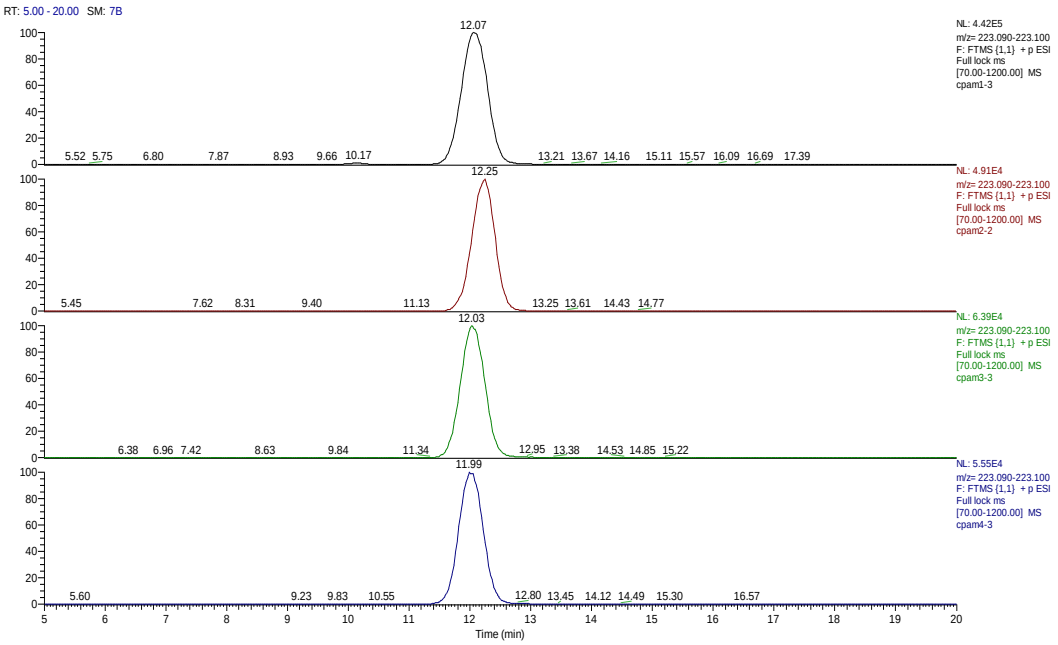


Figure 4.12 EIC for unknown impurity 2 in chlorpheniramine.

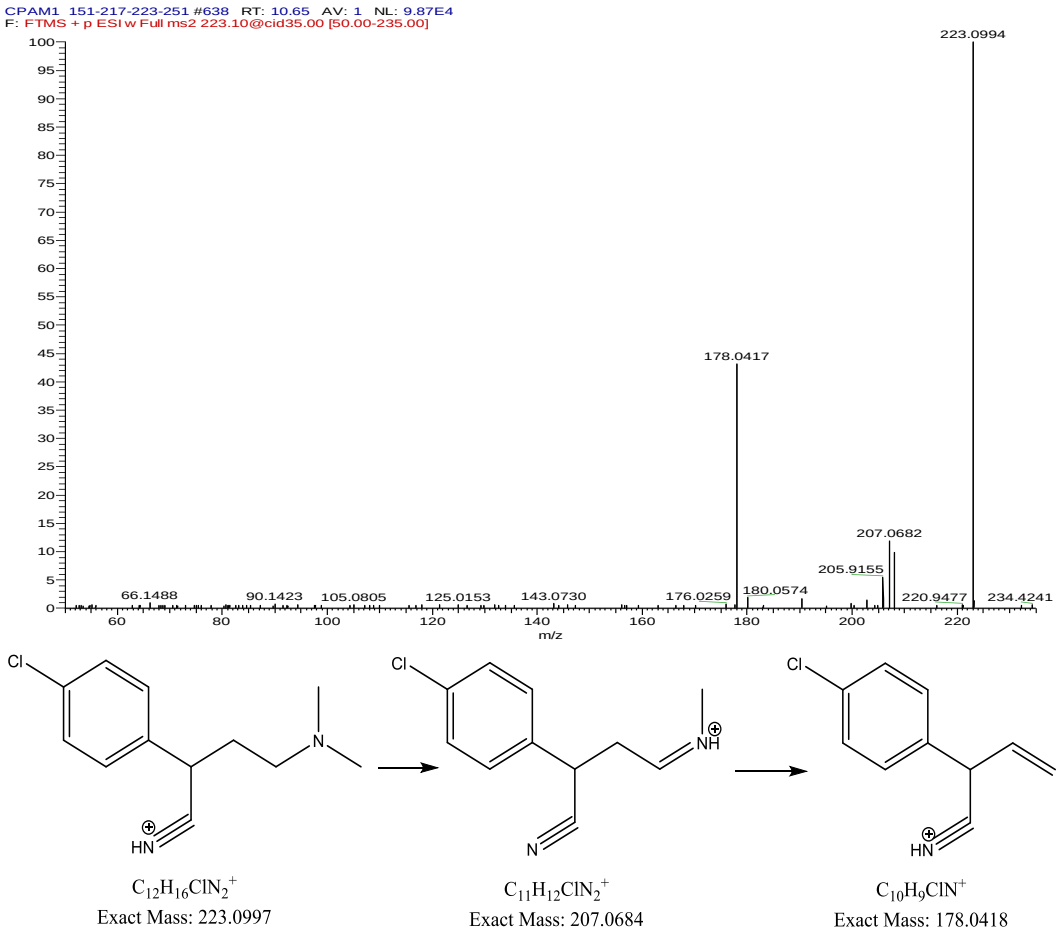


Figure 4.13 MS^2 spectra and proposed structure for unknown impurity 2 at ([M+H] $^+$, m/z 223.0994) and its fragments.

Figure 4.14 shows extracted ion traces for unknown impurity 3. The compound elutes very late in the run and has been partly missed by the HILIC chromatographic method. The HRMS data suggest that unknown impurity at m/z 241.1699 with predicted molecular formula $C_{16}H_{21}N_2$ and with a mass error of 0.4 ppm is due to the absence of a chlorine atom in chlorpromazine. The MS^2 of the ion at m/z 241.1699 was generated to give ion at m/z 196.112 with elemental composition $C_{14}H_{14}N$, resulting from the loss of a dimethylamine group due to α-cleavage (Figure 4.15). The impurity of m/z 241.1699 is also known as the antihistamine drug which is called pheniramine. This impurity could arise via any of the synthetic routes due to the non-chlorinated impurity, benzene, in the chlorobenzene precursor.

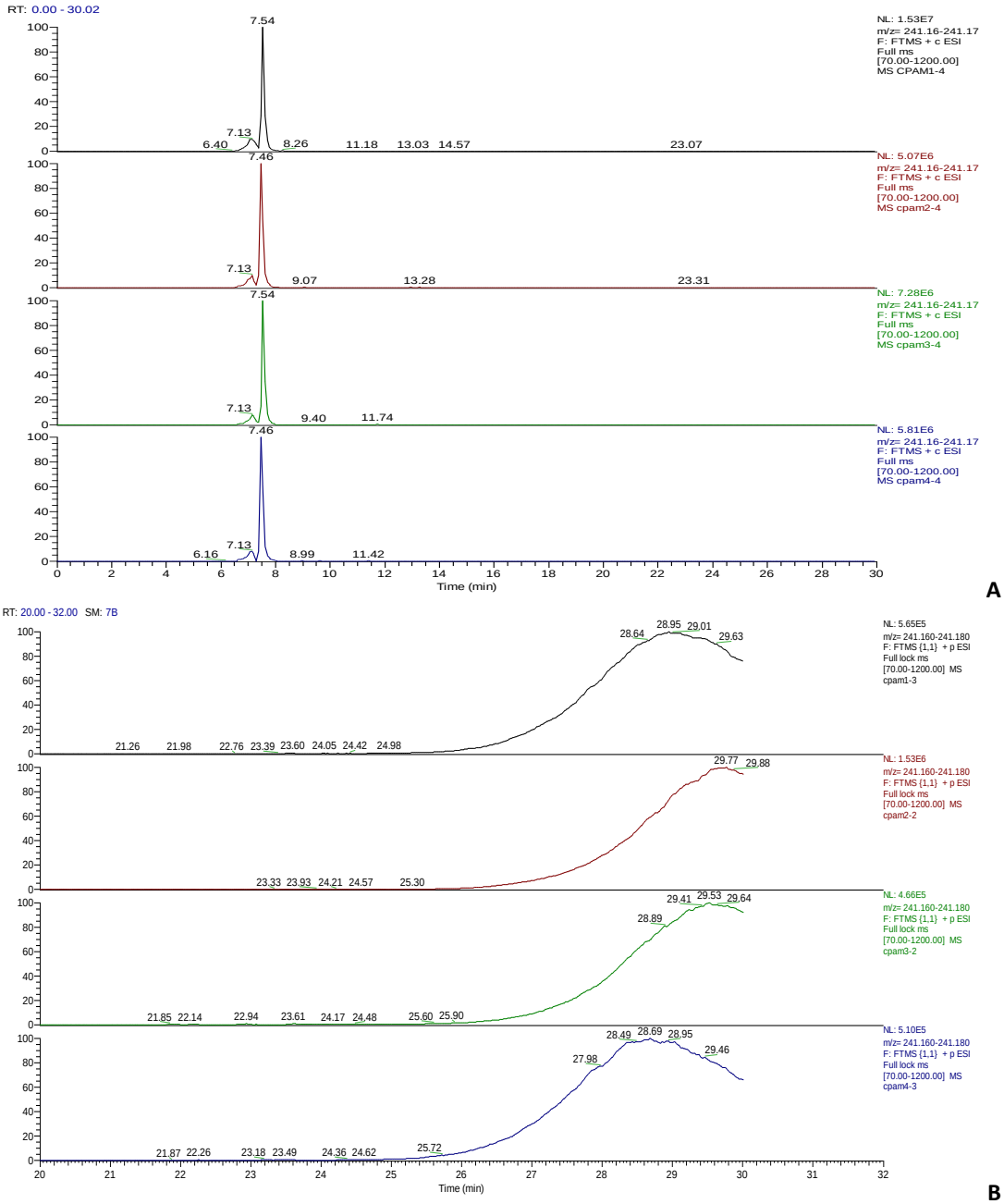


Figure 4.14 Extracted ion trace for unknown impurity 3 in chlorpheniramine from (A) RP-LC (B) HILIC analysis.

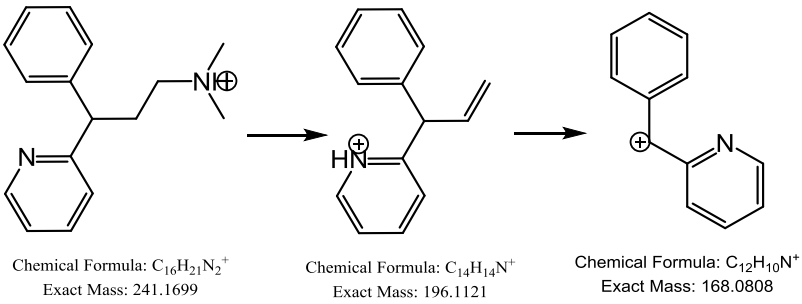


Figure 4.15 MS^2 fragmentation pathway for unknown impurity 3 at ($[M+H]^+$, m/z 241.1699).

The extracted ion trace for unknown impurity 4 is shown in figure 4.16. This impurity has the elemental composition $C_{17}H_{14}N_2Cl$. Its structure is consistent with that shown in figure 4.17. This impurity could arise via the synthetic route (1) shown in figure 4.2 if two chloropyridine moieties react with the cyano group. Its MS^2 spectrum shows the loss of the chlorobenzyl group.

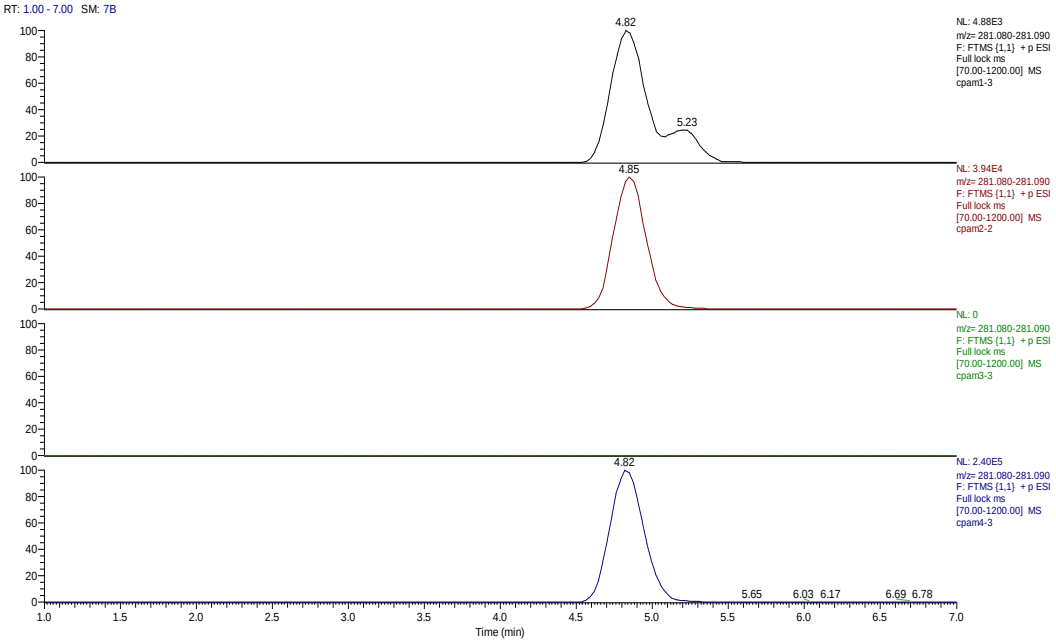


Figure 4.16 EICs for unknown impurity 4 in chlorpheniramine from HILIC.

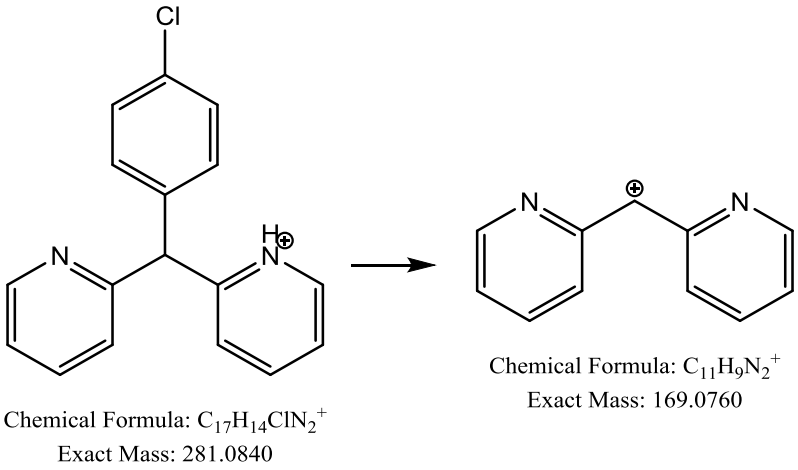


Figure 4.17 Proposed structure and fragmentation of unknown impurity 4 at ([M+H] $^+$, m/z 281.0840).

The elemental composition for unknown impurity 5 is $C_{17}H_{22}N_2Cl$. The EIC trace for this impurity is shown in figure 4.18. The formula consistent with the addition of an extra methyl group to chlorpheniramine. It is not immediately clear how this might arise although it is possible that there might be an impurity in the dimethylamine ethylene chloride used in the synthesis.

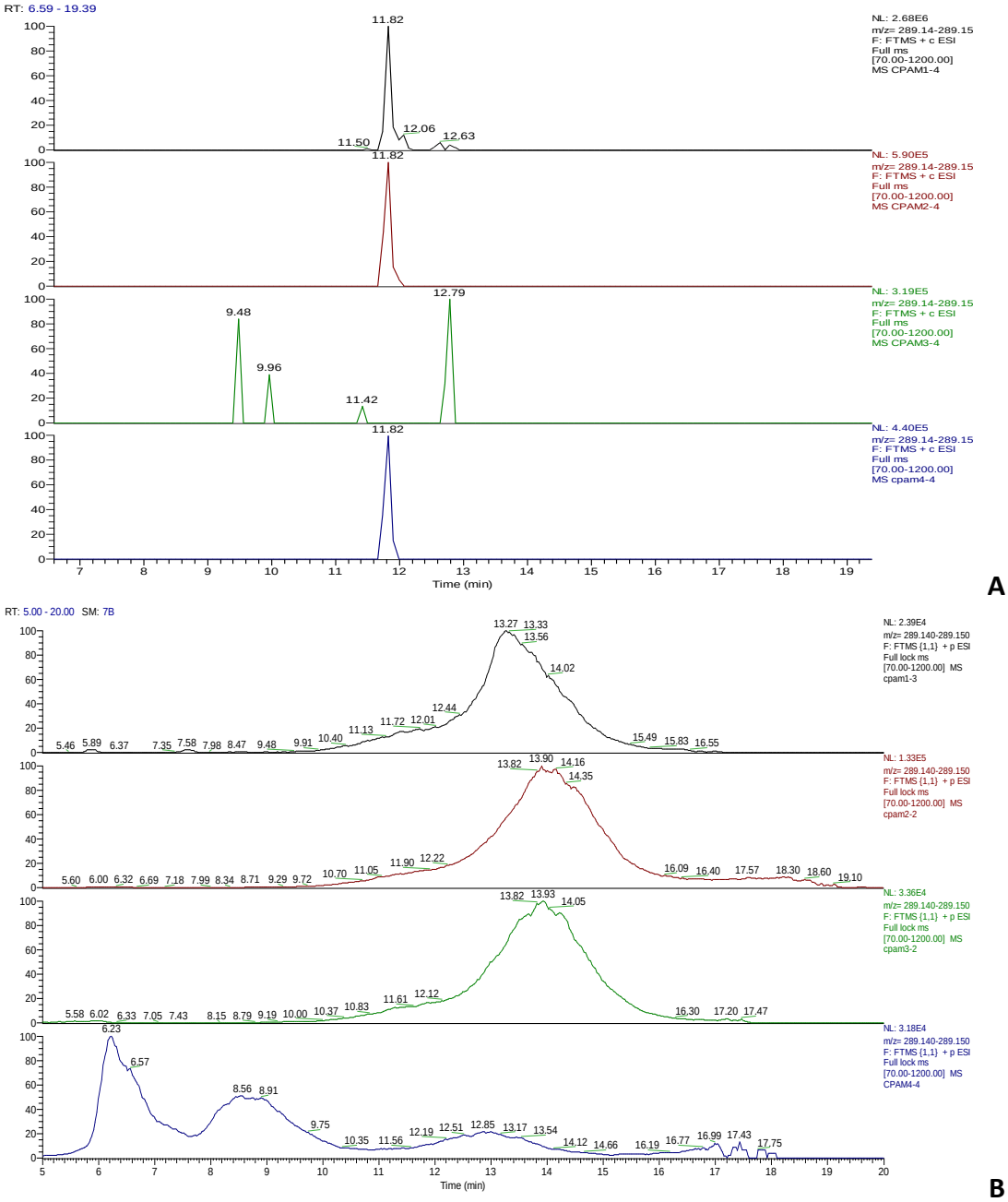


Figure 4.18 EIC trace for unknown impurity 5 in chlorpheniramine from (A) RP-LC (B) HILIC analysis.

Figure 4.19 shows the EIC trace for unknown impurity 6. This impurity has the formula $C_{16}H_{20}N_2OCl$. This is consistent with the addition of an oxygen to chlorpheniramine. This could occur via formation of an N-oxide. The MS^2 spectrum (Figure 4.20) indicates the impurity loses dimethylamine which means that the oxygen must be located on the pyridine nitrogen. This modification could occur via any of the synthetic routes.

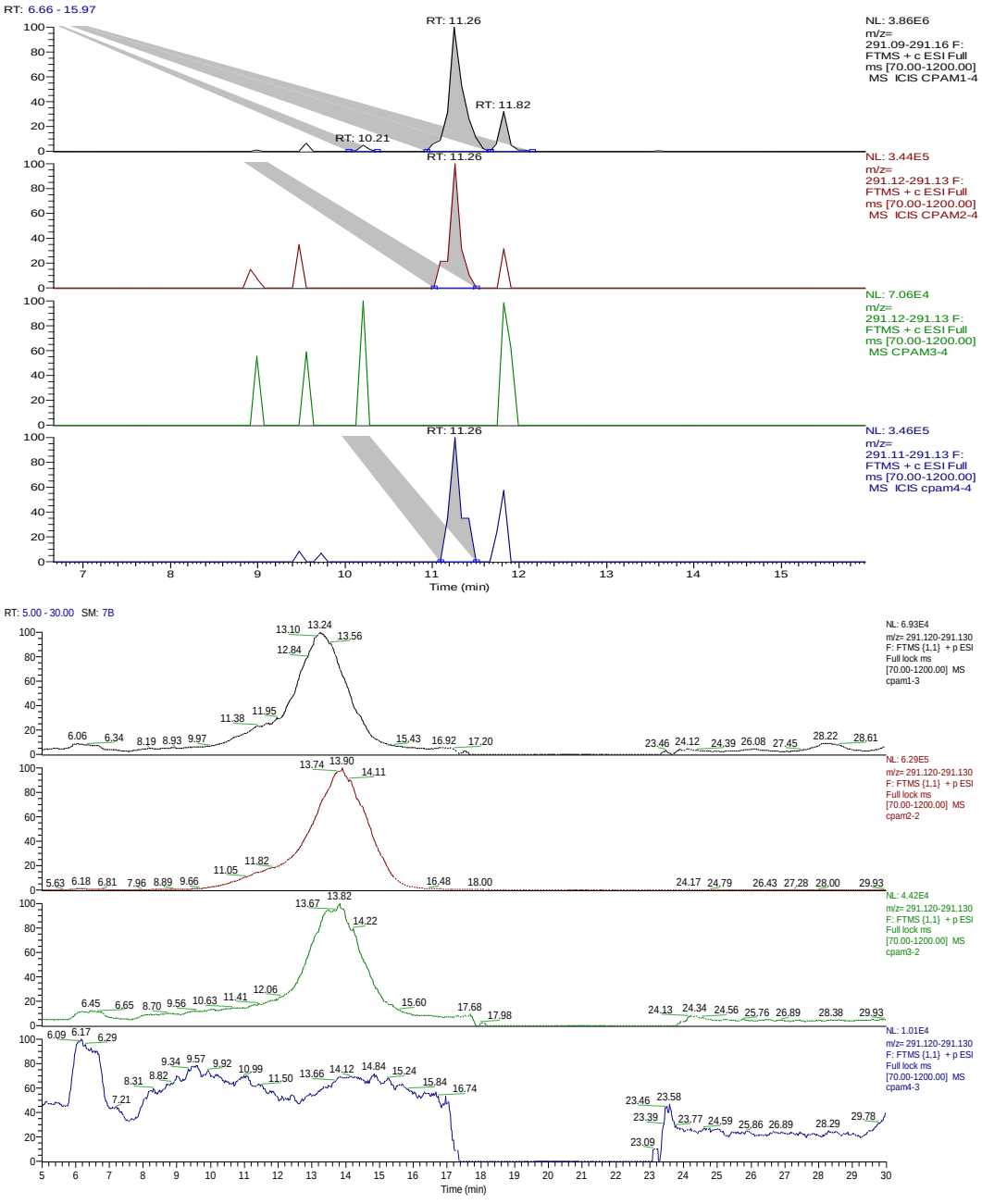


Figure 4.19 EIC trace for unknown impurity 6 in chlorpheniramine from (A) RP-LC (B) HILIC analysis.

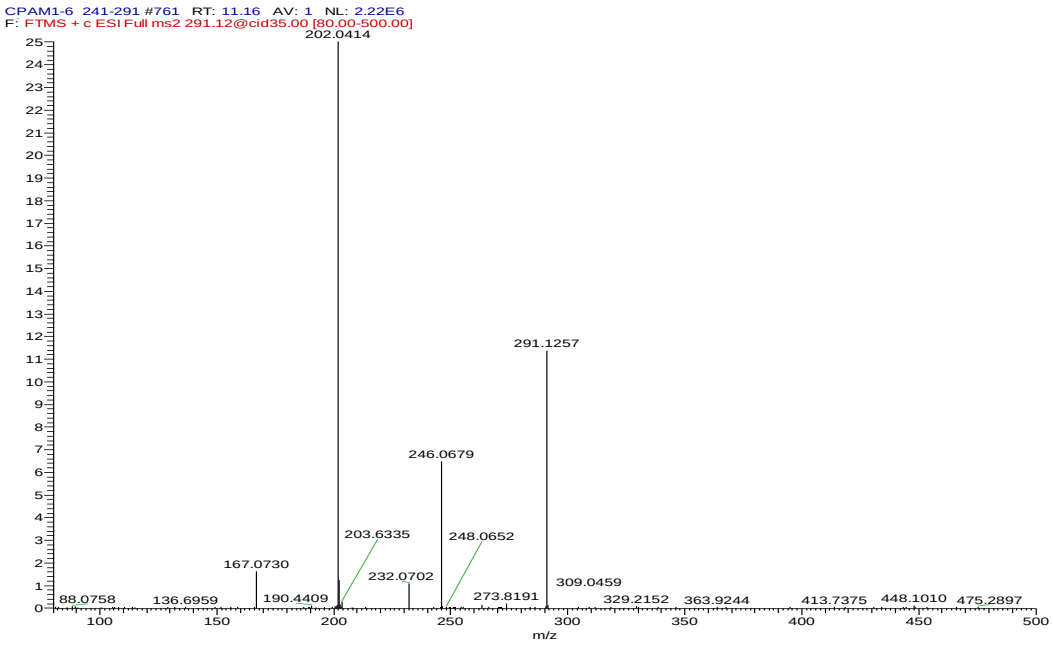


Figure 4.20 MS^2 spectra of unknown impurity 6 at ($[M+H]^+$, m/z 291.1257).

Figure 4.21 shows the EIC trace for unknown impurity 7. The MS^2 data revealed that the impurity had m/z 309.0920 with a mass error of 0.29 ppm has an elemental composition of $C_{16}H_{19}N_2Cl_2$ thus had an extra chlorine atom. The CID produced fragments of the molecular ion at m/z 264.034, which correlated to $C_{14}H_{12}NCl_2$. The fragment peak is due to the loss of dimethylamine due to alpha cleavage (Figure 4.22). This impurity could arise via the synthetic route (1) shown in figure 4.2 if the pyridine moiety reacted at a position not carrying a chlorine atom.

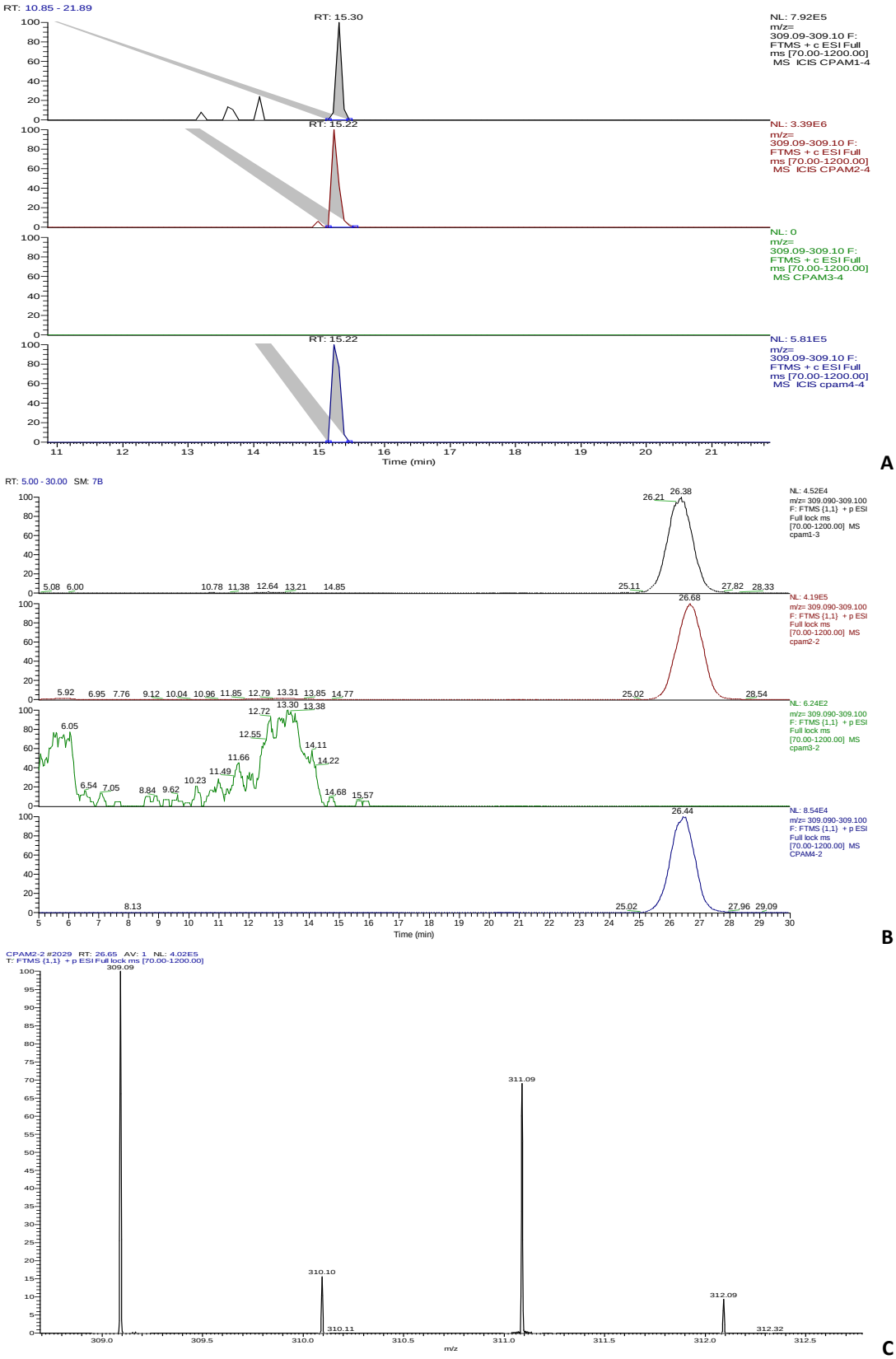
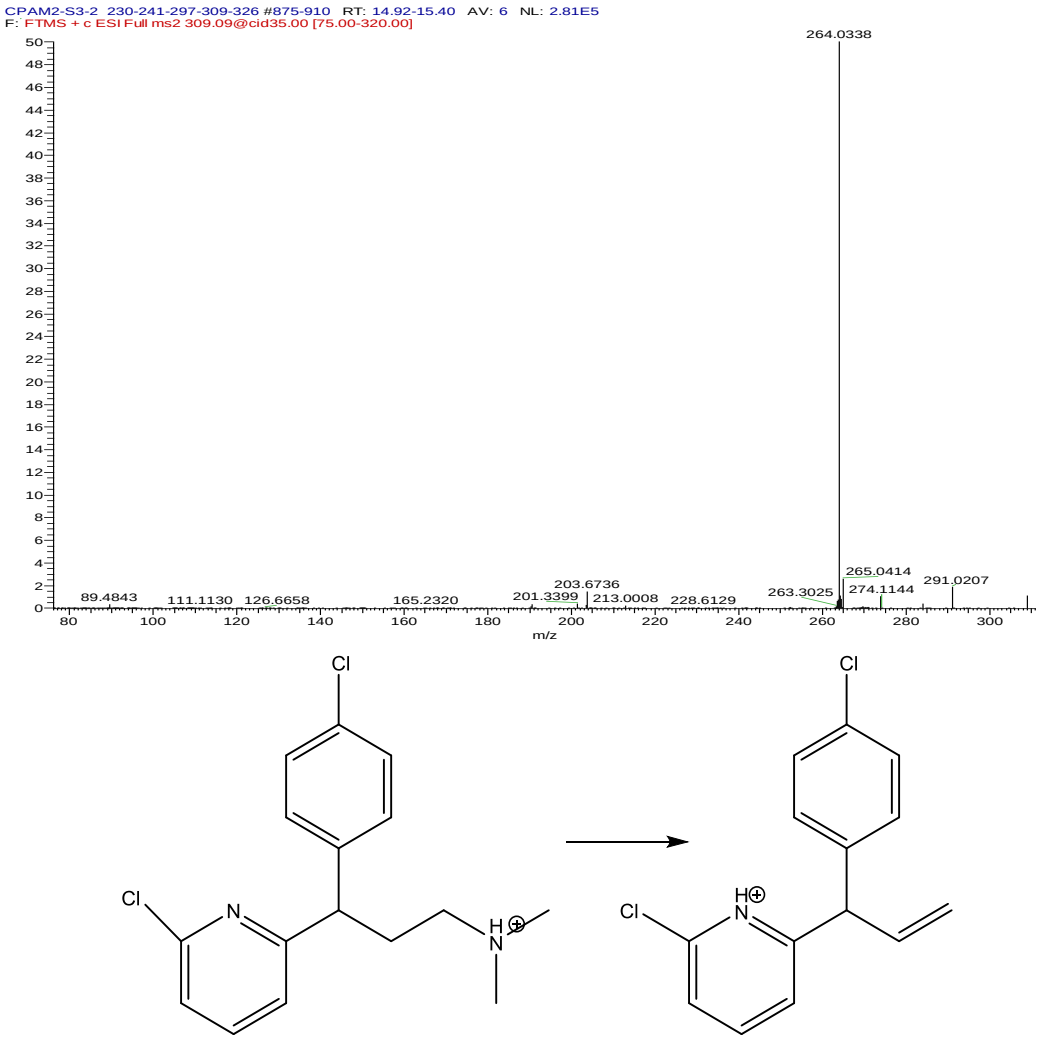


Figure 4.21 EIC for unknown impurity 7 in chlorpheniramine from (A) RP-LC, (B) HILIC analysis (C) MS spectra.



Chemical Formula: $C_{16}H_{19}Cl_2N_2^+$
Exact Mass: 309.0920

Chemical Formula: $C_{14}H_{12}Cl_2N^+$
Exact Mass: 264.0341

Figure 4.22 MS^2 spectra and proposed structure of unknown impurity 7 at ([M+H]$^+$, m/z 309.0920) and its fragment.

Figure 4.23 shows the EIC trace for unknown impurity 8 which has an elemental composition $C_{17}H_{14}N_2OCl$ which is consistent with the attachment of an additional oxygen atom to unknown impurity 4. This could be via hydroxylation of via the formation of an N-oxide.

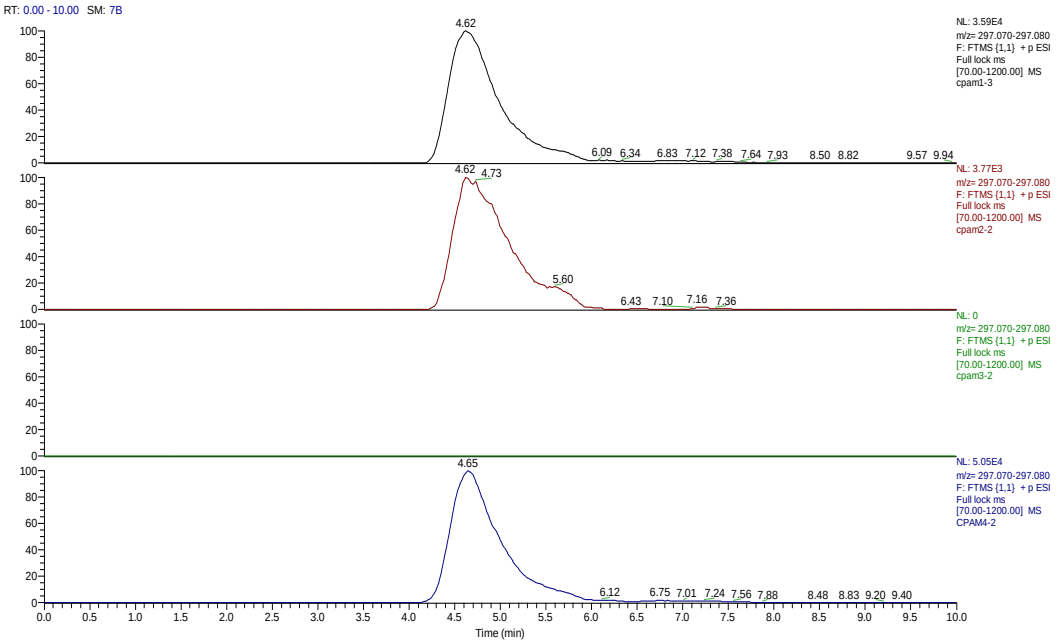


Figure 4.23 EIC trace for unknown impurity 8 in chlorpheniramine from HILIC MS.

The apparent impurity at m/z 230 was the major impurity peak in all four samples. The concentration of this impurity is around 14% and 16% of total API in the HILIC and RP-LC analysis, respectively. The peak is co-elutes at the same retention time with chlorpheniramine peak itself. This impurity is related to loss of C_2H_6N corresponding to a molecular weight of 45 Da. Thus it seems likely that this is not an impurity but rather a fragment of chlorpheniramine produced in the mass spectrometer.

4.3.2 Chemometric modelling of impurity profiles

MZmine 2.20 software was used to process the data from LC-HRMS. Many steps were used to detect the peaks that included peak detection, mass detection, chromatogram deconvolution, deisotoping, normalisation, alignment and gap filling. Different features from HILIC-MS and RPLC-MS were extracted and the data were then transferred to Simca P for chemometric modelling. PCA was applied as an unsupervised method to discriminate the different for the samples. The supervised OPLS-DA model was generated related to the HCA dendrogram of the PCA model.

4.3.2.1 Data analysis of HILIC models

4.3.2.1.1 Unsupervised PCA analysis

A total of 38 different features were extracted and the data were then transferred to Simca P for chemometric modelling. Principal components analysis (PCA) was applied as an unsupervised method to discriminate the different for the samples. The discrimination depends on the differences in the nature and level of the impurities present in each sample. The first principal component (PC1) described 51.8% of the variation in all data and PC2 (23.6%) and PC3 (12.6%) with total PCs (88.0%). The PCA showed high level of model fits to the data with good reproducibility.

A scores plot is used to display the pattern of observations (samples) due to the differences and similarities expressed by their variables (impurities). The scores plot obtained by PCA shows distinct clusters of the four samples each run in four technical replicates and loading plot indicates the scattering of the impurities related to their presence in the samples (Figure 4.24).

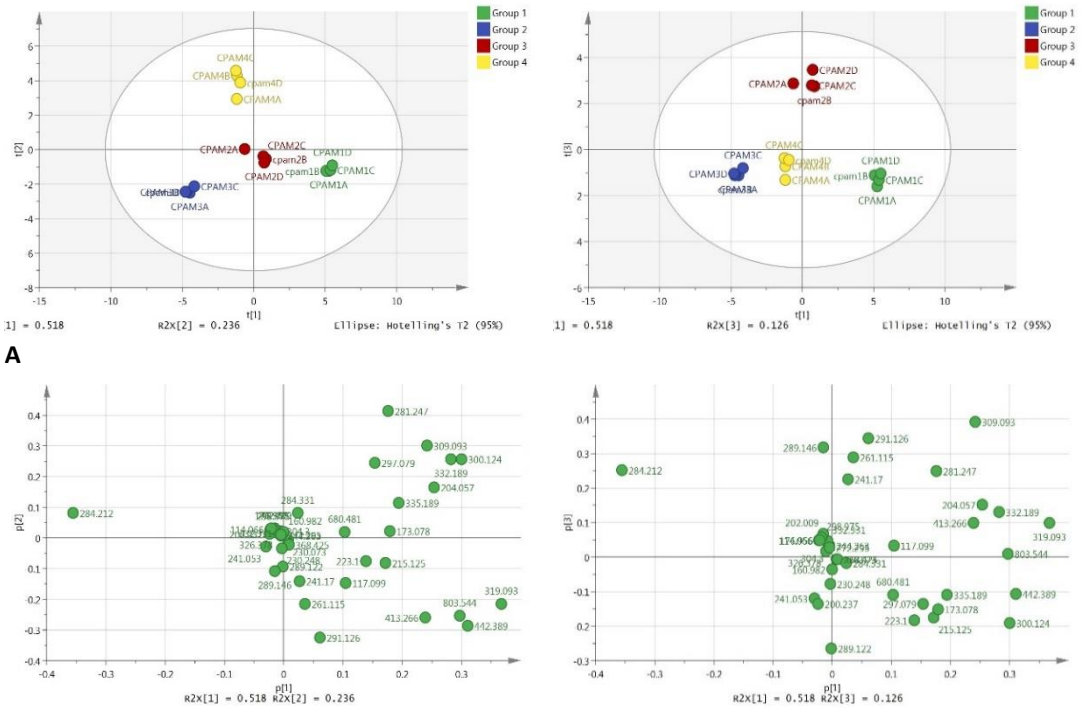


Figure 4.24 (A) PCA scores scatter plot (t1 vs. t2) (left) and (t1 vs. t3) (right) (B) Loading plot (p1 vs. p2) (left) and (p1 vs. p3) (right) for chlorpheniramine samples.

PC1 and PC2 obtained by using Pareto scaling. Three PCs described 88% of the whole variation data (PC1: 51.8%, PC2: 23.6% and PC3: 12.6%). Scores plot shows the discrimination between the samples according to the impurity profile of each sample. Loading plot points out the impurities that govern the position of each sample in the scatter plot.

Hierarchical clustering analysis (HCA) aims to explore which potential impurities would be more efficient in classifying chlorpheniramine samples according to their impurity profile. The HCA shows that the samples can be clustered into four groups (Figure 4.25). The classification pattern has been led by the inter-individual variability which is the outcome of variable intensity happening in the different individuals at a specific sample. The plot classified the samples into four distinct hierarchical groups with a variability index of 160 approximately. The variability index (vertical axis) is expressed as the measure of the degree of closeness between the individual data points or groups. The low value (i.e. small number) subdivisions reveal on the most similar samples. Longer divisions (or high value) show an increase dissimilarity among the groups. Thus CPAM2 and CPAM4 the most similar in terms of impurity profile.

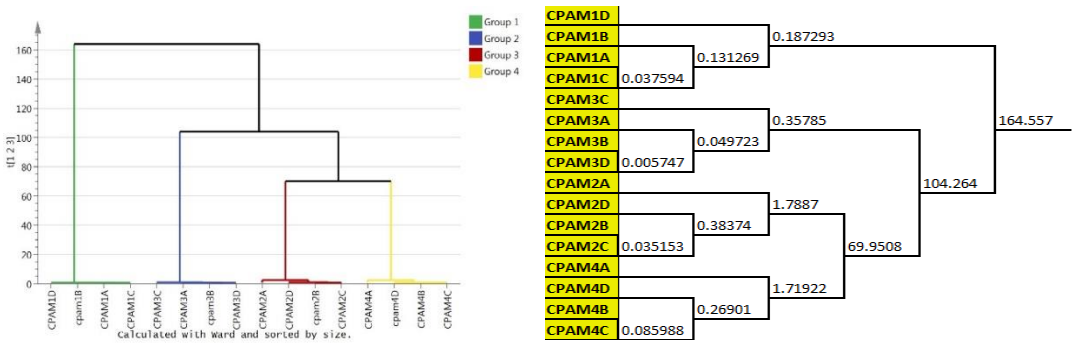
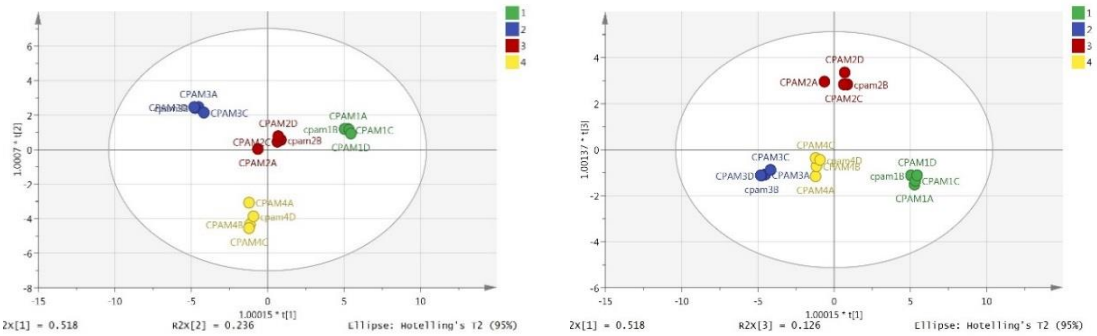


Figure 4.25 Hierarchical Clustering Analysis (HCA) plot for the four samples of chlorpheniramine from HILIC analysis.

The dendrogram shows observations clustered into four hierarchical groups; (CPAM1, green), (CPAM2, blue), (CPAM3, red) and (CPAM4, yellow) by using Ward clustering method. X-axis represents the samples and y-axis shows the variability index (right). The values of the variability index among the samples. The larger values show big differences between group variability and the lower values show more similarity (left).

4.3.2.1.2 Supervised OPLS-DA analysis

The PCA and HCA plots provide information about the similarity of the samples. The separation between groups based on a degree of intervention can be perceived by characterizing the multivariate model using supervised OPLS-DA analysis. The cross validation of the model shows a high value of R^2(cum) = 0.985 (i.e. $R^2 > 0.5$) which means the OPLS-DA model has a high degree of fit and therefore most of the variance related to the response variable (Y) can be explained. In addition, the high value of Q^2(cum) = 0.967 means that the model has a high degree of consistency between the original and predicted data (Figure 4.26). For investigating the impurities which bring about the difference among groups, biplots were applied to illustrate the major contributing variables. It shows which of the variables are more highly correlated to a given group of observations than to the others (Figure 4.27). This can also be confirmed with S-plot for each group.



A B

Figure 4.26 OPLS-DA scores scatter plot (t1 vs. t2): (A) and (t1 vs. t3): (B) from HILIC analysis.
CPAM1 (Green), CPAM2 (Blue), CPAM3 (Red), and CPAM4 (Yellow) (PC1 versus PC2) obtained by using Pareto scale. Three PCs described 92.8% of the whole variation data (PC1: 79.5%, PC2: 7.43% and PC3: 5.83%). Scores plot shows the discrimination between the samples according to the impurity profile of each sample.

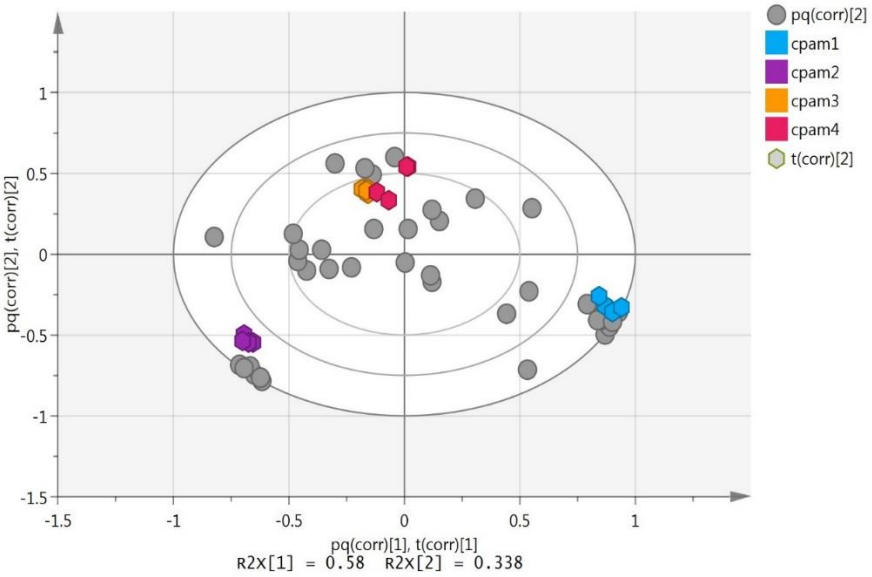


Figure 4.27 Biplot: scores (samples) and loadings (impurity) with low discrimination power on pq(corr)[1] t(corr)[1] vs. pq(corr)[2] t(corr)[2] from HILIC analysis.
The plot shows distribution of 16 observations from four samples and 38 variables. The model includes four groups: CPAM1 (blue), CPAM2 (burble), CPAM3 (orange), and CPAM 4 (red). It consists of one predictive x-score component; component t(corr)[1] and one orthogonal x-score components t(corr)[2] and x-variables component; component pq(corr)[1] and one orthogonal x-variables components pq(corr)[2].

OPLS-DA is used to predict the significant variables (i.e. impurities) having a direct effect on each observation (i.e. sample). Variable importance in projection (VIP) was proposed as a variable selection method. It is a well-established way which was developed specifically for a modelling approach. VIP value larger than 0.5 indicates that the variable is important to the model and the value of less than 0.5 is considered

an unimportant variable. Figure 4.28 contains the significant variables (*red*) was obtained with a VIP more than 0.5 and where jack-knife standard error does not cross zero.

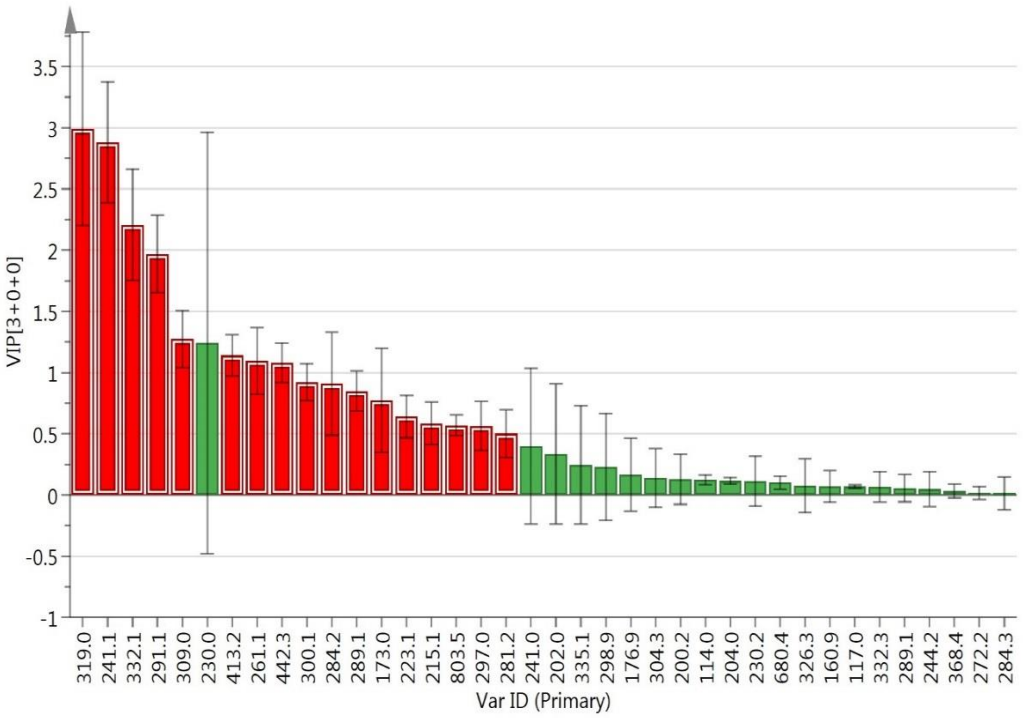
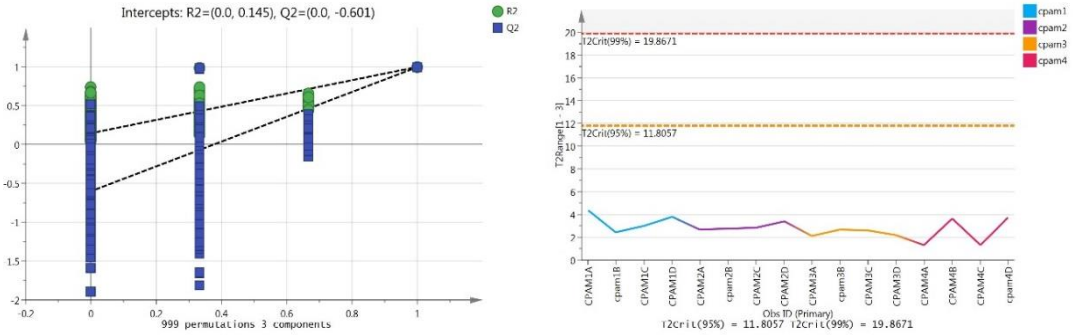


Figure 4.28 VIP plot of the variables of chlorpheniramine samples with the jack-knife standard error at (95%CI) of the VIP computed form all rounds of cross validation from HILIC analysis.
A variable has VIP > 0.5 can be considered as important impurities of chlorpheniramine samples (red column).

The supervised classification model was validated by using a permutation test as a cross-validation method to optimise the model fittings and to assess the performance parameters. As shown in the permutation test (Figure 4.29A), the y-axis gives the value of R^2 = (0.0, 0.145) and Q^2 = (0.0, -0.601) which is related to the goodness of fit and prediction, respectively. The x-axis characterises the correlation between the permuted y-vectors (on the left) and the original y-vector (on the right). The model could be considered statistically significant since the regression line of the Q^2 points intersects the y-axis below zero. In addition, the hotelling's T2 line plot displays the distance from the origin in the model plane (score space) for each selected observation. Values larger than the 95% confidence limit are suspect, and values

larger than the 99% confidence limit can be considered as not significant (Figure 4.29B).



A **B**

Figure 4.29 (A) Validation plot obtained from the permutation test (B) Hotelling's T2 line plot displays the distance from the origin in the model plane (score space) for each selected observation.

Intercepts: R^2= (0.0, 0.145), Q^2= (0.0, -0.601) and the permutation number (n = 999). T2Crit (95%) confidence limit is the warning limit and represented by yellow dotted line, and T2Crit (99%) confidence limit is the action limit and represented by red dotted line. Observations above the action limit considered strong outliers. T2Range is basically calculated as the sum over the selected range of components of the scores in square divided by their standard deviations in square.

The Receiver Operator Characteristic (ROC) curve shows the robustness of the OPLS-DA model obtained from the cross-validated predicted Y (predicted class). It shows that (AUC = 1.0) which means the prediction model of OPLS-DA has great sensitivity and specificity (i.e. AUC > 0.5 corresponds to a good predictive model) (Figure 4.30).

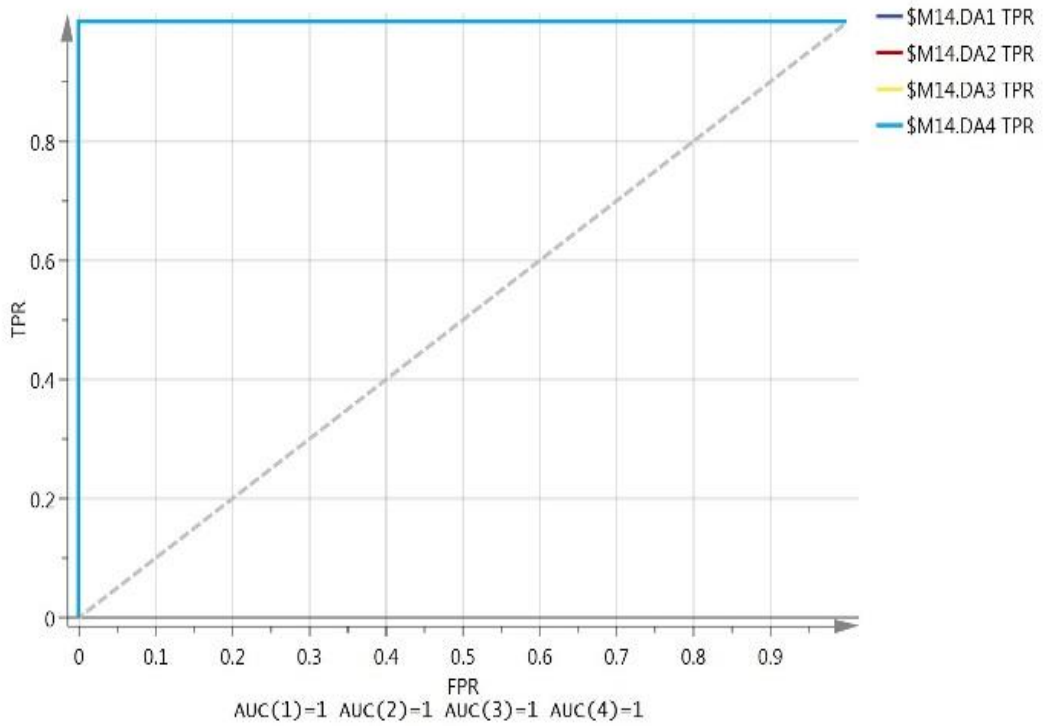
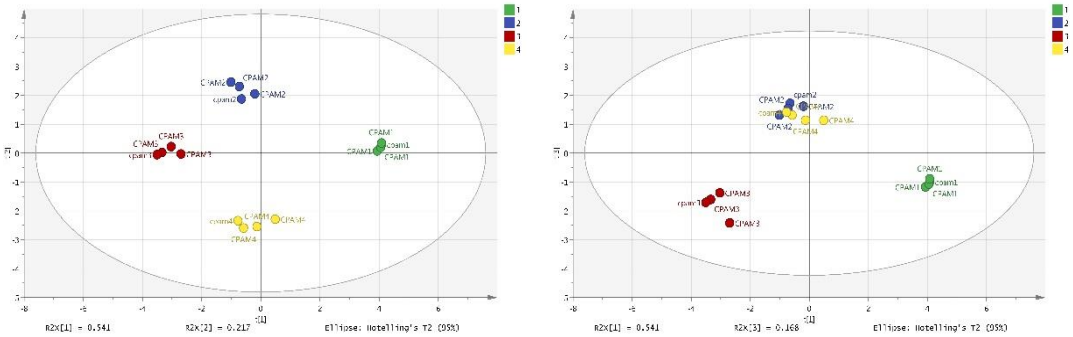


Figure 4.30 Area Under the Receiver Operating Characteristics Curve (ROC) curve calculated from the cross-validated predicted Y-values of the OPLS-DA model from HILIC analysis.
The ROC curve shows sensitivity (true positive rate (TPR)) on the y-axis versus (false positive rate (FPR = 1 - Specificity)) on the x-axis, the value of both normalised to 1 which represents the value of AUROCC for group. All four groups of samples with excellent AUROCC accuracy (i.e. = 1).

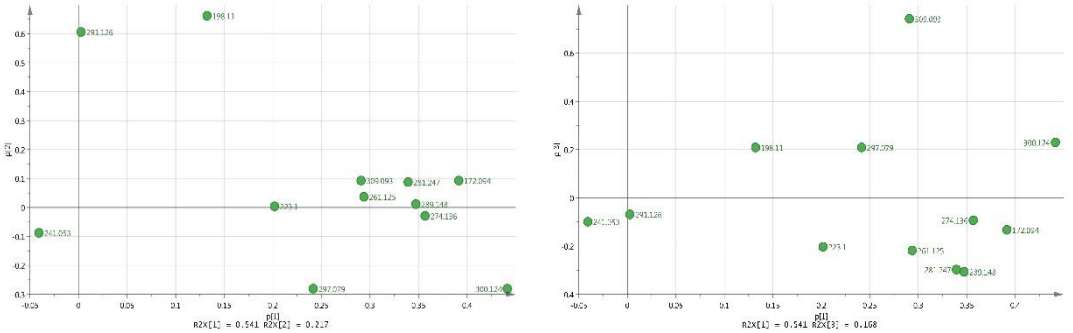
4.3.2.1.3 Selected important variables from HILIC model

For discrimination the synthesis and degradation pathways, new models were applied on the specific impurities shown in Table 4.2 for the HILIC data. PCA and OPLS-DA models were created to compare the samples. The PCA first principal components (PC1) described 54.1% of the variation in all data set and PC2 (21.7%) and PC3 (16.8%) with total PCs (92.6%). The PCA showed high level of model fits to the data with good reproducibility.

The scores plot obtained by PCA shows distinct clusters of the four samples and loadings plot indicates the scattering of the impurities related to their presence in the samples (Figure 4.31).



A



B

Figure 4.31 (A) PCA scores scatter plot (t1 vs. t2) (left) and (t1 vs. t3) (right) (B) Loading plot (p1 vs. p2) (left) and (p1 vs. p3) (right) for chlorpheniramine samples from HILIC analysis.

PC1 and PC2 obtained by using Pareto scale. Three PCs described 92.6% of the whole variation data (PC1: 54.1%, PC2: 21.7% and PC3: 16.8%). Scores plot shows the discrimination between the samples according to the impurity profile of each sample. Loading plot points out the impurities that govern the position of each sample in the scatter plot.

The HCA shows that the samples clustered into four groups (Figure 4.32). The plot classified the samples into four distinct hierarchical groups with a variability index of 90 approximately. The classification using the selected impurities shows no difference in the distribution the samples (Figure 4.32) and CPAM2 and CPAM4 were the most similar in terms of impurity profile analysis according to the HILIC separation.

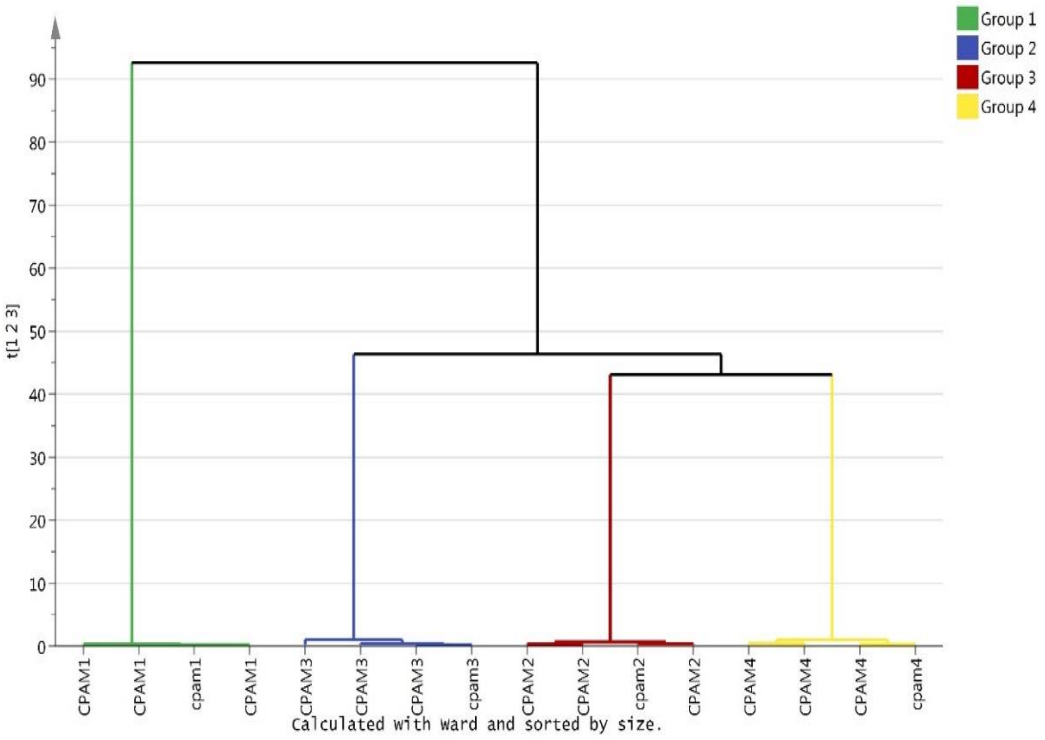
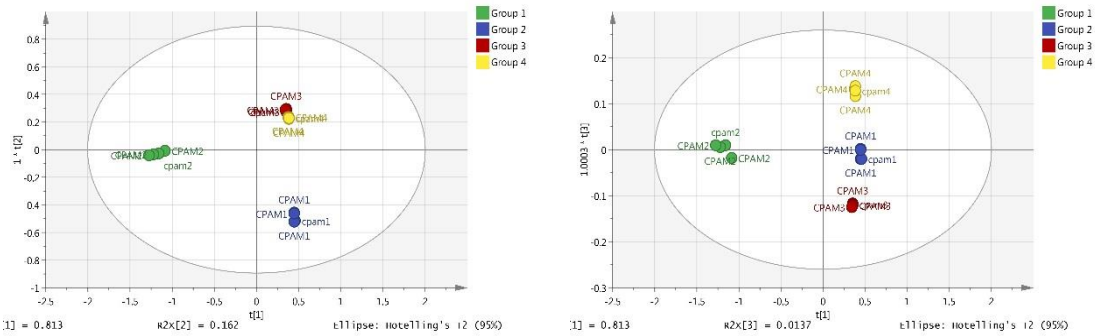


Figure 4.32 Hierarchical Clustering Analysis (HCA) plot for the four samples of chlorpheniramine from HILIC analysis.
The dendrogram shows observations clustered into four hierarchical groups; (CPAM1, green), (CPAM2, red), (CPAM3, blue) and (CPAM4, yellow) by using Ward clustering method. X-axis represents the samples and y-axis shows the variability index. The values of the variability index among the samples. The larger values show big differences between group variability and the lower values show more similarity.

OPLS-DA models were applied based on the PCA and HCA plots which provided information differences between the samples. The cross validation of the model shows a high value of R^2(cum) = 0.985 (i.e. $R^2 > 0.5$) which means the OPLS-DA model has a high degree of fit and therefore most of the variance related to the response variable (Y) can be explained. In addition, the high value of Q^2(cum) = 0.967 means that the model has a high degree of consistency between the original and predicted data (Figure 4.33). For investigating the impurities which bring about the difference among groups, a biplot was applied to illustrate the major contributing variables. It shows which of the variables are more highly correlated to the given group of observations than to the others (Figure 4.34). This also might confirm with S-plot of each group.



A B

Figure 4.33 OPLS-DA scores scatter plot (t1 vs. t2): (A) and (t1 vs. t3): (B) from HILIC analysis.
CPAM1 (Blue), CPAM2 (Green), CPAM3 (Red), and CPAM4 (Yellow) (PC1 versus PC2) obtained by using Pareto scale. Three PCs described 92.8% of the whole variation data (PC1: 79.5%, PC2: 7.43% and PC3: 5.83%). Scores plot shows the discrimination between the samples according to the impurity profile of each sample.

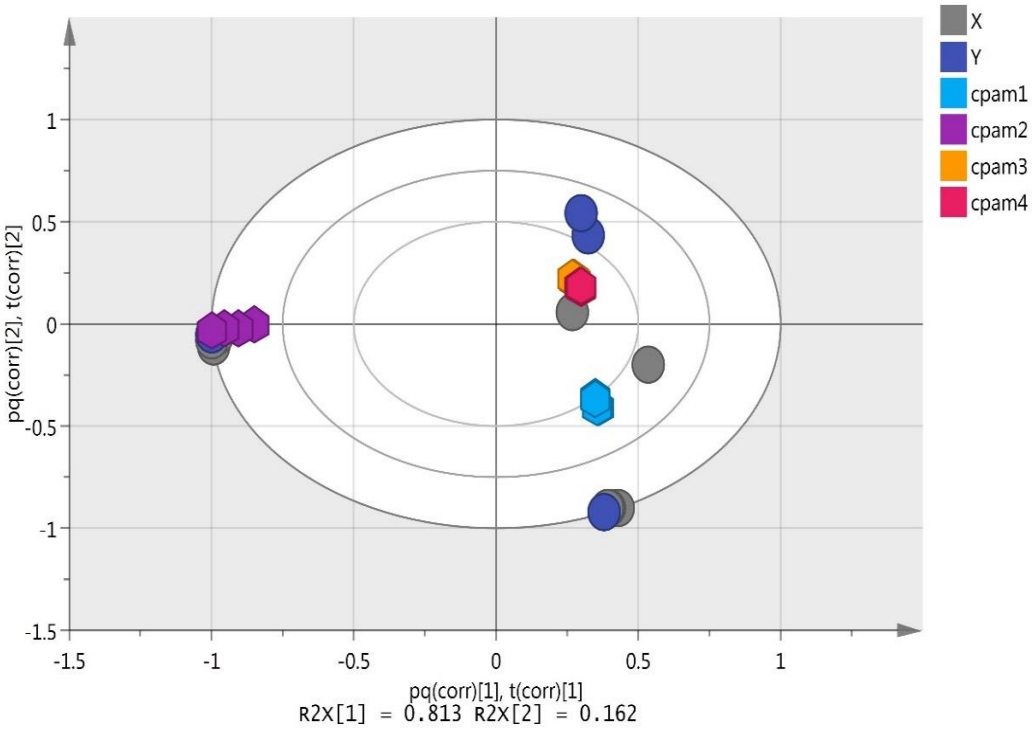


Figure 4.34 Biplot: scores (samples) and loadings (impurity) with low discrimination power on pq(corr)[1] t(corr)[1] vs. pq(corr)[2] t(corr)[2] from HILIC analysis.
The plot shows distribution of 16 observations from four samples and 12 variables. The model includes four groups: CPAM1 (blue), CPAM2 (purple), CPAM3 (orange), and CPAM4 (red). It consists of one predictive x-score component; component t(corr)[1] and one orthogonal x-score components t(corr)[2] and x-variables component; component pq(corr)[1] and one orthogonal x-variables components pq(corr)[2].

OPLS-DA is used to predict the significant variables (i.e. impurities) have a direct effect on each observation (i.e. sample). Variable importance in projection (VIP) was

proposed was as a variable selection method. A VIP value larger than 0.5 indicates that the variable is important to the model and the value of less than 0.5 is considered an unimportant variable. Figure 4.35 contains the significant variables (*red*) was obtained from the VIP more than 0.5 and jack-knife standard error not cross zero.

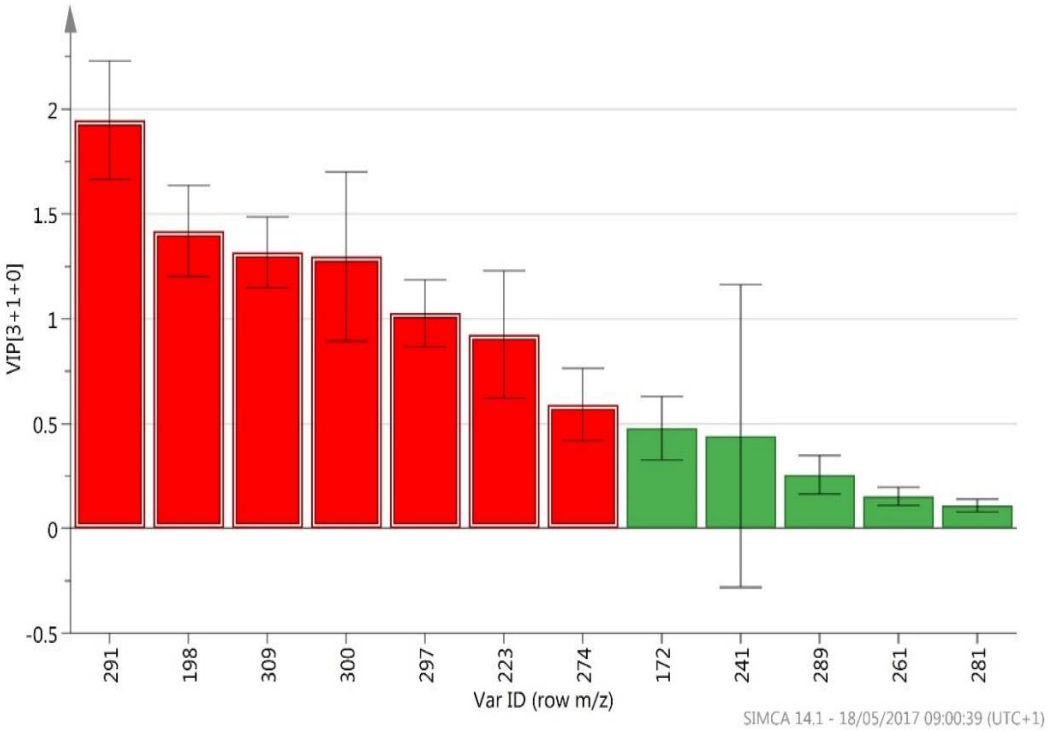
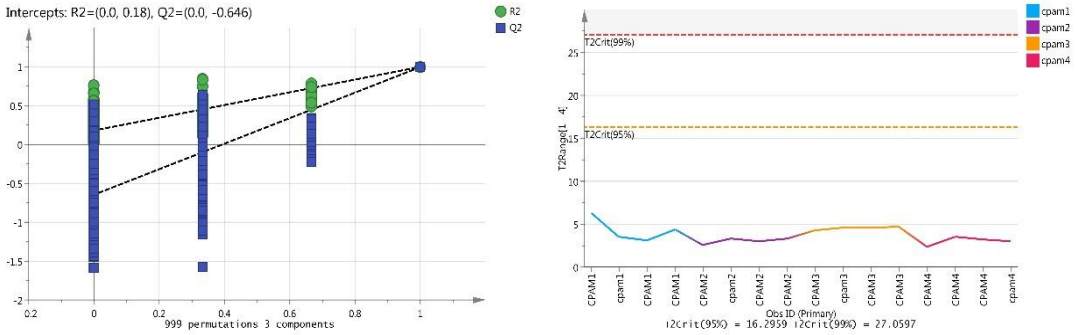


Figure 4.35 VIP plot of the variables of chlorpheniramine samples with the jack-knife standard error at (95%CI) of the VIP computed form all rounds of cross validation from HILIC analysis.
A variable has VIP > 0.5 can be considered as important impurities of chlorpheniramine samples (red column).

The supervised classification model was validated by using permutation test as a cross-validation method to optimise the model fittings and to assess the performance parameters. As shown in the permutation test (Figure 4.36A), the y-axis gives the value of R^2 = (0.0, 0.145) and Q^2 = (0.0, -0.601) which is related to the goodness of fit and prediction, respectively. The x-axis characterises the correlation between the permuted y-vectors (on the left) and the original y-vector (on the right). The model could be considered statistically significant since the regression line of the Q^2 points intersects the y-axis below zero. In addition, the hotelling's T2 line plot displays the distance from the origin in the model plane (score space) for each selected

observation. Values larger than the 95% confidence limit are suspect, and values larger than the 99% confidence limit can be considered as outliers (Figure 4.36B).



A **B**

Figure 4.36 (A) Validation plot obtained from the permutation test (B) Hotelling's T2 line plot displays the distance from the origin in the model plane (score space) for each selected observation from HILIC analysis.

Intercepts: R^2= (0.0, 0.18), Q^2= (0.0, -0.646) and the permutation number (n = 999). T2Crit (95%) confidence limit is the warning limit and represented by yellow dotted line, and T2Crit (99%) confidence limit is the action limit and represented by red dotted line. Observations above the action limit considered strong outliers. T2Range is basically calculated as the sum over the selected range of components of the scores in square divided by their standard deviations in square.

The Receiver Operator Characteristic (ROC) curve shows the robustness of the OPLS-DA model obtained from the cross-validated predicted Y (predicted class). It shows that (AUC = 1.0) which means the prediction model of OPLS-DA has great sensitivity and specificity in classifying the samples according to their impurity profile (i.e. AUC > 0.5 corresponds to a good predictive model) (Figure 4.37).

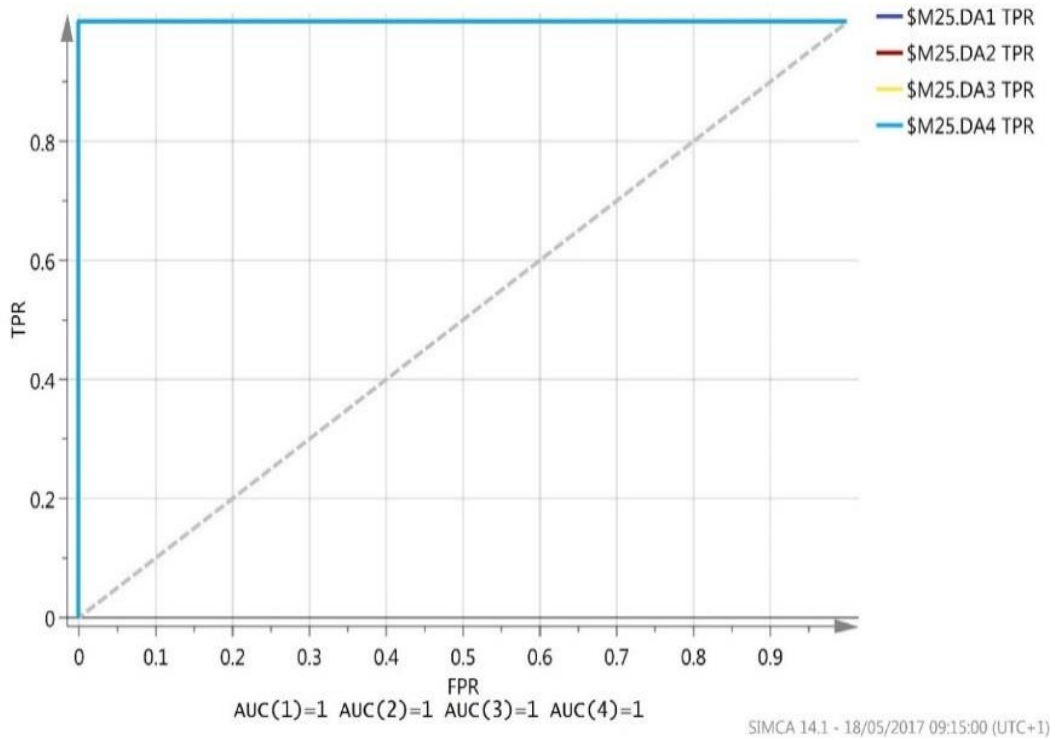


Figure 4.37 Area Under the Receiver Operating Characteristics Curve (ROC) curve calculated from the cross-validated predicted Y-values of the OPLS-DA model from HILIC analysis.
The ROC curve shows sensitivity (true positive rate (TPR)) on the y-axis versus (false positive rate (FPR = 1 - Specificity)) on the x-axis, the value of both normalised to 1 which represents the value of AUROCC for group. All four groups of samples with excellent AUROCC accuracy (i.e. = 1).

4.3.2.1.4 Samples comparison based on VIP

The selective important variables from VIP were deeply investigated in new models to diagnose the variables that make a difference among the chlorpheniramine samples. The comparison was done based on the OPLS-DA hierarchical clustering variability index values (Figure 4.38) between the groups by comparing (CPAM1 vs group of CPAM2, 3 & 4), (CPAM2 vs group of CPAM3 & 4), and (CPAM3 vs CPAM4). The information on the differences in the impurity pattern between the groups based on detecting the significant variables may reveal differences in the manufacturing pathways for each sample. All new OPLS-DA models showed a highly significant cross validation and the permutation test has intercepts for Q^2 below zero which shows that the goodness of fit and prediction performance are good (Table 4.5). The high value of validation in this model indicates that the selective variables are chosen correctly and presents the most important variables in the model.

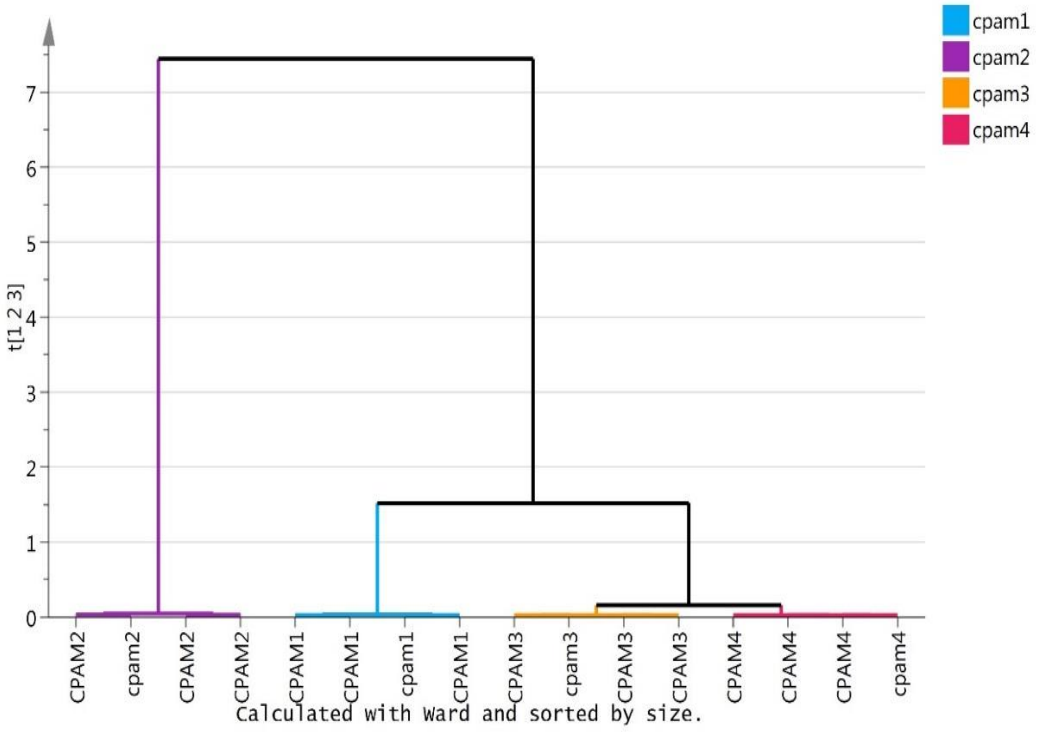


Figure 4.38 Hierarchical Clustering Analysis (HCA) plot for the four samples of chlorpheniramine related to OPLS-DA model from selected HILIC data.
The dendrogram shows observations clustered into four hierarchical groups; (CPAM1, blue), (CPAM2, purple), (CPAM3, orange) and (CPAM4, red) by using Ward clustering method. X-axis represents the samples and y-axis shows the variability index). The larger values show big differences between group variability and the lower values show more similarity.

Table 4.5 The cross validation and permutation test plot with intercepts at the permutation number (n = 999) for chlorpromazine models.

OPLS-DA Model	Cross validation		Permutation test	*P*-value
Component	R^2(cum)	Q^2(cum)	Intercepts	(< 0.05)
CPAM1 Model	0.995	0.992	R^2=(0.0, 0.048), Q^2=(0.0, -0.459)	2.7E-11
CPAM2 Model	0.992	0.986	R^2=(0.0, 0.085), Q^2=(0.0, -0.679)	3.53E-08
CPAM3 vs 4 Model	0.999	0.985	R^2=(0.0, 0.035), Q^2=(0.0, -0.934)	4.40E-05

To obtain on the variables that were important in each group an S-plot was used from OPLS-DA models comparing the groups. From the direct comparison, selective important variables in the S-plot represented the most relevant impurities which contribute in the variation among the groups. These potential impurities were investigated to confirm differences in their average intensity levels by considering the p-values (< 0.05) and 95% CIs.

The important variables were selected related their significance value in the in S-plot of CPAM1 (Figure 4.39), CPAM2 (Figure 4.40), and CPAM3 vs CPAM4 (Figure 4.41). Table 4.6 shows the significant important variables with their VIP value and with jack-knife standard error which makes the variation among the groups.

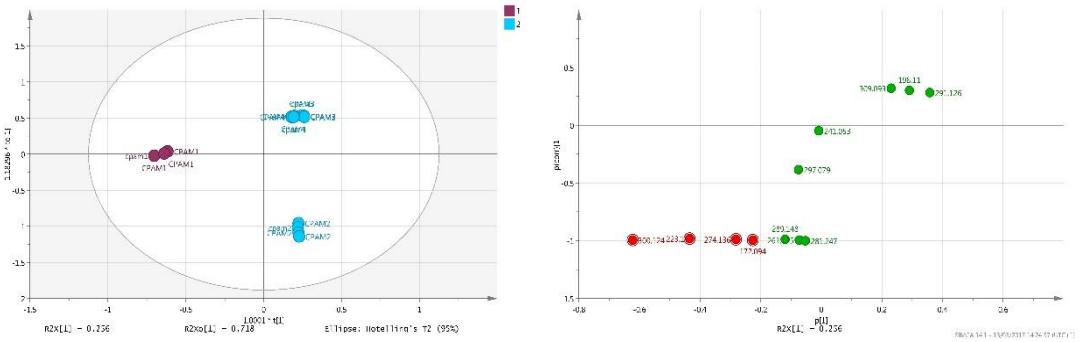


Figure 4.39 the scores plot (left) and S-plots at the first component of OPLS-DA (CPAM1 vs group 2, 3 & 4) from selected HILIC model.

S-plot shows impurities distributed according to their correlation (x-axis) and magnitude of effect (y-axis) of the CPAM1 sample. The red marked points represent selected significant impurities differentiating CPAM 1 from the other samples.

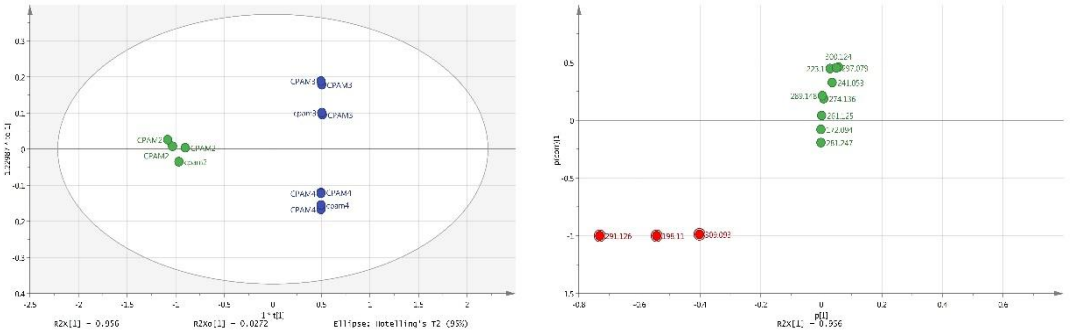


Figure 4.40 the scores plot (left) and S-plots at the first component of OPLS-DA (CPAM2 vs group 3 & 4) from selected HILIC model.

S-plot shows impurities distributed according to its correlation (x-axis) and magnitude of effect (y-axis) of the CPAM2 sample. The red marked points represent selected significant impurities differentiating CPAM 2 from the other samples.

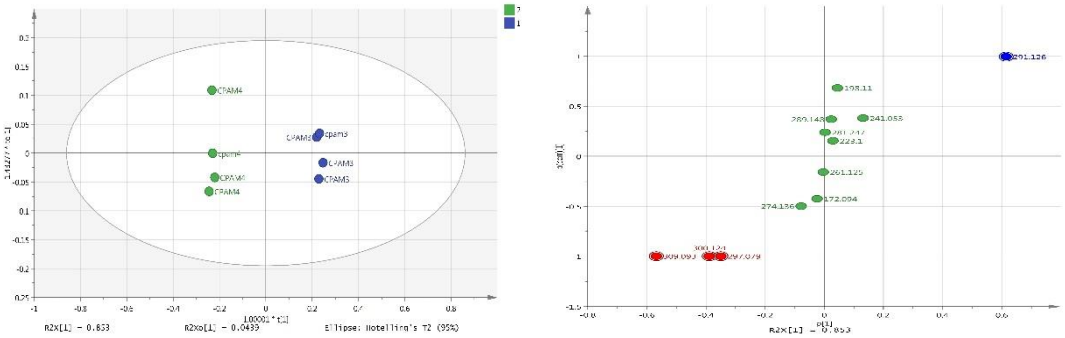


Figure 4.41 the scores plot (left) and S-plots at the first component of OPLS-DA (CPAM3 vs group 4) from selected HILIC model.

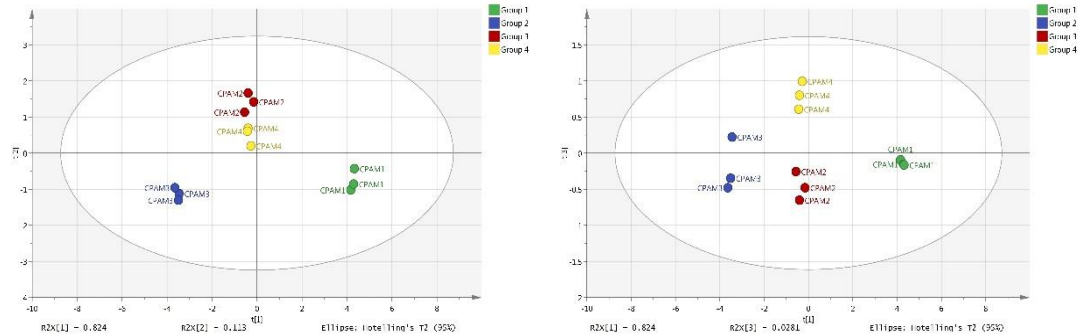
S-plot shows impurities distributed according to its correlation (x-axis) and magnitude of effect (y-axis) of the CPAM3 sample (blue) and the CPAM4 sample (red). The marked points represent selected significant impurities differentiating CPAM 3 and 4 from the other samples.

4.3.2.2 Data analysis of RP-LC models

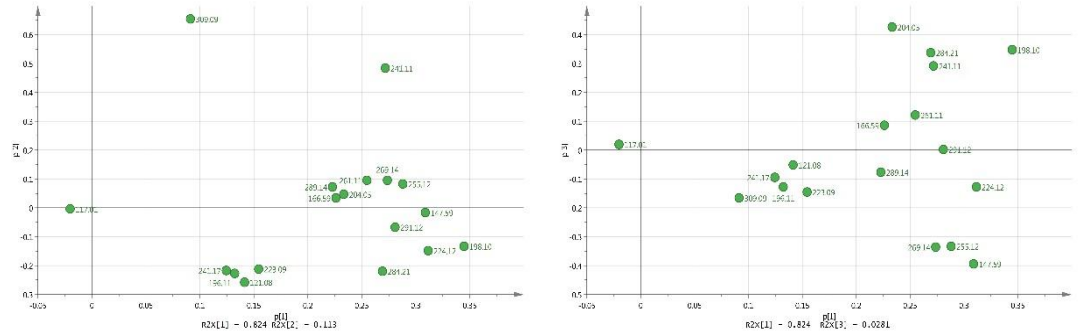
4.3.2.2.1 Unsupervised PCA analysis of RP-LC analysis

Principal components analysis (PCA) was applied as an unsupervised method to discriminate between the samples using the RP data. The discrimination is dependent on impurities that were formed in each sample during manufacture. The total principal components (PCs) were 98.1%. The first principal component (PC1) described 82.4% of the variation in all data set. PC2, PC3 and PC4 described (11.3%), (2.81%) and (1.57%), respectively. The PCA of this model showed a high level of fit to the data with good reproducibility.

A scores plot was used to display the pattern of observations (samples) due to the differences and similarities expressed by their variables (impurities). The scores plot obtained by PCA shows distinct clusters for the four samples and the loading plot indicates the scattering of the impurities related to their presence in the samples (Figure 4.42).



A



B

Figure 4.42 (A) PCA scores scatter plot (t1 vs. t2) (left) and (t1 vs. t3) (right) (B) Loading plot (p1 vs. p2) (left) and (p1 vs. p3) (right) for chlorpheniramine samples from RPLC-MS analysis.
PC1 and PC2 obtained by using Pareto scale. Three PCs described 98.1% of the whole variation data (PC1: 82.4%, PC2: 11.3%, PC3: 2.8% and PC4: 1.5%). Scores plot shows the discrimination between the samples according to the impurity profile of each sample. Loading plot points out the impurities that govern the position of each sample in the scatter plot.

The HCA shows that the samples can be clustered into four groups (Figure 4.43). The plot classified the samples into four distinct hierarchical groups with a variability index of 75 approximately. The variability index (vertical axis) is expressed as the measure of the degree of closeness between the individual data points or groups.

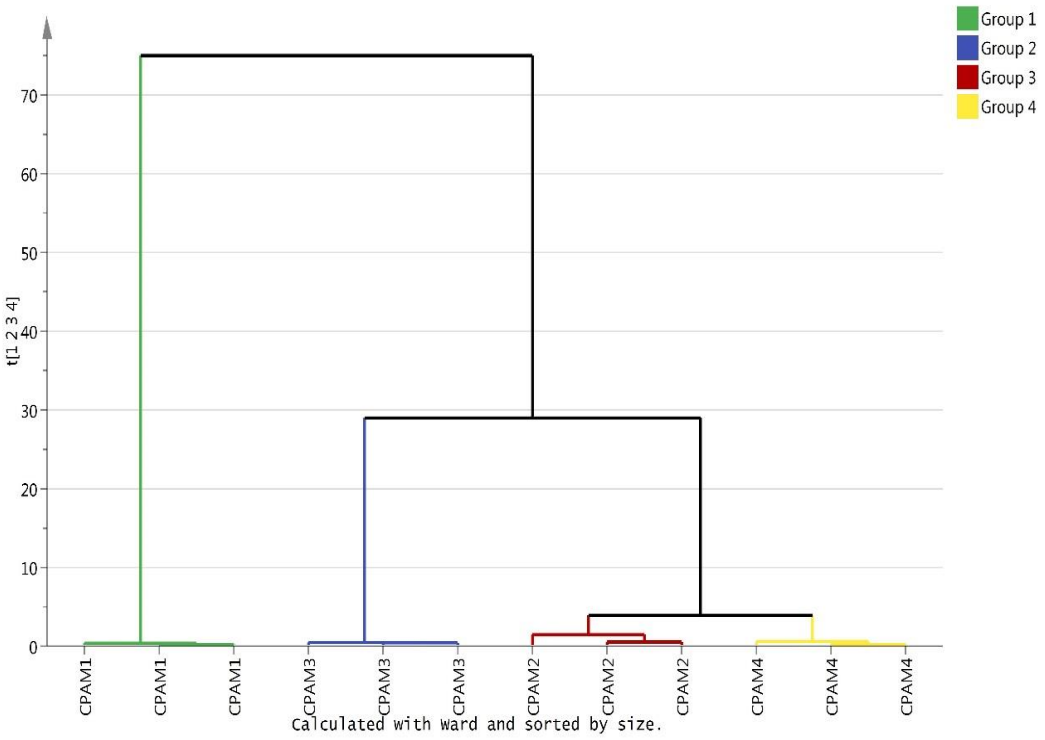


Figure 4.43 Hierarchical Clustering Analysis (HCA) plot for the four samples of chlorpheniramine from RPLC-MS analysis.
The dendrogram shows observations clustered into four hierarchical groups; (CPAM1, green), (CPAM2, red), (CPAM3, blue) and (CPAM4, yellow) by using the Ward clustering method. X-axis represents the samples and y-axis shows the variability index. The values of the variability index among the samples. The larger values show big differences between group variability and the lower value shows more similarity.

4.3.2.2.2 Supervised OPLS-DA analysis

The PCA and HCA plots provide information about any similarities among the samples based on unsupervised analysis. The separation between groups based on a degree of intervention can be by using supervised OPLS-DA analysis. The cross validation of the model in the current case showed a high value of R^2(cum) = 0.995 which means the OPLS-DA model has a high degree of fit and therefore most of the variance related to the response variable (Y) can be explained. In addition, the high value of Q^2(cum) = 0.758 means that the model has an acceptable degree of consistency between the original and predicted data (Figure 4.44). For investigating the impurities which bring about the difference among groups, Biplot was applied to illustrate the major contributing variables. It shows which of the variables are more highly correlated to

the given group of observations than to the others (Figure 4.45). This can also be confirmed with an S-plot of each group.

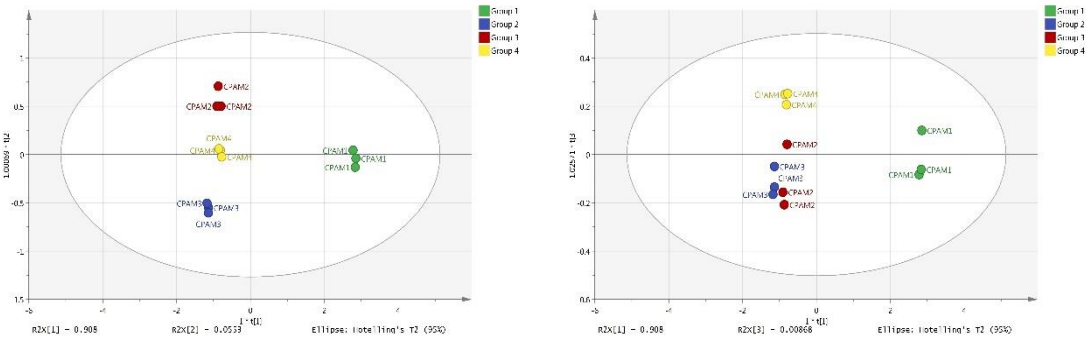


Figure 4.44 OPLS-DA scores scatter plot (t1 vs. t2) (left) and (t1 vs. t3) (right) for the four samples of chlorpheniramine from RPLC-MS analysis.

The plot shows distribution of 12 observations based on readings of 18 variables. The model includes four groups: CPAM1 (Green), CPAM2 (Red), CPAM3 (Blue), and CPAM4 (Yellow). It consists of one predictive x-score component; component t[1] and one orthogonal x-score components to[1]. (A) t[1] explains 90.5% of the predictive variation in x, to[1] explains 5.5% of the orthogonal variation in x, R^2X (cum) = 0.995, R^2Y (cum) = 1, R^2 (cum) = 0.922, Accuracy of prediction Q^2 (cum) = 0.758.

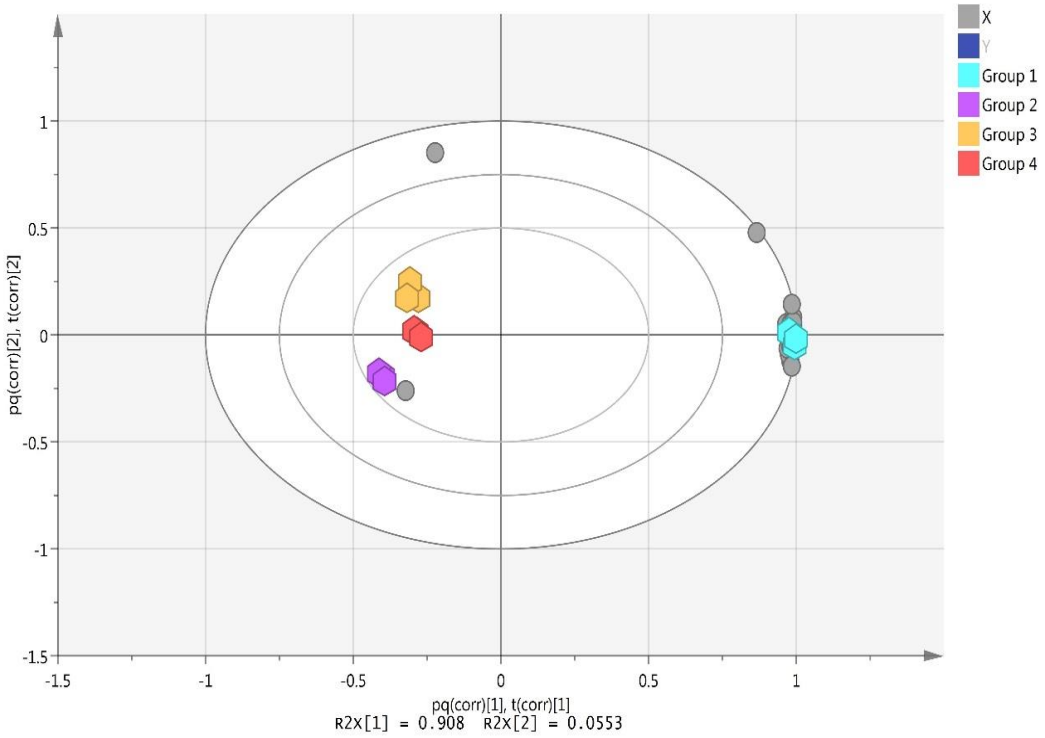


Figure 4.45 Biplot: scores (samples) and loadings (impurity) with low discrimination power on t1 vs. t2 from RPLC-MS analysis.

The plot shows distribution of 12 observations from four samples and 18 variables. The model includes four groups: CPAM1 (Blue), CPAM2 (Purple), CPAM3 (Orange), and CPAM4 (Red). It consists of one predictive x-score component; component t(corr)[1] and one orthogonal x-score components t(corr)[2] and x-variables component; component pq(corr)[1] and one orthogonal x-variables components pq(corr)[2].

OPLS-DA is used to predict the significant variables (i.e. impurities) which have a direct effect on each observation (i.e. sample). Variable importance in projection (VIP) was proposed as a variable selection method. A VIP value > 0.5 indicates that the variable is important to the model and the value < 0.5 is considered an unimportant variable. Figure 4.46 contains the significant variables (*red*) which was obtained from for VIP value > 0.5 and for where the jack-knife standard error did not cross zero.

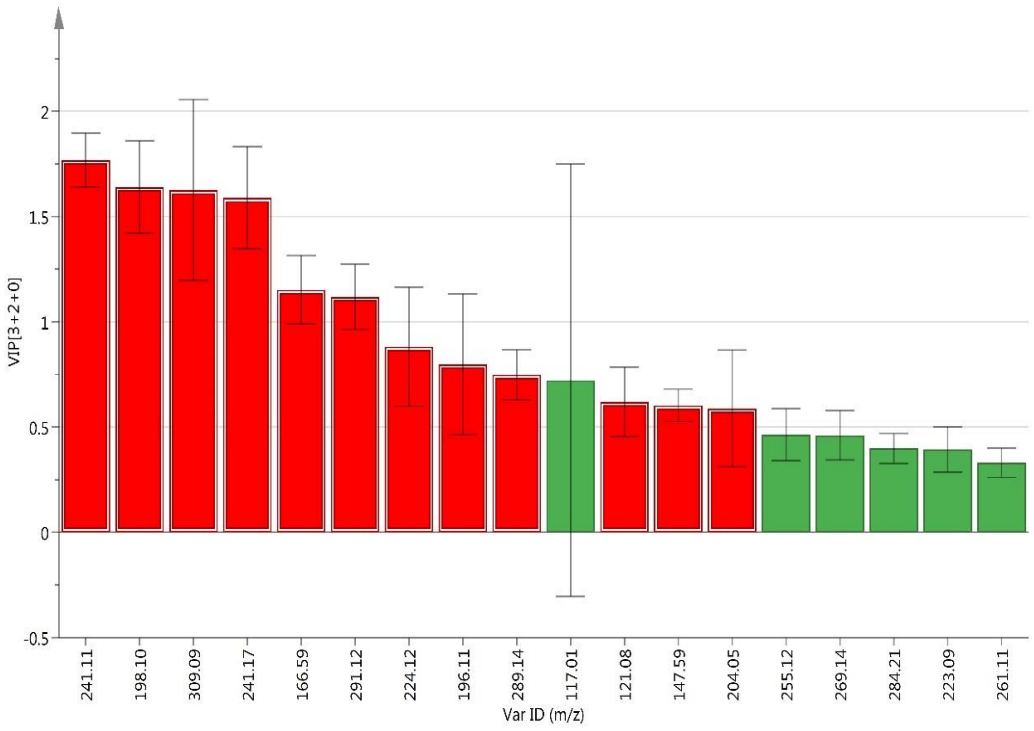
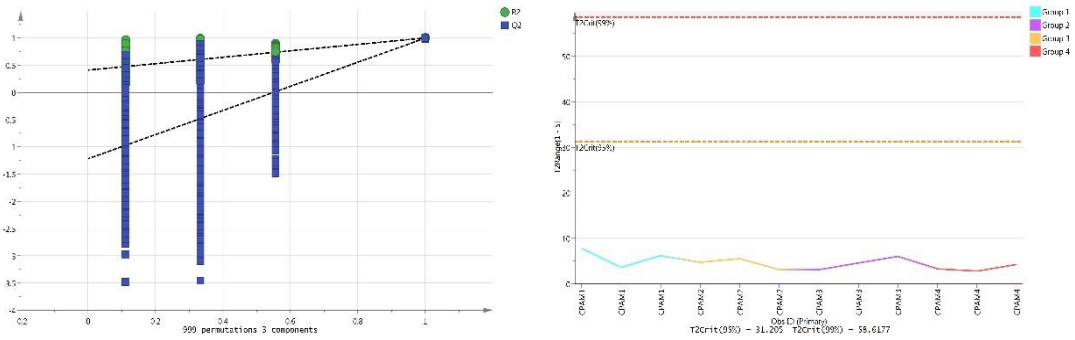


Figure 4.46 VIP plot of the variables of chlorpheniramine samples with the jack-knife standard error at (95%CI) of the VIP computed form all rounds of cross validation from RPLC-MS analysis model.
The X-axis represents the m/z ions and y-axis shows the VIP value. The red columns are significant variables have (VIP > 0.5) can be considered as important impurities.

The supervised classification model was validated by using permutation test as a cross-validation method to optimise the model fittings and to assess the performance parameters. As shown in the permutation test (Figure 4.47(A)), the y-axis gives the value of intercepts R^2 = (0.0, 0.414), Q^2 = (0.0, -1.2) which is related to the goodness of fit and prediction, respectively. The x-axis characterises the correlation between

the permuted y-vectors (on the left) and the original y-vector (on the right). The model could be considered statistically significant since the regression line of the Q^2 points intersects the y-axis below zero. In addition, the hotelling's T2 line plot displays the distance from the origin in the model plane (score space) for each selected observation. Values larger than the 95% confidence limit are suspect, and values larger than the 99% confidence limit can be considered outliers. (Figure 4.47(B)).



A B

Figure 4.47 (A) Validation plot obtained from the permutation test (B) Hotelling's T2 line plot displays the distance from the origin in the model plane (score space) for each selected observation from RPLC-MS analysis model.

Intercepts: R^2 = (0.0, 0.414), Q^2 = (0.0, -1.2) and the permutation number (n = 999). T2Crit (95%) confidence limit is the warning limit and represented by yellow dotted line, and T2Crit (99%) confidence limit is the action limit and represented by red dotted line. Observations above the action limit considered strong outliers. T2Range is basically calculated as the sum over the selected range of components of the scores in square divided by their standard deviations in square.

The Receiver Operator Characteristic (ROC) curve shows the robustness of the OPLS-DA model obtained from the cross-validated predicted Y (predicted class). It shows that (AUC = 1.0) which means the prediction model of OPLS-DA has great sensitivity and specificity (i.e. AUC > 0.5 corresponds to a good predictive model) (Figure 4.48).

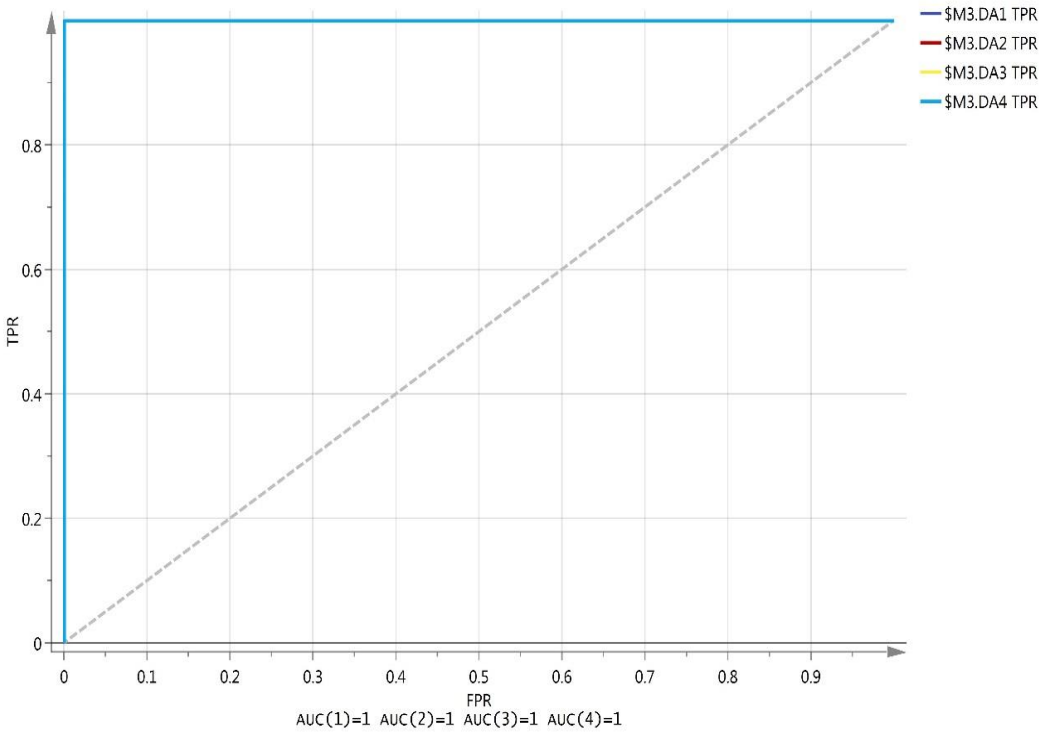


Figure 4.48 Area Under the Receiver Operating Characteristics Curve (ROC) curve calculated from the cross-validated predicted Y-values of the OPLS-DA model from RPLC-MS analysis.
The ROC curve shows sensitivity (true positive rate (TPR)) on the y-axis versus (false positive rate (FPR = 1 - Specificity)) on the x-axis, the value of both normalised to 1 which represents the value of AUROCC for group. All four groups of samples with excellent AUROCC accuracy (i.e. = 1).

4.3.2.2.3 Selected important variables from RP-LC analysis

The selected important variables according to VIP values were thoroughly investigated for the RPLC-MS data. The models were built based on the division of the samples in OPLS-DA hierarchical clustering variability index values (Figure 4.49). Three models were generated including comparison between the groups (CPAM1 vs group of CPAM2, 3 & 4), (CPAM3 vs group of CPAM2 & 4), and (CPAM2 vs CPAM4). All new OPLS-DA models showed a high cross validation and the permutation test has intercepts for Q^2 below zero (y-axis) which shows the goodness of fit and prediction performance (Table 4.6). The high value of validation in this model indicates that the selected variables are chosen correctly and presents the most important variables in the models.

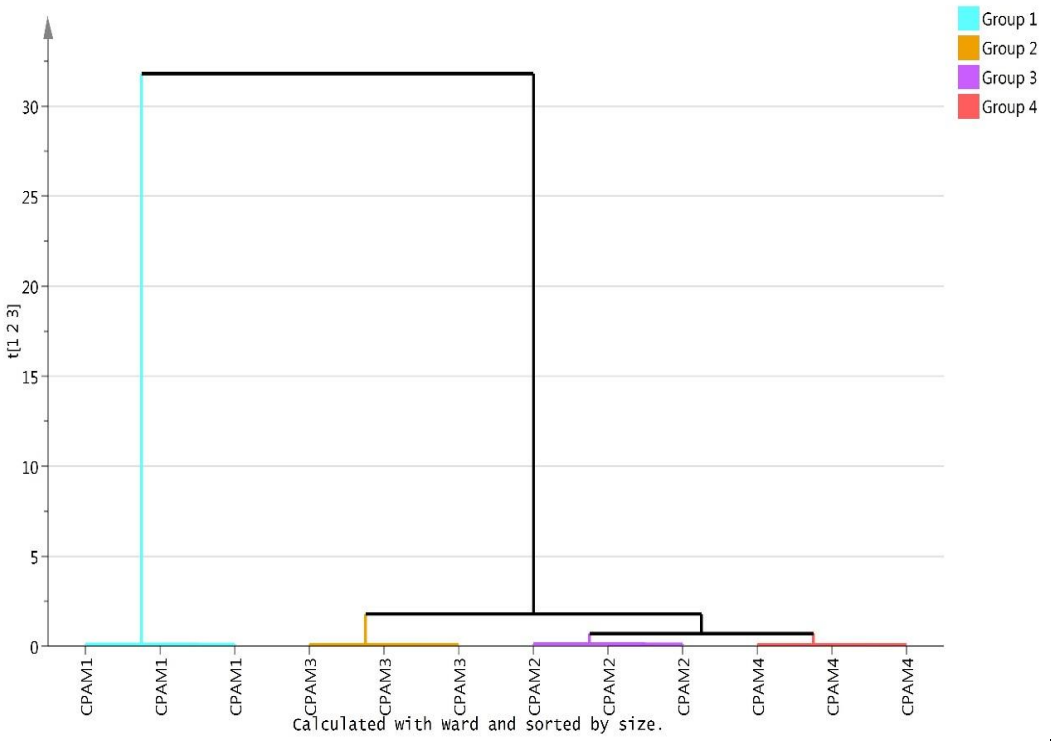


Figure 4.49 Hierarchical Clustering Analysis (HCA) plot for the four samples of chlorpheniramine related to OPLS-DA model from RPLC-MS analysis

The dendrogram shows observations clustered into four hierarchical groups; (CPAM1, blue), (CPAM2, purple), (CPAM3, orange) and (CPAM4, red) by using Ward clustering method. X-axis represents the samples and y-axis shows the variability index). The larger values show big differences between group variability and the lower values show more similarity. Thus CPAM 2 and 4 are very close in character which CPAM 1 is quite different from the other 3 samples.

Table 4.6 The cross validation and permutation test plot with intercepts at the permutation number (n = 999) for chlorpromazine models from RPLC-MS analysis.

OPLS-DA Model	Cross validation		Permutation test	*P*-value
Component	R^2(cum)	Q^2(cum)	Intercepts	(< 0.05)
CPAM1 Model	0.965	0.997	R^2=(0.0, 0.157), Q^2=(0.0, -0.749)	2.79E-08
CPAM3 Model	0.995	0.979	R^2=(0.0, 0.433), Q^2=(0.0, -0.78)	1.28E-03
CPAM2 vs 4 Model	0.997	0.912	R^2=(0.0, 0.77), Q^2=(0.0, -0.693)	0.12905

S-plot was implemented from OPLS-DA models to obtain on the variables that were important in each group. The selected important variables (red points) in the S-plot shows the most relevant variables that contribute in the variation in the model. These potential variables were investigated to confirm differences in their average intensity levels by considering the p-values (< 0.05) and 95% CIs.

Figures 4.50 and 4.51 show the important variables which are selected related their significance value in the in S-plot for CPAM1 and CPAM3, respectively. The important variables select based on their VIP value (VIP > 1) with jack-knife standard error.

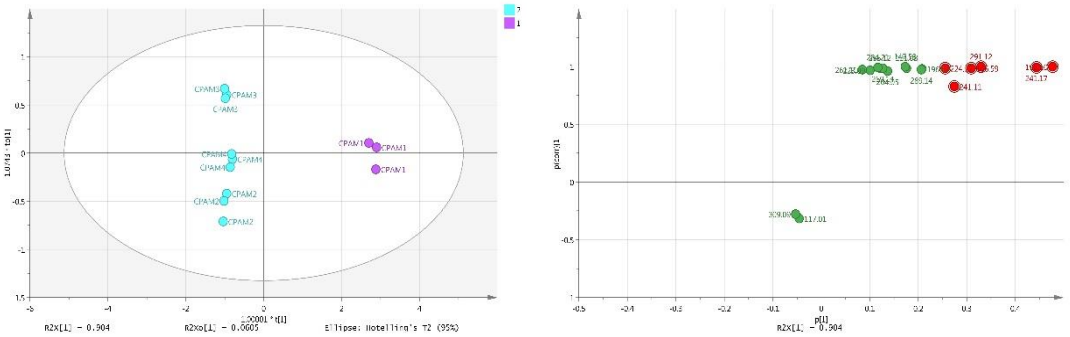


Figure 4.50 the scores plot (left) and S-plots at the first component of OPLS-DA (CPAM1 vs group 2, 3 & 4) from RPLC-MS analysis.
S-plot shows impurities distributed according to its correlation (x-axis) and magnitude of effect (y-axis) of the CPAM1 sample. The red marked points represent selected significant impurities.

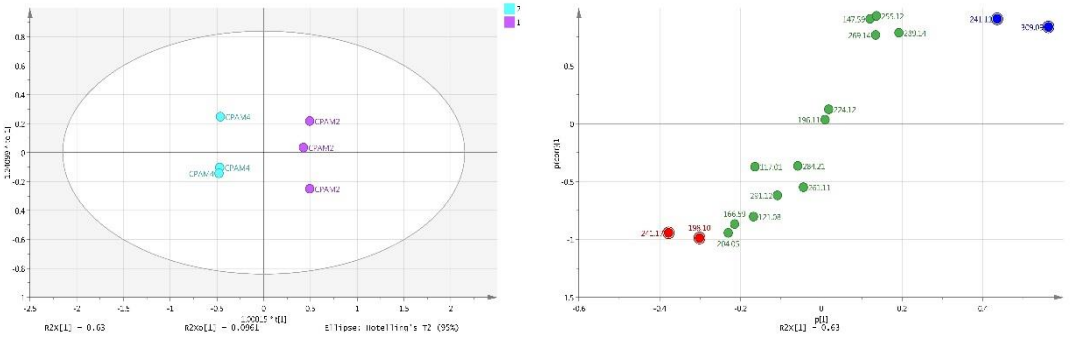


Figure 4.51 the scores plot (left) and S-plots at the first component of OPLS-DA (CPAM3 vs group 4) from RPLC-MS analysis.
S-plot shows impurities distributed according to its correlation (x-axis) and magnitude of effect (y-axis) of the CPAM2 sample (blue) and the CPAM4 sample (red). The marked points represent selected significant impurities.

Although, the model of CPAM 2 & 4 is considered a valid model as judged by the cross validation and permutation test, it shows an insignificant p-value. That is due to the similarity of the impurity profiles of the two samples.

4.3.3 Synthesis and degradation pathways

Table 4.7 shows the important variables with their VIP values with jack-knife standard errors which discriminate between the samples for both HILIC and RPLC analysis.

Table 4.7 The selected important variables in the CPAM1, CPAM2, CPAM3 and CPAM4 samples.

Sample	HILIC-MS model				RPLC-MS model			
	m/z	*p-value**	VIP	VIPcvSE**	m/z	*p-value**	VIP	VIPcvSE**
CPAM1	172.094	1.35E-17	0.62	0.18	166.598	3.20E-08	1.3	0.13
	223.100	2.50E-11	1.19	0.4	198.104	3.74E-15	1.99	0.11
	274.136	3.08E-13	0.78	0.17	241.110	9.95E-04	1.26	0.25
	300.124	5.75E-13	1.71	0.55	241.170	2.05E-10	1.86	0.25
					291.126	4.42E-13	1.37	0.03
CPAM2	198.110	4.96E-13	1.87	0.02	241.110	1.92E-02	1.74	0.94
	291.126	3.41E-12	2.52	0.05	309.092	3.51E-02	2.34	1.38
	309.093	1.23E-08	1.39	0.12				
CPAM3	291.126	3.03E-07	2.11	0.28	241.17	1.22E-02	1.21	0.35
CPAM4	297.079	6.84E-10	1.2	0.04	241.17	5.37E-03	1.51	0.61
	300.124	2.37E-08	1.33	0.09	309.092	3.51E-02	2.34	1.38
	309.093	3.36E-06	1.94	0.07				

*p-value < 0.05, **VIPcvSE: Jack-knife standard error of VIP.

Chlorpheniramine sample 1 had the highest number of significant impurities. From both HILIC and RPLC analysis, seven ions had VIP value more than 1. Impurity D at m/z 300 and UK2 at m/z 223.10, which have an additional CN group on the γ-position, were the major ions from HILIC producing a separation. UK1, UK3 and UK6, which lack pyridine, chlorine and result from oxidation of chlorpheniramine, respectively, are shown as the most significant impurities from analysis of the RPLC data. The proposed related substances pathway of sample CPAM1 is shown in figure 4.52. These lead to the strong suggestion that the synthetic route of CPAM1 is via route (1) (Figure 4.2).

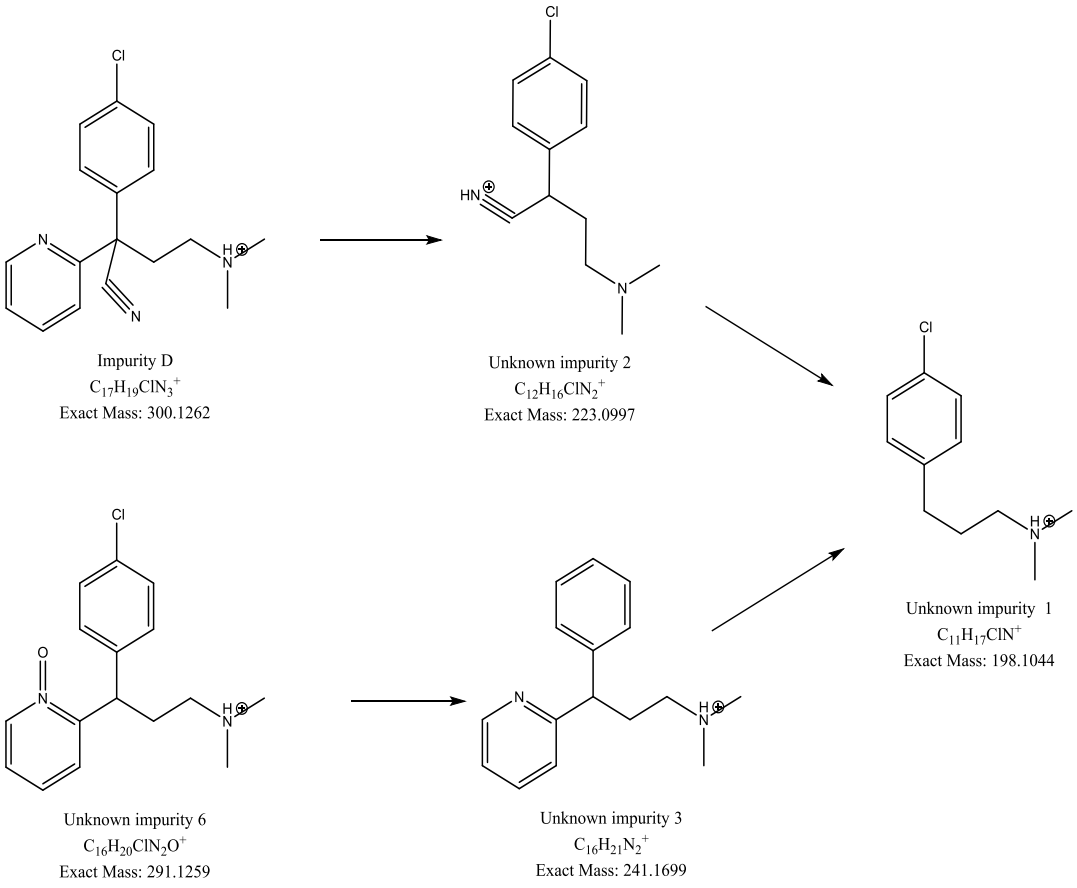


Figure 4.52 The proposed related substances pathway of CPAM1.

Chlorpheniramine sample 2 shows that the highest VIP value for impurities at m/z 291.12 (VIP = 2.52) and m/z 198.10 (VIP = 1.87) which are related to the oxidation product (UK6) product and the absence of pyridine in chlorpheniramine. In addition, the impurity at m/z 309.093 which is significant in both HILIC and RPLC analysis contains two chlorine atoms that might arise from small amounts of a dichlorobenzyl in the chlorobenzyl starting material. Such an intermediate could be present in all the starting materials used in routes 1-3.

Chlorpheniramine sample 3, is the S-(+) form of chlorpheniramine. The synthesis of the pure isomer form is achieved by using D-phenylsuccinic acid to separate the racemic forms of chlorpheniramine (Vardanyan and Hruby 2006).

The sample has fewer total impurities among the samples. The impurity at m/z at 241.17 (UK3) is related to RPLC analysis and ion at m/z 291.12 (UK6) was shown to have a high VIP value from HILIC analysis.

Chlorpheniramine sample 4 shows a unique impurity for this sample at m/z 297.08 containing an extra pyridine ring replacing the aliphatic chain. The presence of Impurity D, UK3 and UK7 suggests that the manufacturing process for CPAM4 is via route (1) (Figure 4.2).

4.4 Conclusions

In this work, novel impurities and degradation products were detected and identified by using orthogonal methods. HILIC-MS, RPLC/MS and LC-MS/MS methods were used in a purpose of structure elucidation of chlorpheniramine impurities. Some samples were successfully classified by SIMCA-P in accordance to their impurity profiles associated with their manufacturer/source synthetic route. Thus it is clearly possible to classify such generic materials according to their impurity profiles and this would make possible to spot fake branded products using this type of technique on the basis of the pattern of impurities in the API.

Chapter 5

Application of ion chromatography with suppressed conductivity detection for profiling of trace ionic impurities in pharmaceutical products

5 Application of ion chromatography with suppressed conductivity detection for profiling of trace ionic impurities in pharmaceutical products

5.1 Introduction

The requirement for systematic appraisal of impurity profiles of drug substances within the pharmaceutical industries, as well as regulatory requirements and compendial monographs, has increasingly led to high demands for suitable methods for the analysis, characterisation, and control of impurities at trace levels in order to ascertain the quality and safety of APIs. Currently, counterions are widely used with polar ionisable drug substances which necessitates control of the purity of the salt form in the API (Stahl and Wermuth, 2002). The counterion may be an inorganic ion, organic acid or base which is then paired with the ionisable drug substance, which itself may be acidic or basic, to form a salt form with adequate physiochemical and biopharmaceutical properties such as aqueous solubility and bioavailability. The salt form has a high impact on the quality, safety and efficacy of drug substances (Bansal et al., 2008, Stahl and Wermuth, 2002, Rocheleau, 2008, Stepanova and Kasicka, 2014) and a pharmaceutical drug substance which is contaminated with impurities can cause undesirable effects that may include risks of toxicity (Ahuja, 2007, Stepanova and Kasicka, 2014).

Organic impurities can arise during the process of synthesis, purification, and storage of the API. The presence of inorganic impurities is usually related to the process of manufacturing. Therefore, assessment of these impurities is required during the drug development process in order to comply with pharmacopoeial standards and toxicity levels (ICH-Q3A(R2), 25 October 2006). Effectively minimizing the potential of genotoxic impurities during the synthesis of salt forms is a challenge for the pharmaceutical industry. One such challenge is the detection of the formation of several alkyl sulphonates which result from the reaction between sulphonates and an alcohol solvent used during the manufacturing process. This is because the

sulphonic acid salt-forming moiety is difficult to analyse (Elder and Snodin, 2009). It has been suggested that all the impurities that can possibly be formed from the synthetic route and the preparation of final salt form of a drug substance should be evaluated in terms of their toxic potential (Verbeeck et al., 2006). In addition, the crystallization process can be significantly affected by the existing trace amounts of impurities in the API. The impurities may affect the primary and secondary nucleation kinetics or change the solubility and/or crystal shape of the API by altering the crystallographic surface growth rate (Nagy et al., 2008, Mullin, 2001). Thus, appropriate methods of analysis to identify and quantify the exact counterions and their stoichiometry at all stages of drug discovery, development and consistency between manufactured batches of drug substances are required (Stepanova and Kasicka, 2014). Several techniques have been employed in the evaluation of small organic acids and bases, and inorganic ions during pharmaceutical analysis. The main methods used are ion chromatography (IC), flame atomic absorption spectrometry, potentiometry with an ion–selective electrode, and flask-based methods such as titration (F. A. Chmilenko et al., 2000, Fritz, 2000). Capillary ion electrophoresis (CIE) with indirect UV detection has also been applied to the determination of counterions content in APIs (Rocheleau, 2008). Additionally, CIE with conductivity detection has been used for the screening of specific inorganic impurities and short chain carboxylic acids (Wiliams et al., 1997) and potassium (Williams and Boucher, 2000) in pharmaceutical drug substances.

Ion chromatography is the main technique for the separation of organic and inorganic ions (Weiss, 2004). The suppressed conductivity detector has been used widely with IC since it provides greater sensitivity and stability. It can decrease the background conductivity of the eluent which leads to signal enhancement of the components, thus offering more efficiency and robustness (Haddad et al., 2003). Ion chromatography with conductivity detection is often used for the separation of counterions including inorganic ions, and acidic and basic organic compounds. Techniques such as capillary electrophoresis and indirect UV detection may be used as alternative techniques alongside hydrophilic interaction chromatography (HILIC)

with evaporative light-scattering detector (ELSD) or refractive index detection for separation of counterions in pharmaceuticals. The use of suppressed conductivity detectors with IC can provide greater detection sensitivity since the eluents used can be changed from a high conductivity eluent in non-suppressed IC to water or low-conductivity carbonic acid. The baseline and drift are then significantly reduced which allows detection levels to reach low ppm values relative to the drug substance (Rocheleau, 2008).

The aim of this investigation was to provide an assessment of inorganic, and acidic organic impurities of selected drug salt forms for several pharmaceutical drug substances which were commercially available. The API standards that were chosen for the investigation contained different counterions in order to determine whether or not the chosen counterion had a bearing on the levels of anionic impurities. A high-pressure capillary ion chromatography (IC) with suppressed conductivity detection method was used for the separation and determination of the profile of the inorganic and organic counterions, in the APIs. This technique was applied to achieve a high separation resolution of the counterions and was optimized to obtain a high sensitivity of detection.

5.2 Methodology

5.2.1 Samples and materials

All the samples used in the study were commercially available drug substance reference standard materials. Oxytocin acetate hydrate, leucine enkephalin acetate, lanreotide acetate and chlorhexidine acetate were all purchased from Sigma-Aldrich (Dorset, UK). The citrate-containing samples, alverine citrate and orphenadrine citrate, were purchased from Sigma Aldrich (Dorset UK). Pergolide mesylate and methane sulphonic acid were obtained from Sigma Aldrich (Dorset UK). The (±)-propranolol HCl, R-(+)-propranolol HCl, S-(-)-propranolol HCl and chlorpromazine HCl were obtained from Sigma-Aldrich (Dorset UK). The study on chlorpheniramine maleate used three samples of (±)-chlorpheniramine maleate purchased separately

one from Alfa Aesar (Morecambe UK) and two from Sigma (Dorset UK) respectively. S-(+)-chlorpheniramine maleate was purchased from Sigma Aldrich (Dorset UK).
The seven anion mixture was commercially available (ThermoFisher, Hemel Hempstead UK) and contained fluoride (20 mg/L), chloride (100 mg/L), nitrite (100 mg/L), bromide 100 mg/L, nitrate (100 mg/L), phosphate (200 mg/L), and sulphate (100 mg/L) and was used for identification of anions and for the preparation of a calibration curve.

Other standard solutions were prepared in-house using ammonium citrate, ammonium lactate, ammonium bromide, ammonium formate, sodium trifluoroacetate, sodium fumarate dibasic, butyric acid and ammonium acetate.

5.2.2 Samples preparation

All samples were diluted to a level of 20 mg/L in deionized water for IC analysis (Sigma Aldrich, Dorset UK.), and to a level of 50 mg/L for the samples containing acetate.

5.2.3 Quantification of anions

The amount of anions was detected after subtracting the area of the water peak (if applicable) from the anion according to their relative retention time (RRT) (relative to chloride). Then the concentration of anion was proportionally calculated to the response factor (RF) of each anion, which measured the ratio between the concentration of the standard anions being analysed and the peak area of that anion. At the end, the concentration of anion was multiplied by the dilution factor to provide the actual amount of anion in each compound. The unknown anions were calculated according to the RF of chloride.

5.2.4 Ion chromatography system

The anions were analysed using a capillary-IC system (see section 2.7).

5.3 Results and Discussion

A capillary ion chromatography system with suppressed conductivity detection method was used to characterize several pharmaceutical drug substances in terms of the profiles of small acidic and inorganic anionic impurities that might be present. The standard drug substances investigated contained different counterions such as chloride, maleate, acetate, citrate and mesylate. Relative retention times (RRTs), calculated by dividing the retention time of a chloride ion by that of the counterion, were used to characterise unknown impurities. The percentage composition of each anionic impurity in the sample was calculated relative to the level of the counterion concentration in the sample based on their response factors obtained for the standard mixture.

Peak identification in the samples of drug substances analysed was done by using known standard organic acids and inorganic compounds similarly treated and analysed by the same method. The RRT of these standards were determined and compared with those of unknown impurities to be identified. Table 5.1 summarises the concentrations of the anions present in the blank which was deionized water for ion chromatography, and figure 5.1 shows the chromatogram obtained for the deionized water blank. Thus in this type of analysis the LOD is defined by the background ions in the blank which is significant even in highly purified water.

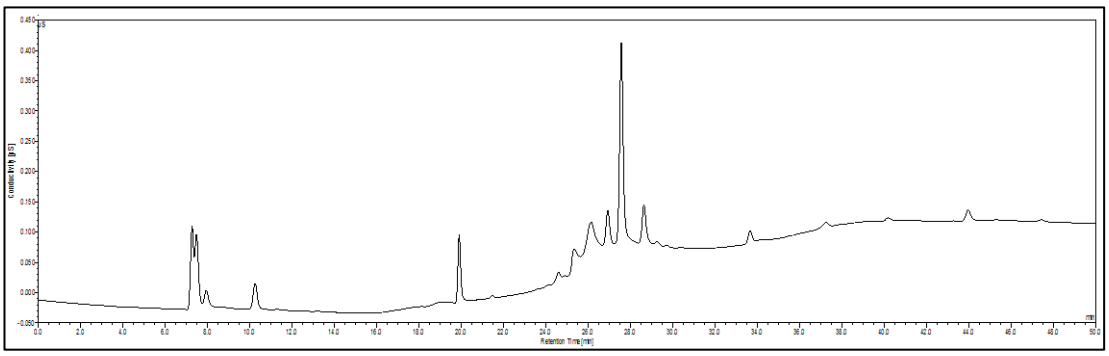


Figure 5.1 Chromatogram of the deionized water blank.

Table 5.1 The levels of anions present in the deionized water blank.

Anion	Blank IC water				Slope	3.3*SD/S	10*SD/S
	RT mins	RRT	Intensity µS	Intensity SD		LOD (µg/L)	LOQ (µg/L)
Unknown	7.28	0.37	0.07779	0.01162			
Acetate	7.94	0.41	0.01346	0.00077			
Unknown	10.34	0.52	0.04484	0.00133			
Chloride	19.9	1	0.10448	0.01652	0.69164	0.0788	0.2388
Nitrite	21.45	1.08	0.02273	0.00530			
Bromide	26.04	1.31	0.03835	0.00128	1.68205	0.0025	0.0076
Nitrate	26.86	1.35	0.02052	0.00120	1.11723	0.0036	0.0108
Sulphate	27.48	1.38	0.05597	0.00135	0.68503	0.0065	0.0197
Phosphate	33.53	1.69	0.01273	0.00263			

The separation of the mixture of anion standards is shown in figure 5.2. The method was calibrated over the range 0.5-10 µg/mL using a standard mixture containing seven anions that are, fluoride, chloride, bromide, nitrite, nitrate, sulphate and phosphate. The serial dilutions for the calibration were made by diluting the standard mixture with deionized water. Plus calibration curves were prepared for some additional ions not in the standard mixture.

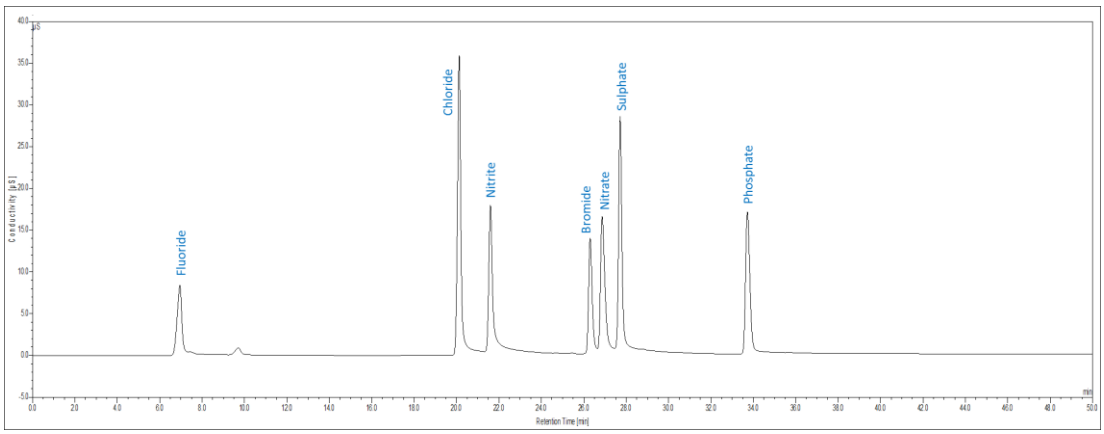


Figure 5.2 Chromatograms of separation for the mixture of seven anion standards
containing: fluoride (1 µg/L); chloride (5 µg/L); nitrite (5 µg/L); bromide (5 µg/L); nitrate (5 µg/L); phosphate (10 µg/L); sulphate (5 µg/L).

The calibration curve was made using anions standard materials with serial dilution to reach the expected concentration levels of the anion impurities (Table 5.2).

Table 5.2 Calibration curve results for the counterions*.

Anion	Range (µg/L)	R^2	Slope	Intercept
Bromide	0.5-10	0.9999	1.6572	0.1653
Chloride	0.5-10	0.9996	0.6916	0.0470
Nitrate	0.5-10	0.9998	1.1094	0.1136
Nitrite	0.5-10	0.9996	0.9154	0.1001
Sulphate	0.5-10	0.9932	0.6850	0.1428
Citrate	0.5-10	0.9995	0.3665	0.0332
Acetate	0.5-10	0.9983	0.3688	0.1094
Mesylate	0.5-20	0.9997	0.6766	-0.1754
Maleate	0.5-10	0.9999	1.3227	-0.0191

*Calibration curve based on linear regression for concentration levels (0.5, 1, 2, 5, and 10) µg/L of the three replicates analysis.

The retention time precision was determined for all the impurities detected in each sample. It is important for identification of trace levels of anionic impurities to have a high precision for the retention times for all detected peaks to minimise misidentification since retention times can be altered by the presence of the large amount of the major counter ion to the API. Retention time precision for each impurity peak was assessed in each sample analysed based on relative standard deviation (RSD) of the retention time relative to the chloride peak retention (Table 5.3).

Table 5.3 The retention time precision for the counterions.**

Anion	RT (mins)	RRT	SD	RSD (n=6)
Acetate	8.17	0.41	0.0208	0.2637
Mesylate	13.11	0.65	0.2924	2.32
Chloride	20.08	1	0.1419	0.7084
Nitrite	21.63	1.08	0.0833	0.3869
Trifluoroacetate	24.84	1.23	0.0087	0.0351
Bromide	26.24	1.31	0.0337	0.128
Nitrate	27.15	1.35	0.0626	0.2328
Maleate	27.18	1.36	0.0064	0.0237
Sulphate	27.73	1.38	0.0417	0.1505
Fumarate	28.82	1.44	0.0413	0.1434
Phosphate	33.88	1.69	0.0664	0.197
Citrate	35.73	1.78	0.0319	0.0903

** RRT related to the retention time of chloride.

5.3.1 Anionic Impurities in Propranolol HCl

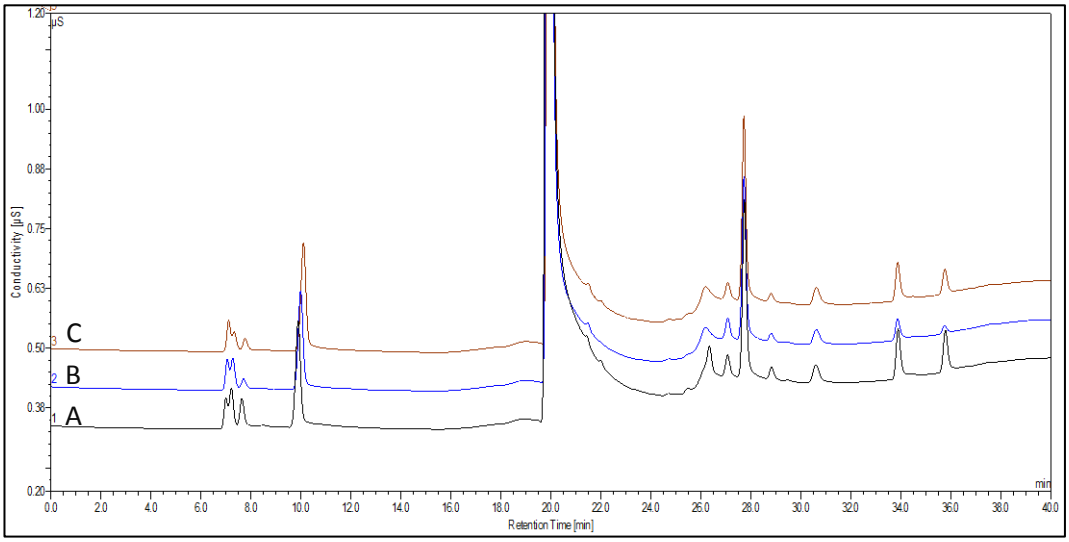


Figure 5.3 Chromatograms of the propranolol HCl samples. (A) (+/-) Propranolol HCl, (B) (R)-(+)-Propranolol HCl, (C) (S)-(-)-Propranolol HCl.
The retention time (min) presents on the x-axis and conductivity (µS) on the y-axis.

Three samples of Propranolol HCl were investigated, a racemic mixture and the R and S isomers and the major anionic impurities were identified. The IC for propranolol hydrochloride samples is shown in the figure 5.3. Impurities at RRT 0.52 and RRT 1.31

were detected as the most abundant anionic impurities in all propranolol HCl samples at levels up to 3% of the response of the chloride counterion peak. The peak at RRT 0.51 was an unknown impurity and RRT 1.31 was identified as bromide. Impurities RRT 1.35 and RRT 1.38 corresponded to nitrate and sulphate respectively, were found to be present at levels between 0.5 - 1.2% of the chloride peak. Impurity RRT 1.08 was putatively identified to be nitrite and was detected only in samples containing (R)-(+)-Propranolol HCl at levels of 6.8% of the chloride peak. Phosphate and citrate which are the impurities RRT 1.69 and RRT 1.78 were present in (±) Propranolol HCl and (S)-(-)- Propranolol HCl samples. Unknown impurities at RRT 1.53 and RRT 2.03 were detected in some samples (Figure 5.4). The anionic impurities results were calculated at the level of µg/L for the samples containing 1 mg/ml concentration of drug substance (Table 5.4). The retention time precision was measured and the RSD did not exceed 2% for all impurities.

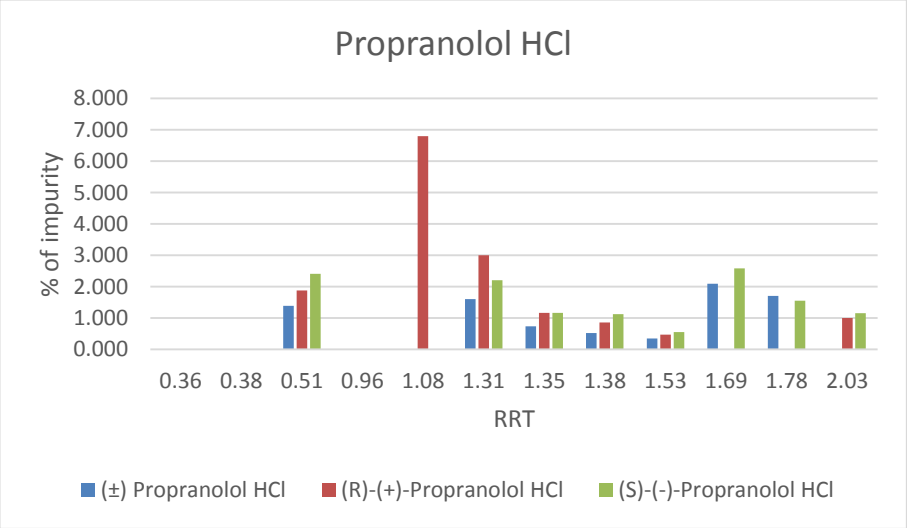


Figure 5.4 the percentage of trace anionic impurities in propranolol HCl samples relative to chloride counterion. The relative retention time (RRT) presents on the x-axis and % of impurity on the y-axis.

Table 5.4 The amounts of chloride and impurities (µg/L) of propranolol HCl samples.

Anion	RRT	Concentration of the impurity in each API (µg/L)		
		(±) Propranolol HCl	*(R)-(+)-Propranolol HCl*	*(S)-(-)-Propranolol HCl*
Unknown	**0.52**	2.352	2.088	2.675
Chloride	**1.00**	168.823	111.285	110.952
Nitrite	**1.08**	0.000	7.568	0.000
Bromide	**1.31**	2.714	3.341	2.443
Nitrate	**1.35**	1.249	1.299	1.291
Sulphate	**1.38**	0.886	0.952	1.252
Unknown	**1.53**	0.587	0.518	0.613
Phosphate	**1.69**	3.537	0.000	2.862
Citrate	**1.78**	2.873	0.000	1.718

The amount of chloride ion is 116.6% of the expected content suggesting that this sample is quite pure. However, the levels of chloride in the pure enantiomers is much lower. It is interesting that the level of anionic impurities is greater in the pure enantiomers which have a more complex processing history than the racemic mixture since they will have been resolved by fractional crystallisation of their diastereomeric salts.

5.3.2 Anionic Impurities in Chlorpromazine HCl

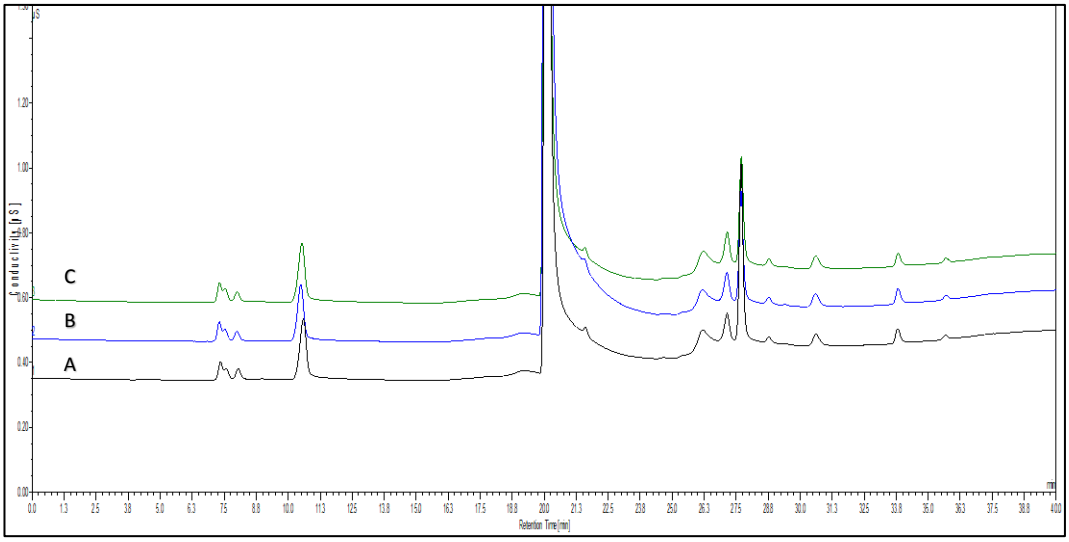


Figure 5.5 Chromatograms of the Chlorpromazine HCl samples.
(A) Chlorpromazine HCl-1, (B) Chlorpromazine HCl-2, (C) Chlorpromazine HCl-3. The retention time (*min*) presents on the x-axis and conductivity (μS) on the y-axis.

Chlorpromazine HCl samples were analysed to obtain an impurity profile of the small organic and inorganic anions that might be present. Three commercial standards of the Chlorpromazine HCl drug substance were analysed to detect any trace amounts of anionic impurities. As shown in Figures 5.5 and 5.6, there was disparity between the levels of impurities detected in the samples tested. Impurity RRT 1.08, which was putatively identified to be nitrite anion, and was detected as the most abundant present in Chlorpromazine HCl samples 1 and 3 at level up to 6.4% of the expected counterion peak chloride.

Chlorpromazine HCl-1 shows a higher level of impurity RRT 1.38, which corresponding to sulphate, than the other samples. Other impurities such RRT 1.31, RRT 1.35 and RRT 1.53 showed at levels up to 4.4% and their RRTs showed that they could be identified as bromide, nitrate and ketomalonate, respectively. The unknown impurity RRT 0.52 was also detected in the all samples and phosphate (RRT 1.69) was only detected in the Chlorpromazine-1. Chlorpromazine-2 shows lower levels of all impurities compared to the other chlorpromazine samples: this might be related to differences in the manufacturing sources. The anionic impurity results were

calculated at the level of µg/L for the samples which contained a 1 mg/L concentration of drug substance (Table 5.5). The RSD of retention times of each of the impurities was not more than 1.38 % which showed that there was stability in the chromatography.

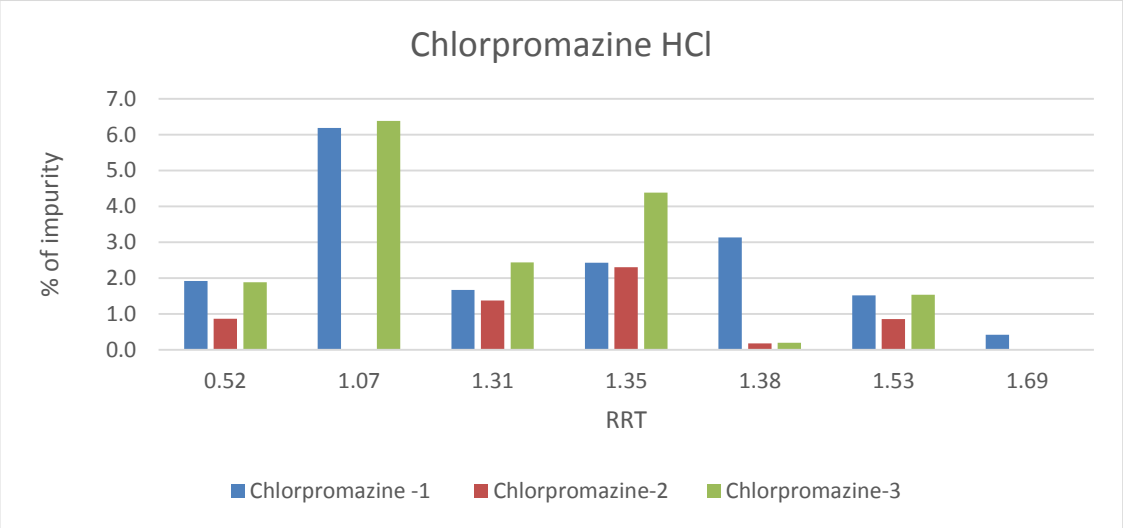


Figure 5.6 the percentage of trace anion impurities in Chlorpromazine HCl samples relative to chloride counterion.
The relative retention time (RRT) presents on the x-axis and % of impurity on the y-axis.

Table 5.5 The amounts of chloride and impurities (µg/L) in Chlorpromazine HCl samples.

Identified impurity	RRT	Concentration of the impurity in each API (µg/L)		
		Chlorpromazine-1	Chlorpromazine-2	Chlorpromazine-3
Unknown	**0.52**	2.402	1.942	2.367
Chloride	**1.00**	124.835	127.875	125.373
Nitrite	**1.08**	7.732	0.000	8.008
Bromide	**1.31**	2.271	3.349	3.328
Nitrate	**1.35**	1.084	1.839	1.961
Sulphate	**1.38**	3.915	0.392	0.244
Ketomalonate	**1.53**	0.617	0.628	0.628
Phosphate	**1.69**	1.625	0.000	0.000

5.3.3 Anionic Impurities in Chlorpheniramine maleate

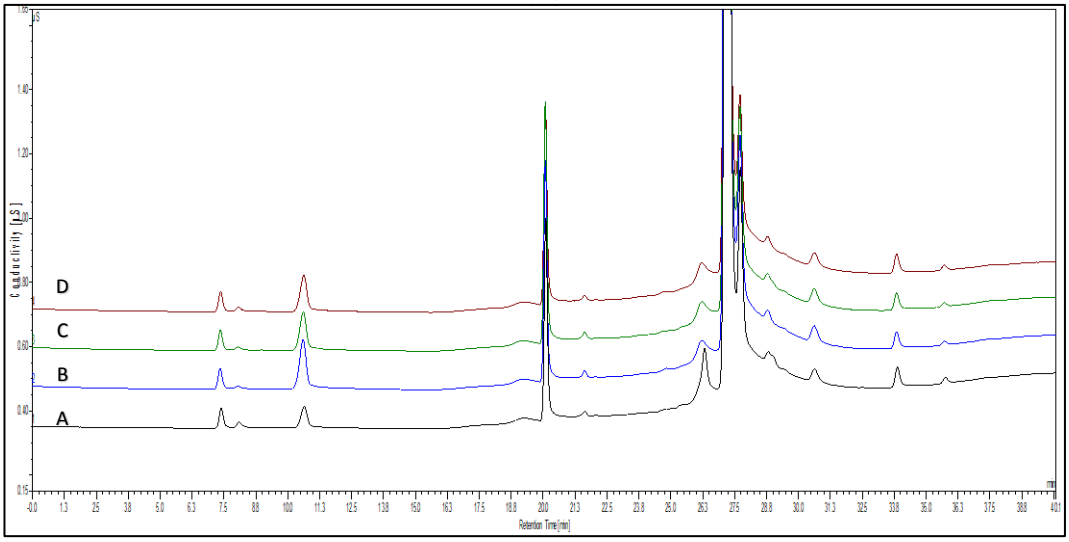


Figure 5.7 chromatograms of the Chlorpheniramine maleate samples.
(A) (±)-Chlorpheniramine maleate-1, (B) (±) Chlorpheniramine maleate-2, (C) S-(+)-Chlorpheniramine maleate-3 (D) Chlorpheniramine maleate-4. The retention time (*min*) presents on the x-axis and conductivity (*μS*) on the y-axis.

Four commercial drug substances of chlorpheniramine maleate were tested to determine the levels of the trace anionic impurities present. The ICs for chlorpheniramine maleate samples are shown in the Figure 5.7. In all four samples tested, the highest level of anionic impurity did not exceed 2.6% of the response of the drug counterion. Impurities RRT 1.00, RRT 1.31 and RRT 1.38 were detected as the most abundant impurities at levels between 0.5 - 2.6 % of the level of the counterion peak and were identified as chloride, bromide and sulphate respectively. The impurities at RRT 1.44 (fumarate), and RRT 1.53 (ketomalonate) were detected at levels of less than 0.4% of the level of the counterion peak (Figure 5.8). The anionic impurities results were calculated at the level of µg/L for the samples containing 1 mg/L concentration of drug substance (Table 5.6). The retention time precision measurement showed most of the impurities to have RSD less than 0.8% with the exception of the unknown impurity RRT 0.52 where the retention time RSD was 3.4%. The purity of the chlorpheniramine maleate with respect to anionic contamination was good.

Table 5.6 The amounts of maleate and impurities (µg/L) in Chlorpheniramine maleate samples.

Anion	RRT	Concentration of the impurity in each API (µg/L)			
		(±)-Chlorpheniramine 1	(±)-Chlorpheniramine 2	S-(+)-Chlorpheniramine 3	Chlorpheniramine 4
Unknown	**0.52**	0.263	1.728	1.372	1.355
Chloride	**1**	4.287	4.927	5.718	4.766
Bromide	**1.31**	4.078	2.010	2.433	2.420
Maleate	**1.36**	435.934	496.991	480.455	459.351
Sulphate	**1.38**	11.237	11.663	11.600	10.518
Fumarate	**1.44**	0.501	0	0.051	0
Ketomalonate	**1.53**	0.369	0.481	0.465	0.411
Phosphate	**1.69**	0.767	0.662	0.672	0.776

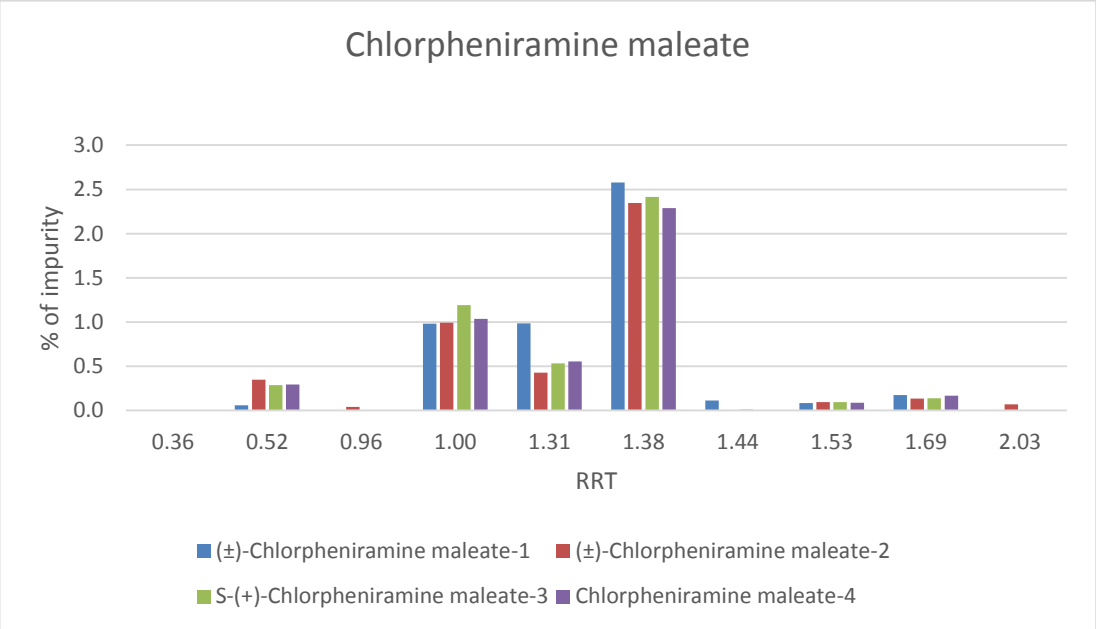


Figure 5.8 the percentage of trace anion impurities in Chlorpheniramine maleate samples relative to maleate counterion.
The relative retention time (RRT) presents on the x-axis and % of impurity on the y-axis.

5.3.4 Anionic Impurities in Drugs with Acetate as the Counterion

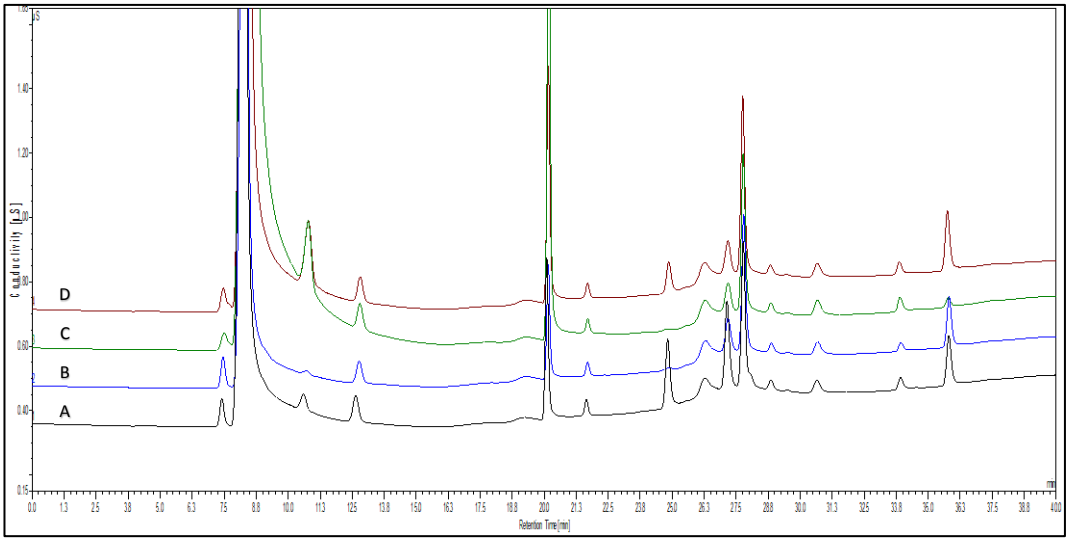


Figure 5.9 chromatograms of the Acetate samples.
(A) Oxytocin Acetate salt hydrate, (B) Leucine Enkephalin Acetate, (C) Lanreotide acetate, (D) Chlorhexidine acetate. The retention time (*min*) presents on the x-axis and conductivity (μS) on the y-axis.

Four different standard drug substances containing acetate counterions were analysed to investigate levels of any trace anionic impurities present. The hypothesis here was that acetate might be readily displaced by anions of stronger acids during processing leading to a lower than expected acetate level, or higher levels of the other strong anions. For this purpose, oxytocin acetate, leucine enkephalin acetate, lanreotide acetate and chlorhexidine acetate were tested. The IC for acetate samples is shown in the figure 5.9. The retention time of the chloride counterion was used as denominator to calculate the RRT. The major trace anionic peaks RRT 1.00, RRT 1.31, RRT 1.35, RRT 1.38 and RRT 1.78 were identified as chloride, bromide, nitrate, sulphate, and citrate, respectively.

There were disparities in measured impurity levels between the samples tested. The impurities level of nitrate is the highest anion peak and shows varying levels in the acetate samples. It was detected in oxytocin acetate and leucine enkephalin acetate at approximately 5.3% and 2.5%, respectively, and less than 0.5% in the remaining samples (Figure 5.10). The impurities RRT 1.08, RRT 1.24, and RRT 1.53 - had shown a very low level - were putatively identified as nitrite, trifluoroacetate and

ketomalonate, respectively. Unknown impurities RRT 0.53 and RRT 0.63 were detected in some samples. The anionic impurities results were calculated at the level of µg/L for the samples containing 1 mg/L concentration of drug substance (Table 5.7). The RSD of retention time precision did not exceed 2% for all impurities. The results did not bear out the hypothesis that acetate might be displaced by anions of stronger acids since the impurities were no higher than for instance that those in the chlorides discussed above. Although there about 5% of the anion composition is something other than counterion, the acetate samples have the high number of impurities. The lowest impurity levels were in chlorhexidine which is a strongly basic drug and thus will produce alkaline solutions due to salt hydrolysis ensuring that acetate remains fully ionised.

Table 5.7 The amounts of acetate and impurities (µg/L) in samples containing Acetate counterion.

		Concentration of the impurity in each API (µg/L)			
Identified impurity	**RRT**	**Oxytocin Acetate**	**Leucine Enkephalin Acetate**	**Chlorohexidine Acetate**	**Lanreotide Acetate**
Acetate	**0.41**	111.403	95.462	322.464	169.689
Unknown	**0.53**	2.137	0.000	8.416	3.775
Unknown	**0.63**	0.599	0.468	0.387	1.133
Chloride	**1.00**	1.269	0.765	4.802	2.004
Nitrite	**1.08**	0.000	0.000	0.511	0.000
Trifluoroacetate	**1.24**	0.931	0.000	0.000	0.432
Bromide	**1.31**	0.334	1.313	0.253	0.246
Nitrate	**1.35**	5.839	2.236	0.239	0.325
Sulphate	**1.38**	0.544	0.486	1.663	2.062
Ketomalonate	**1.53**	0.246	0.247	0.293	0.277
Citrate	**1.78**	2.225	2.059	0.000	2.479

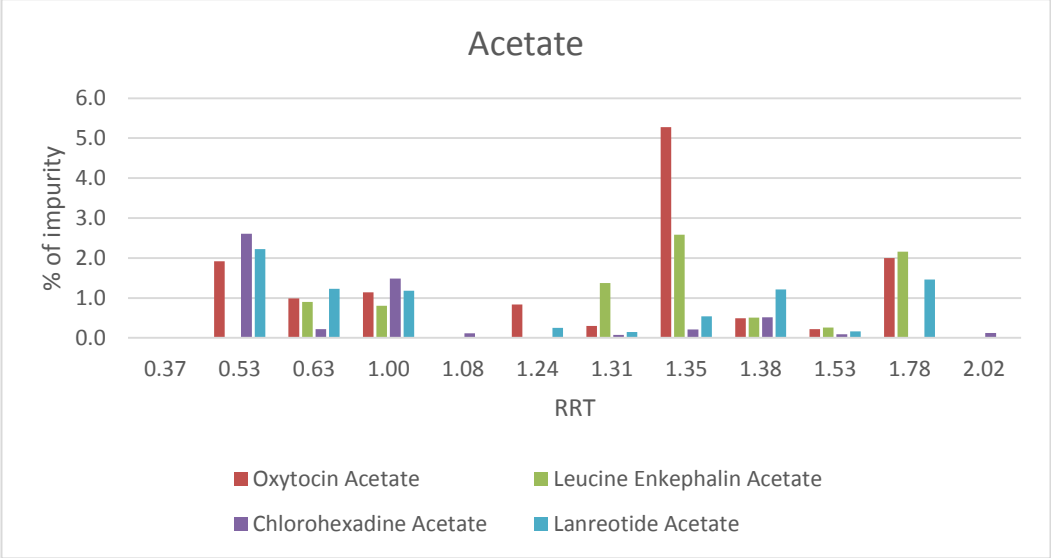


Figure 5.10 the percentage of trace anion impurities in samples containing Acetate counterion. The relative retention time (RRT) presents on the x-axis and % of impurity on the y-axis.

5.3.5 Anionic Impurities in Drugs with Citrate as the Counterion

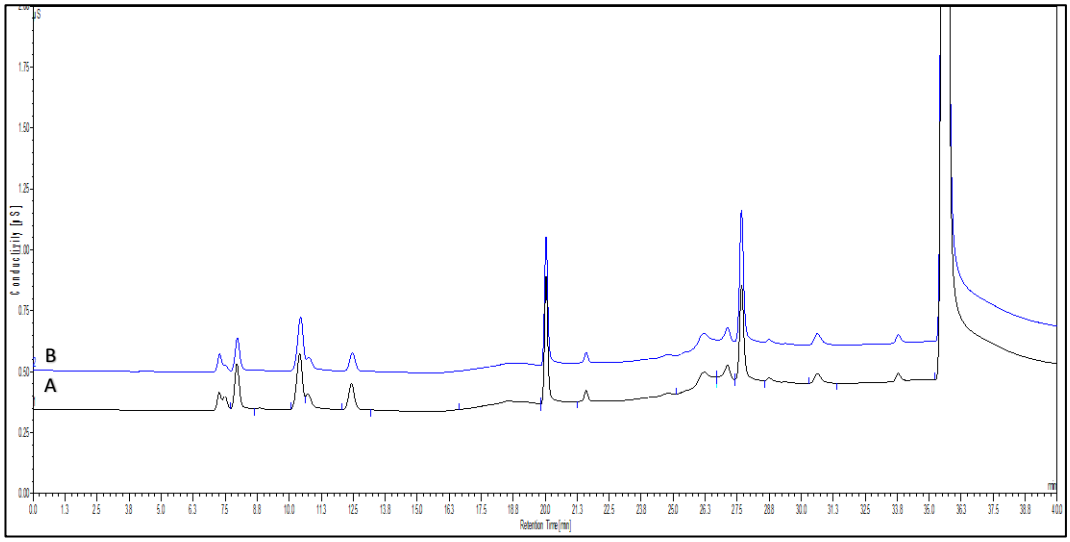


Figure 5.11 chromatograms of the samples containing citrate counterion.
(A) Alverine Citrate, (B) Orphenadrine citrate. The retention time (min) presents on the x-axis and conductivity (µs) on the y-axis.

The investigation of trace anions in two samples of drug substances each containing the citrate counterion was carried out. The IC for citrate samples is shown in the figure 5.11. The anionic impurities in alverine citrate and orphenadrine citrate drug substances were detected and identified using relative retention times and response factors relative to the citrate counterion peak. The major impurity in the samples was

RRT 0.41 with abundances of approximately 2% of the counterion peak in alverine citrate and 1.3% of the counterion peak in orphenadine citrate and it was putatively identified as acetate. The impurities RRT 1.00, RRT 1.31, RRT 1.35, RRT 1.38 and RRT 1.53 were observed in both samples with less than 1% (Figure 5.12) and they were identified as chloride, bromide, nitrate, sulphate and ketomalonate, respectively. Unidentified impurities RRT 0.52 and RRT 0.62 were detected in both samples. The anionic impurities results were calculated at the level of µg/mL for the samples containing 1 mg/L concentration of drug substance (Table 5.8). The retention time precision measurement showed RSD not exceeding 1% for all impurities detected in the samples. The overall impurities not shows difference between the two samples with total level 4.5%.

Table 5.8 The amounts of citrate and impurities (µg/L) in samples containing citrate counterion: Alverine Citrate and Orphenadrine citrate

Identified impurity	RRT	Concentration of the impurity in each API (µg/L)	
		Alverine Citrate	Orphenadrine Citrate
Acetate	**0.41**	8.286	5.526
Unknown	**0.52**	1.873	2.323
Unknown	**0.62**	1.400	0.985
Chloride	**1.00**	3.280	3.303
Bromide	**1.31**	2.464	2.767
Nitrate	**1.35**	1.497	1.621
Sulphate	**1.38**	0.398	1.955
Ketomalonate	**1.53**	0.653	0.765
Citrate	**1.78**	417.648	425.700

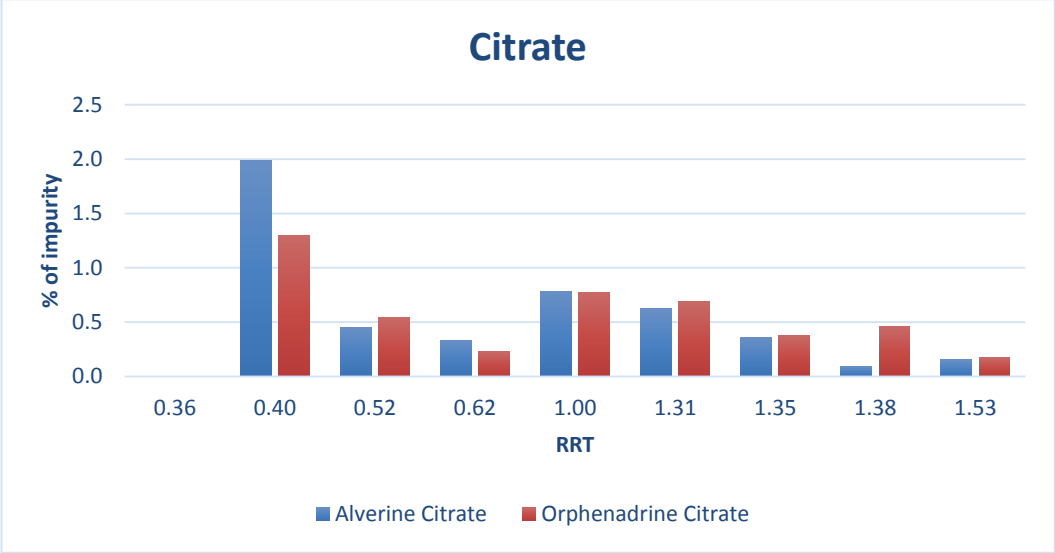


Figure 5.12 the percentage of trace anion impurities in samples containing citrate counterion: Alverine Citrate and Orphenadrine citrate. The relative retention time (RRT) presents on the x-axis and % of impurity on the y-axis**.**

5.3.6 Anionic impurities in a drug with Mesylate as the counterion

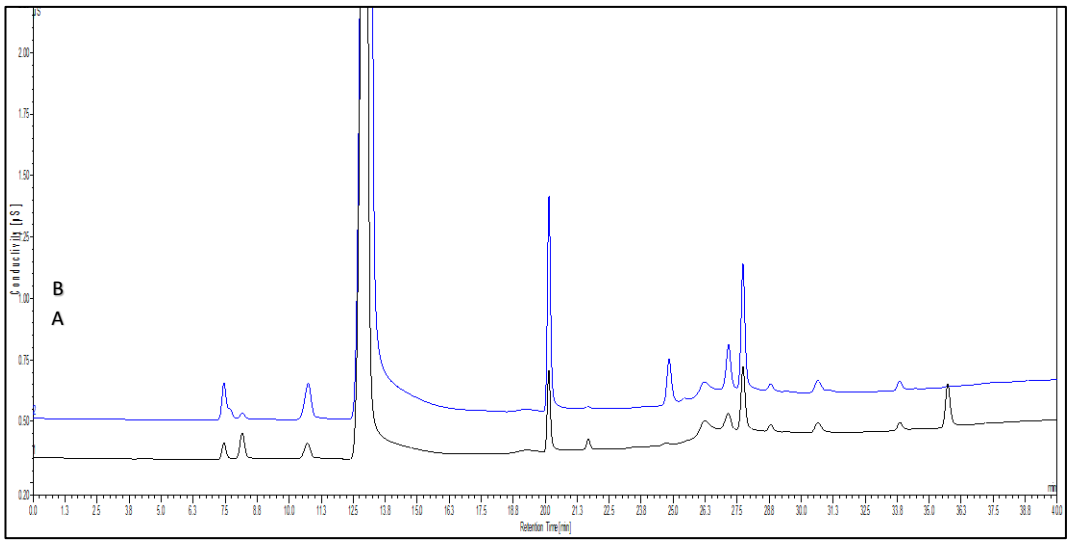


Figure 5.13 chromatograms of the samples containing mesylate counterion.
(A) Pergolide mesylate, (B) Methane sulphonic acid. The retention time presents (*min*) on the x-axis and conductivity (*µS*) on the y-axis.

The IC for trace anionic impurities was also determined for pergolide mesylate, in addition a (MSA) standard was run as shown in the figure 5.13. The major trace anionic impurities in pergolide mesylate had RRT 0.41 and RRT 1.78 with the levels up to 4.9% relative to the counterion peak which were putatively identified to be

acetate and citrate respectively. These impurities did not appear in the MSA sample. An impurity with RRT 1.35 was putatively identified as nitrate and was one of the major impurity in both samples. The level of nitrate impurity was at least 4% in pergolide mesylate and at least 2% in MSA. This suggests that nitrate might be related to methane sulphonic acid and during crystallisation may associate more strongly with the drug than MSA. Impurities RRT 1.00 and RRT 1.31 which appeared in the both samples corresponded to chloride and bromide, respectively. The impurity RRT 1.38 which was related to sulphate was only found in the MSA at the level of 1% of the counterion peak. Unknown impurities RRT 0.54 and RRT 1.23 were detected in both samples (Figure 5.14). The anionic impurity results were calculated at the level of µg/L for the samples containing 1 mg/L concentration of drug substance (Table 5.9). The retention time precision calculation gave RSD of < 1.4% for all the impurities in the samples.

Table 5.9 The amounts of mesylate and impurities (µg/L) in samples containing mesylate counterion: Pergolide mesylate and MSA.

Identified impurity	RRT	Concentration of the impurity in each API (µg/L)	
		Pergolide mesylate	MSA
Unknown	**0.37**	0.000	2.274
Acetate	**0.41**	5.905	0.000
Unknown	**0.54**	1.520	8.382
Mesylate	**0.65**	128.599	999.254
Chloride	**1.00**	1.808	22.176
Trifluoroacetate	**1.23**	2.590	7.386
Bromide	**1.30**	3.423	7.777
Nitrate	**1.35**	5.297	19.080
Sulphate	**1.38**	0.000	10.509
Ketomalonate	**1.52**	0.000	3.432
Citrate	**1.77**	6.283	0.000

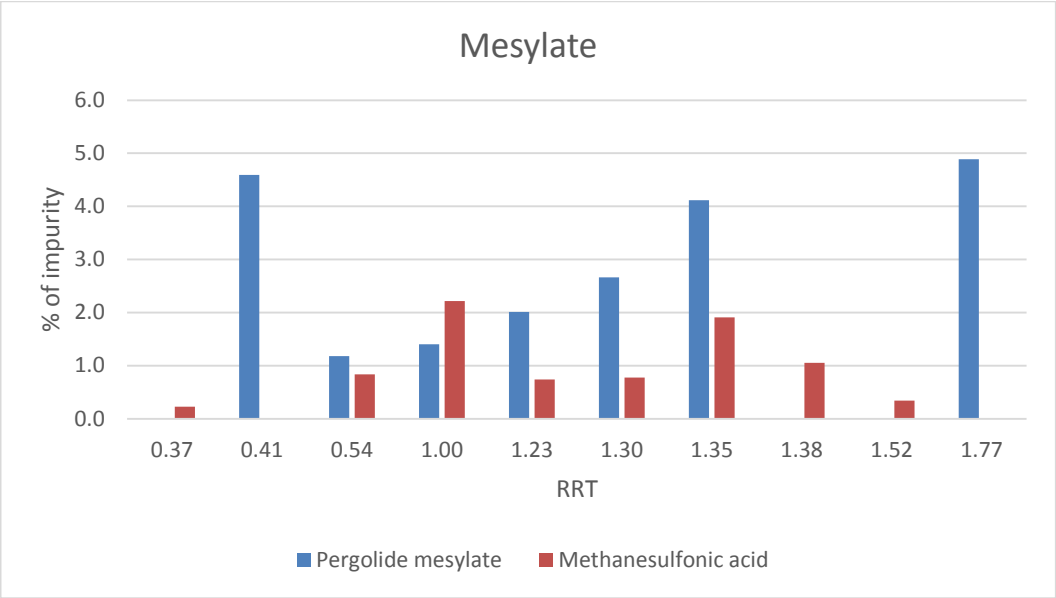


Figure 5.14 the percentage of mesylate and trace anionic impurities in samples containing mesylate counterion: Pergolide mesylate and MSA.
The relative retention time (RRT) presents on the x-axis and % of impurity on the y-axis.

5.4 Conclusion

This work provided an assessment of inorganic, and acidic and basic organic impurities of selected drug salt forms for several pharmaceutical drug substances which were commercially available. It was found that there were both qualitative and quantitative differences in anionic impurities in the different of drugs. It was also observed that the level of anionic impurities was greater in pure enantiomers - which have a more complex processing history - than in corresponding racemic mixtures. The results show that the acetate counterion is not displaced by anions of stronger acids since, in general, the levels of impurities were no higher than those in the chlorides. The lowest levels of impurities in samples containing acetate were observed in chlorhexidine which is a strongly basic drug. The major trace anionic impurity in mesylate containing samples was nitrate at least 4% and 2% in both pergolide mesylate and methane sulphonic acid (MSA) standard, respectively. This method was found to be robust enough for widespread use for fast determination and profiling of inorganic anions and organic acid impurities in APIs which might have an impact on large scale processing. In a related project, assistance was provided to another student who was examining variation in anionic impurities according to the crystallisation method used. It was observed that the crystallisation procedure had an impact on anionic impurities and also that trace anions affected the time required for induction of crystal formation (Natalia Dabrowska, PhD thesis, University of Strathclyde 2016).

Chapter 6

Determination of source of high 'Acid Clarity' values in 6-Aminopenicillanic acid (6-APA) powder

6 Determination of source of high 'Acid Clarity' values in 6-Aminopenicillanic acid (6-APA) powder

6.1 Introduction

6-aminopenicillanic acid (6-APA) is the penicillin core structure, and was first synthesised by removing the side (acyl) chain of penicillins isolated from a culture of *Penicillium chrysogenum*. 6-APA was produced by using Penicillin G acylase, a heterodimeric N-terminal nucleophile (Ntn) serine hydrolase, to induce an enzymatic deacylation of penicillin G (Shewale and Sivaraman, 1989, Wegman et al., 2001). This enzymatic reaction had low productivity related to the large reactions volumes of alkali (ammonia) and acid (sulphuric acid) used, and it generated an excessive amount of waste (Wang et al., 2006). Therefore, an efficient deacylation of penicillin G as a chemical synthesis was developed with the selective cleavage of the secondary amide bond. In the meantime, penicillin G acylase was obtained by screening of deacylase-producing microorganisms and extracting the enzymes from bacteria and actinomycetes that efficiently broke down the side-chain of penicillin G and penicillin V, respectively (Rolinson and Geddes, 2007). Recombinant DNA technology was employed to improve the production, stability and cost-efficiency. These enzymes were combined with efficient immobilization on a solid support which made the biocatalyst recyclable (Shewale and Sivaraman, 1989, Parmar et al., 2000).

The chemical synthesis of the 6-APA starts from the benzyl penicillin (Penicillin G) (Structure 1 in Figure 6.1) or Penicillin V by converting the carboxylic acid group to the corresponding silyl ester (Structure 2 in Figure 6.1) using trimethylsilyl chloride. This product is treated with phosphorus pentachloride in the presence of pyridine as base in results in the production of an imino chloride compound (Structure 3 in Figure 6.1). Solvolysis of the product with butanol produces the iminoether product (Structure 4 in Figure 6.1) and under the reaction conditions the silylether converts to the free acid. Hydrolysis to remove the imino ether function by using aqueous acid

affords the 6-APA (5) with free primary amine as seen as Structure 5 in Figure 6.1 (Weissenburger and Van Der Hoeven, 1970).

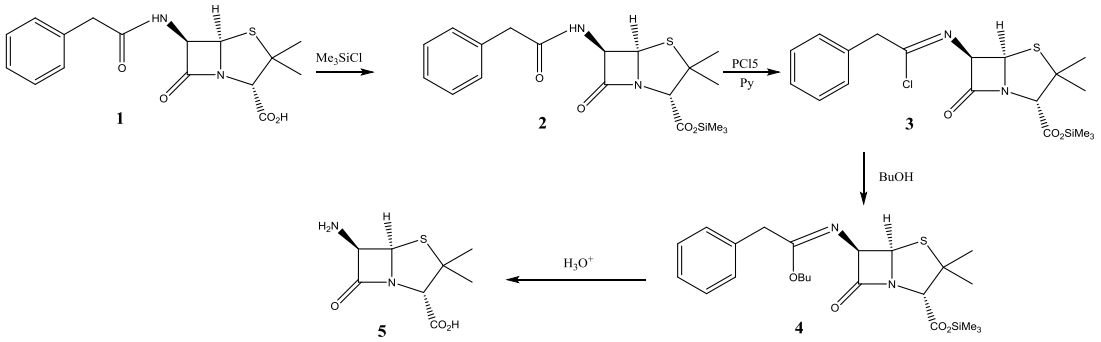


Figure 6.1 synthesis of 6-aminopenicillinanic acid (6-APA) (Weissenburger & Van Der Hoeven, 1970).

6-APA was confirmed as an intermediate in the synthesis of penicillin and this led to huge progress in the development and production of a large number of semisynthetic penicillins by acylation of 6-APA with many acid derivatives (Wegman et al., 2001).

Due to the high polarity and thermal instability of penicillins (Kantiani et al., 2009), HPLC techniques can provide a valuable tool for characterizing various antibiotics. Since many penicillins contain ionisable groups, the earliest separations were obtained using LC methods.

Reversed phase methods using bonded stationary phase materials have been coupled with HRMS. MS/MS has been used to thoroughly analyse the β-lactam ring structure (Canzani, Hsieh, Standland, Hammack, & Aldeek, 2017; Daeseleire, De Ruyck, & Van Renterghem, 2000; Granelli & Branzell, 2007; Heller & Ngoh, 1998; Pozo et al., 2006; Rabbolini, Verardo, Da Col, Gioacchini, & Traldi, 1998; Riediker & Stadler, 2001; Seifar et al., 2008) using mobile phases of water and acetonitrile both with and without the addition of 0.1% formic acid (De Baere & De Backer, 2007; Makeswaran, Patterson, & Points, 2005; Riediker & Stadler, 2001; Turnipseed, Andersen, Karbiwnyk, Madson, & Miller, 2008). HILIC can be used as a valuable alternative to the LC mode for separating small polar compounds (Jovanović et al., 2012; Liu et al., 2011) and it may be more appropriate for antibiotic analysis than using the traditional RP chromatography (Kahsay, Song, Van Schepdael, Cabooter, & Adams,

2014). ESI-MS is a common detection technique applied with HILIC separation for the many antibiotics (Kahsay et al., 2014; Schiesel, Lämmerhofer, & Lindner, 2010; Xu, Xie, Miller-Stein, & Woolf, 2009).

Ion chromatography (IC) is a well-developed technique for the separation of cationic and anionic ions and polar molecules based upon their charge (Joachim Weis & Weiss, 2004). IC coupled with conductivity detection has been widely used in pharmaceutical analysis (Cassidy, Demarest, Wright, & Zimmerman, 2004; DeBorba, Rohrer, & Bhattacharyya, 2004; Page, Stevenson, & Powell, 2014) due to it's high sensitivity and stability. Decreasing the background conductivity of the eluent leads to signal enhancement of the components of interest, thus offering more efficiency and robustness (Haddad, Jackson, & Shaw, 2003).

Gas chromatography (GC) is the most appropriate technique for volatile components and organic solvent detection. It can provide high level of resolution and selectivity and has the capacity to detect low solvent limits. Compound identification can be confirmed when GC is coupled with a mass spectrometer. The identity and quantity of residual solvents, or other organic volatile impurities, can be determined by using headspace (HS) sampling systems coupled with GC (B'Hymer, 2003; Grodowska & Parczewski, 2010).

This work describes analytical investigations of a series of samples of 6-APA which were provided by Glaxo Smith Kline (GSK) that had varying acid clarity values. During the manufacture of 6-APA, occasional batches are found to have high levels of acid clarity (i.e. high turbidity) leading to analytical/QC failure of the batch. The work provided a thorough investigation of the root cause of the high acid clarity values in the 6-APA materials.

6.2 Methodology

6.2.1 Samples

All samples were supplied by GSK (Irvine, UK) as a powdered form. The samples contained a range of acid clarity values and were made up fresh each time before analysis following the standard operating procedure (SOP)/analytical procedure supplied by GSK.

6.2.2 Samples preparation

The samples were diluted as required for the respective analysis.

For LC-MS analysis, the samples were prepared by dissolving 25 mg of the material in 50 mL of 1:1 water/methanol to a final concentration of 0.5 mg/mL.

For IC analysis, the samples containing 1mg of the 6-APA were diluted to a level of 20 µL/mL with deionized water.

For HS-GC-MS analysis, the samples 100mg of 6-APA were dissolved in water and 1 ml of solution into a 20 ml headspace vial

6.2.3 Hydrophilic interaction chromatography (HILIC) conditions

In the HILIC system, the stationary phase column was SeQuant ZIC-HILIC (3.5µm, 200Å) 150 × 4.6 mm PEEK metal-free HPLC column from Merck (Darmstadt, Germany). The guard column was SeQuant ZIC-HILIC Guard 20 × 2.1 mm from Merck (Darmstadt, Germany). The separation was carried out with solvent A (0.1% v/v formic acid in water) and B (0.1% v/v formic acid in acetonitrile). The gradient was started at 80% B at 0 min, then linearly decreased to 20% B with a flow rate of 0.3 ml/min for 30 min.

6.2.4 Reverses Phase liquid chromatography (RP-LC) conditions

In the RP-HPLC system, the reversed phase column used was ACE Super C18 (150 × 4.6 mm, 3 µm) from HiChrom (Reading, UK). The mobile phases were the same as for the HILIC method above. The gradient was started at 10% B at 0 min, then linearly increased to 100% B in 30 min and a flow rate of 0.3 ml/min. For MS detection the full mass range was 70 - 1200 m/z. Xcalibur software version 2.1.0 (Thermo Fisher Scientific, CA, USA) was employed for data acquisition and processing.

6.2.5 Liquid chromatography- Mass spectrometry

See section 2.3.

6.2.6 Liquid chromatography- MS^n mass spectrometry

The RP-HPLC system (see section 2.4) was used (with the column and gradients system as mentioned above). Protonated ($[M+H]^+$) and deprotonated ($[M-H]^-$) molecular ions were applied for detecting the elemental composition.

6.2.7 Ion chromatography

The anions and cations were analysed using capillary-HPIC system (see section 2.7). The method of calculation for anions and cations as seen in the section 5.2.3.

6.2.8 Headspace GC-MS

See section 2.8.

6.2.9 Microscopy analysis

The acid clarity solution was prepared according to GSK procedure by dissolving the samples in 1 molar hydrochloric acid (100 mg/mL) then stirring for 5 min. A 0.5 ml sample drop was placed on the microscope slide and investigated under laser scanning confocal microscopy at magnifications of 50x.

6.2.10 NMR

See section 2.9.

6.3 Analysis of Impurities and Correlation to High Acid Clarity Value

This section constitutes the main body of work within this project. A range of columns coupled with mass spectroscopy were used to identify the different impurities present within the 6-APA samples and to then correlate the level of the impurities with the high acid clarity samples using SIMCA-P 14 software and chemometric techniques. Three different chromatographic techniques were used to examine the anionic, cationic and neutral/lypophillic species present in the samples and the techniques are described in more detail in the following sections.

6.3.1 Anionic Impurity Analysis by Ion chromatography

A capillary ion chromatography system coupled with suppressed conductivity detection was used to characterize different batches of 6-APA for the profiling of small acidic and inorganic anionic impurities. The relative retention times (RRTs) of known anion peaks were used to identify unknown impurities. The conductivity levels of each impurity in the sample were determined for the unknown and knowns impurities. Figure 6.2 shows the ion chromatography of the 16 6-APA samples.

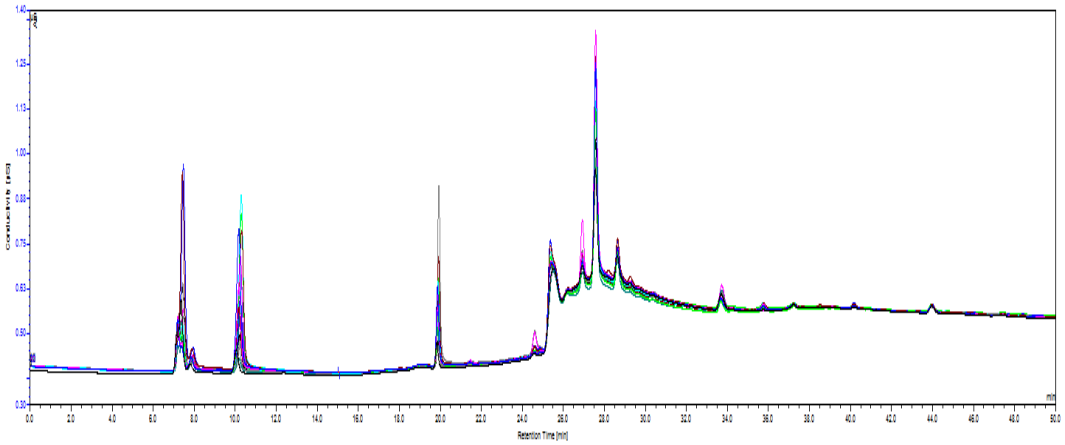


Figure 6.2 Overlapped anionic ion chromatograms of the 16 samples of 6-APA were obtained by Reagent-Free Capillary HPIC system of an ICS-5000+ instrument with a potassium hydroxide electrolytic eluent generation module. The retention time (0 to 50 min) is shown on the x-axis and the conductivity (µS) level is shown on the y-axis.

The level of anions associated with different levels of acid clarity of the samples can be seen below (Figure 6.3), and shows that over half of the anions were identified as commonly known anions with chloride, acetate, formate/butyrate, nitrate, sulphate and phosphate being the most common ions. The major anionic impurities found are related to the unknown impurity at RRT 1.28, formate/butyrate and sulphate with percentages of 23.66, 19.17 and 16.35 %, respectively (Figure 6.4).

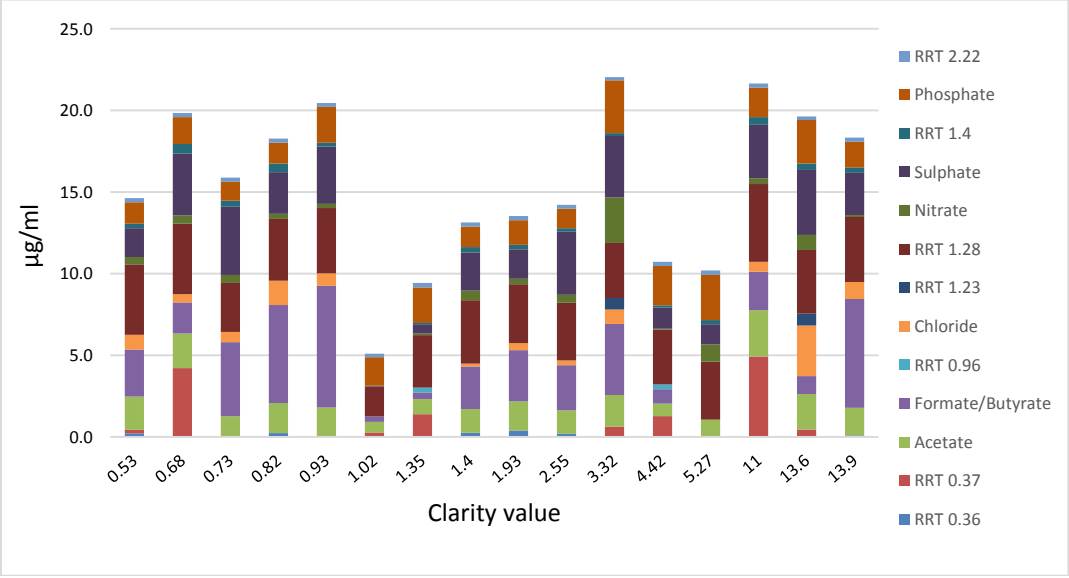


Figure 6.3 Ion Chromatographic analysis for anionic contaminants of the 6-APA samples with identification of the known anions shows the level of anions in each sample. The clarity values are shown on the x-axis and the amount of anions (µg/ml) is shown on the y-axis.

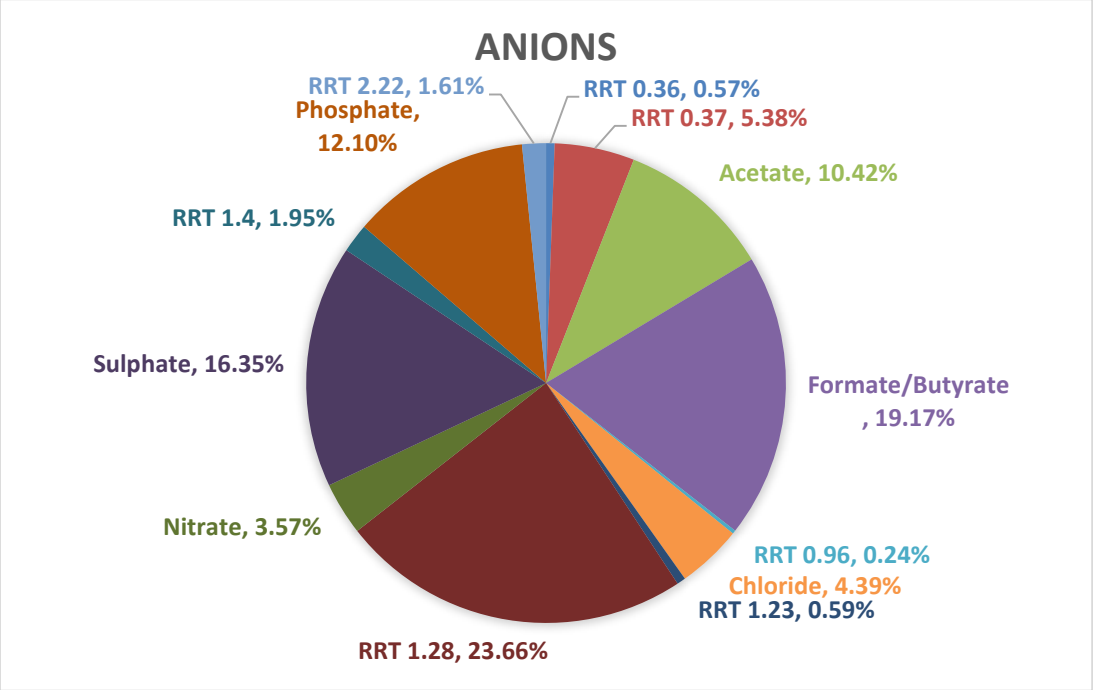


Figure 6.4 Pie plot shows the average percentage of organic acid and inorganic anions in the acid clarity samples of 6-APA.

This anionic impurity profile was used to generate a PCA model to investigate the correlations between anionic impurities and the acid clarity value of the samples. As shown in the figure 6.5, the PCA biplot shows the distribution of the acid clarity values over the anions. The scores (i.e. samples) were randomly spread on the model and the loadings (i.e. anions) show that even though some are significantly raised in some samples, the increase still does not show any relation to the clarity value of the samples. HCA was used as method of cluster analysis to create a binary tree of the data that successively merges similar groups of samples (Figure 6.6). As a result, the biplots and HCA plot show that acid clarity value of the samples has no statistically relevant correlation to the distribution of the anionic impurities. In addition, the PCA model shows very weak cross validation with $R^2X(cum)=0.57$ and $Q^2(cum)=0.065$ which shows weak productivity and reproducibility.

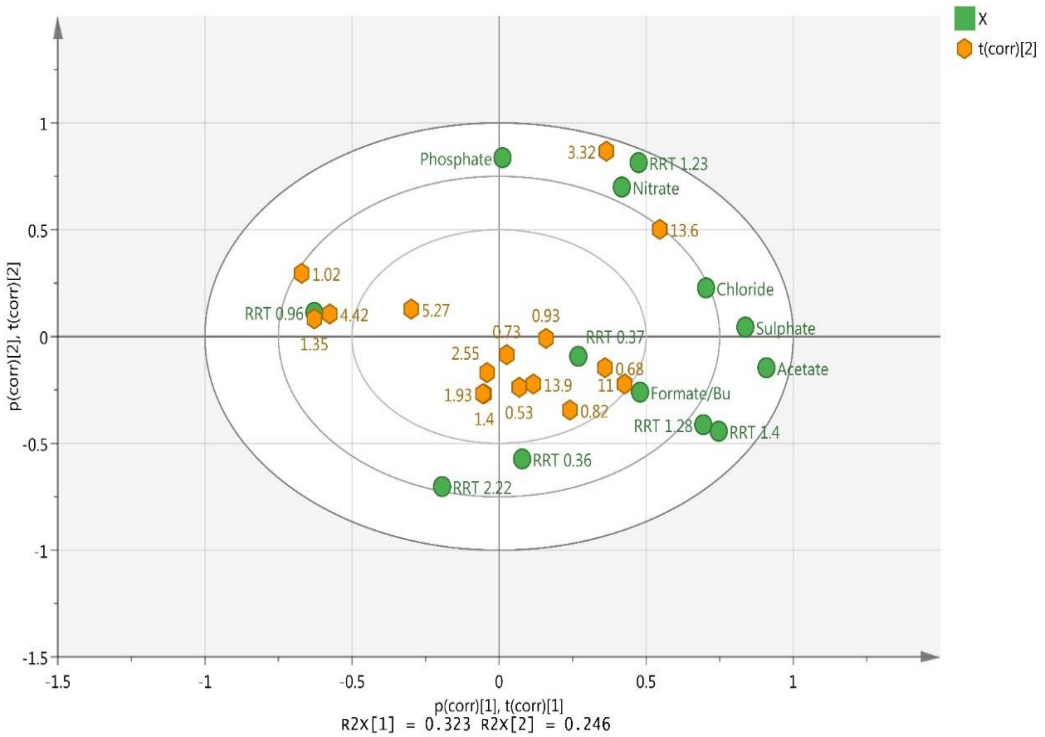


Figure 6.5 Biplot of anions ion chromatography analysis: scores (samples) and loadings (anions) with low discrimination power on pq(corr)[1] t(corr)[1] vs. pq(corr)[2] t(corr)[2].

The plot shows distribution of 16 observations plus acid clarity values(orange) and 13 variables (green). It consists of one predictive x-score component; component t(corr)[1] and one orthogonal x-score components t(corr)[2] and x-variables component; component pq(corr)[1] and one orthogonal x-variables components pq(corr)[2].

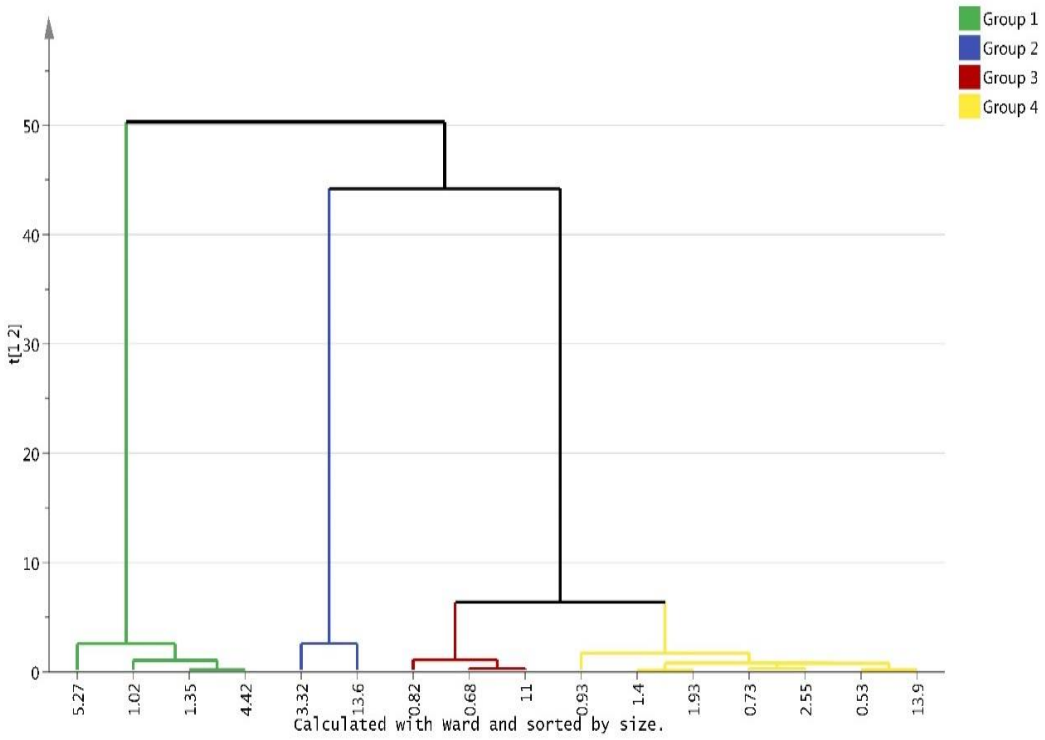


Figure 6.6 Hierarchical Clustering Analysis (HCA) plot for the 16 samples of 6-APA from PCA model of anions analysis.
The dendrogram shows observations clustered into four hierarchical groups; (green; 5.27, 1.02, 1.35 and 4.42), (blue; 3.32 and 13.6), (red; 0.82, 0.68 and 11) and (yellow; 0.93, 1.4, 1.93, 0.73, 2.55, 0.53 and 13.9) by using Ward clustering method. X-axis represents the samples and y-axis shows the variability index. The values of the variability index among the samples. The larger values show big differences between group variability and the lower values show more similarity.

The most reasonable conclusion of these results is that anionic impurities are not related to the clarity issue observed for the 6-APA samples. As no correlation was found, no additional time was spent in identifying the unknown anions.

6.3.2 Cationic impurity Analysis by Ion chromatography

A capillary ion chromatography system coupled with suppressed conductivity detection was used to characterize different batches of 6-APA for profiling of cationic impurities. Figure 6.7 shows the cationic analysis by ion chromatography. The RTs of given cation peaks were applied to identify the impurities. The conductivity signals of the impurity peaks were used to measure the level of the unknown and known impurities.

The analysis investigated specifically cationic impurities and used PCA to look for correlations between cationic impurities and the level of clarity of the samples. Many of the cations were identified as commonly known cations with sodium, ammonium, potassium and calcium being the most common. As can be seen below (Figure 6.8), the levels of cations vary among the samples but there does not seem to be any correlation between the sample clarity values and the distribution of the cations. Although, the major cationic impurities found are related to the ammonium and sodium with percentages of 46 and 24%, respectively (Figure 6.9).

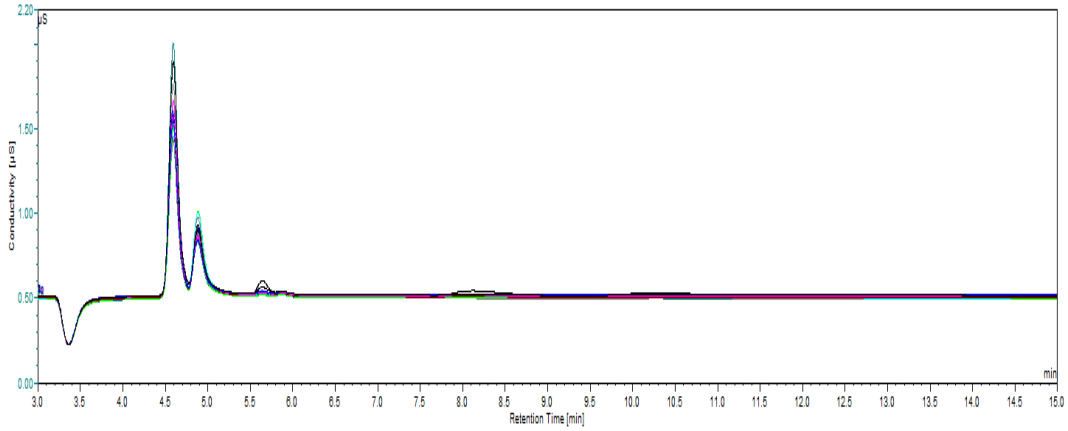


Figure 6.7 Overlapped cationic ion chromatograms of the 16 samples of 6-APA were obtained by Reagent-Free Capillary HPIC system of an ICS-5000+ instrument with methane sulphonic acid electrolytic eluent generation module. The retention time (0 to 20 min) is shown on the x-axis and the conductivity (µS) levels are shown on the y-axis.

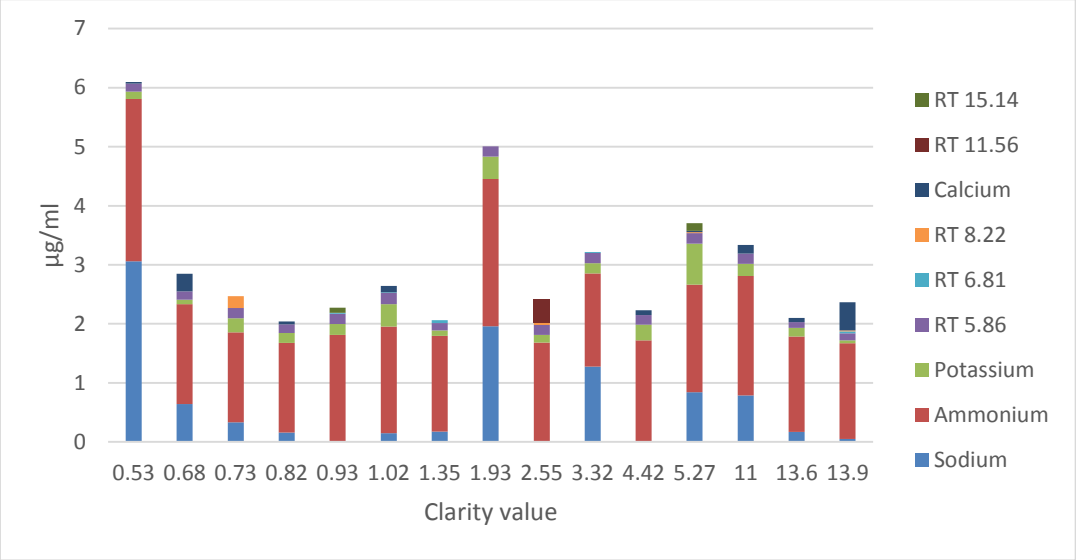


Figure 6.8 Ion Chromatography analysis for cations of 6-APA samples
with identification of known cations shows the level of cations in each sample. The clarity values are shown on the x-axis and the amount of the cations (µg/ml) is shown on the y-axis.

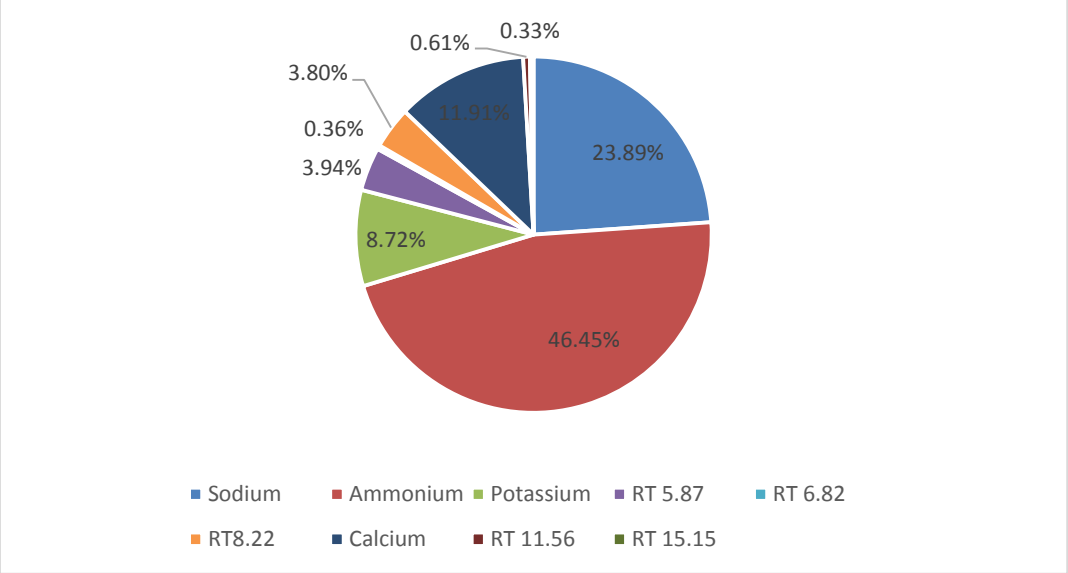


Figure 6.9 Pie plot shows the percentage of average cations in the acid clarity samples of 6-APA.

These results were then analysed using PCA to determine if there was any correlation between cationic impurities and the clarity value of the samples. Figure 6.10 shows the PCA biplot for cations spreading over the clarity values for the samples. The distribution of the observations (i.e. samples) has no relation to their clarity values. Figure 6.11 shows the results from this analysis from HCA which shows the very low

level of variability index (y-axis) which means that the minor differences among the samples are related to their cation levels.

The sample with clarity value 1.4 is likely considered serious outlier observation since it larger than the 99% confidence limit. The hotelling's T2 plot is seen in (Figure S6.1, appendix)

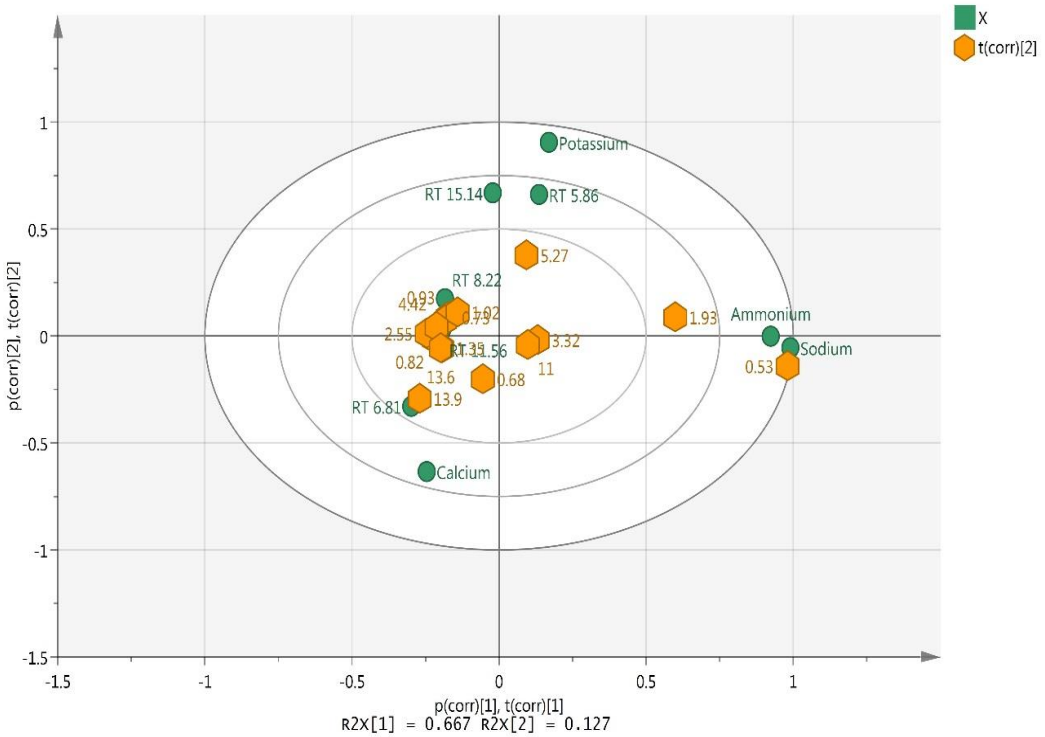


Figure 6.10 Biplot of cation ion chromatography analysis: scores (samples) and loadings (cations) with low discrimination power on pq(corr)[1] t(corr)[1] vs. pq(corr)[2] t(corr)[2].
The plot shows distribution of 15 observations plus acid clarity values(orange) from four samples and 9 variables (green). It consists of one predictive x-score component; component t(corr)[1] and one orthogonal x-score components t(corr)[2] and x-variables component; component pq(corr)[1] and one orthogonal x-variables components pq(corr)[2].

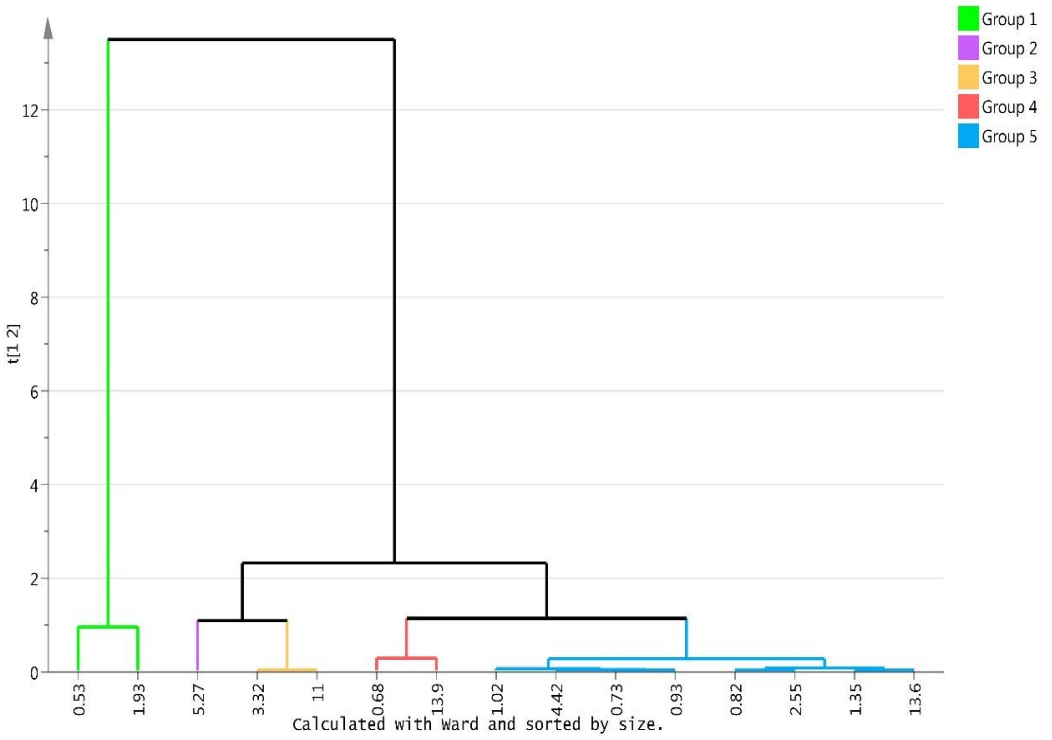


Figure 6.11 Hierarchical Clustering Analysis (HCA) plot for the 15 samples of 6-APA from PCA model of cations analysis.
The dendrogram shows observations clustered into five hierarchical groups; (green; 0.53 and 1.93), (purple; 5.27), (orange; 3.32 and 11), (red; 0.68 and 13.9) and (blue; 1.02, 4.42, 0.73, 0.93, 0.82, 2.55, 1.35 and 13.6) by using Ward clustering method. X-axis represents the samples and y-axis shows the variability index. The values of the variability index among the samples. The larger values show big differences between group variability and the lower values show more similarity.

These results show that there is no statistically relevant correlation between cationic impurity levels and clarity values. The most reasonable conclusion of these results is that cationic impurities are not directly related to the clarity issue observed for the 6-APA samples. As no correlation was found, no additional time was spent identifying the unknown cations.

6.3.3 Volatile Organic Compounds Analysis

During the batch development, the plant facility recognised that the original analysis would be likely to miss several volatile organic compounds, especially low boiling point (bpt) alcohols and the levels of some of these was known to vary; particularly butanol. Therefore, headspace GC-MS analysis was carried out to investigate the change in the butanol level in the samples.

The main peak identified by MS was butanol, as was expected. The butanol levels (as peak area) were plotted vs the acid clarity values for the samples. As can be seen in figure 6.12, there is no correlation between acid clarity value and butanol levels.

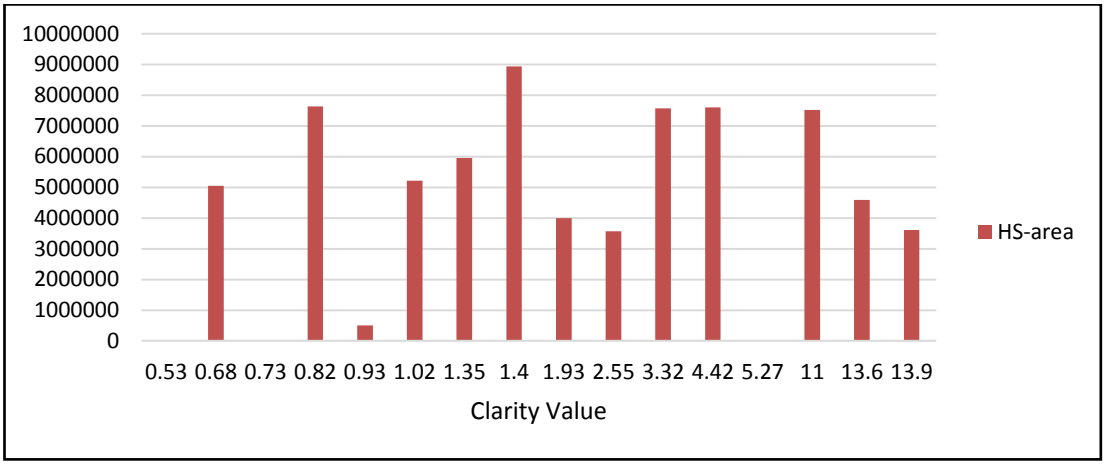


Figure 6.12 Level of butanol in 6-APA samples plotted as peak area butanol vs acid clarity value. X-axis represents the clarity value of the samples and y-axis shows the peak area.

The results show that while the majority of the samples contain residual butanol to some extent, there is no correlation between levels of residual butanol and clarity of the samples.

6.3.4 Neutral/Organic Impurity Analysis

HILIC chromatography hyphenated to a high resolution mass spectrometer was used for the purpose of obtaining a comprehensive characterisation of the organic impurity profile. The chromatograms for HILIC system were generated using the method conditions described in section 6.2.3.

The LC/MS data were captured in the Xcalibur data format. The data acquired from Xcalibur software was used to obtain the empirical formulae. The data of exact masses of the molecular ions with error margin of less than 3ppm combined with the information from nitrogen rule, double bonds (RDB), and isotopic patterns, were used for interpretation of the data. In addition, adduct formation was investigated

for the purpose of confirming apparent impurities arising from the molecular ion in some cases.

MZmine 2.20 software was used to process the data from LC-HRMS. Many steps were used to detect the peaks including: peak detection step, mass detection, chromatogram deconvolution, deisotoping, normalisation, alignment and gap filling. The extracted components were modelled by using SIMCA-P software. The total ion current (TIC) of HILIC chromatogram profile for each sample is shown in (Figure 6.13). Both the positive and negative mode data were analysed.

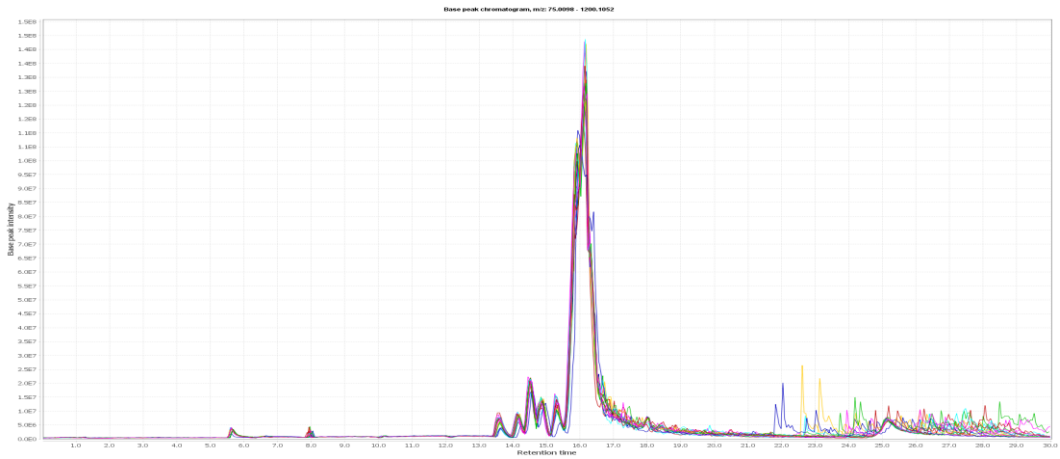


Figure 6.13 Overlapped total ion current chromatograms (TICs) of HILIC analysis of 6-APA samples characteristics by MZmine 2.23. X-axis represents the retention time (0 -30 min) and y-axis shows the base peak intensity**.**

6.3.4.1 Negative ion mode mass spectrometry

The negative ion mode determines the deprotonated organic species which are ionised by the negative ion mode used in the mass spectrometer, rather than only looking for negative ion impurities.

A PCA model was generated to investigate the correlations between impurities from negative ion mode and acid clarity value of the samples. The total PCs show cross validation ($R^2X(cum)$= 0.963, $Q^2(cum)$= 0.772) which are considered an acceptable level of model fit to the data with good reproducibility. In this case the grouping in the PCA plot shows a low likelihood of finding any statistically significant correlations

to the high clarity reading samples. The PCA scores and loadings plots (Figure 6.14) and HCA (Figure 6.15) for negative mode HILIC-MS data show that the samples have not logically grouped in accordance with their clarity value.

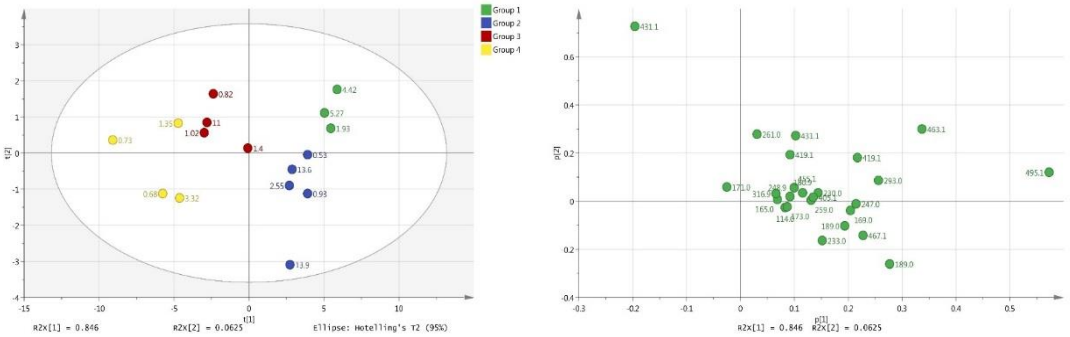


Figure 6.14 PCA scores scatter plot (t1 vs. t2) (left) loadings plot (p1 vs. p2) (right) of PCA analysis of organic impurities of the clarity samples by negative ionisation mode HILIC-MS.
PC1 and PC2 obtained by using Pareto scale. Three PCs described 96.3% of the whole variation data (PC1: 84.6%, PC2: 6.25% and PC3: 5.45%). Scores plot shows the discrimination between the samples according to the impurity profile of each sample. Loading plot points out the impurities that govern the position of each sample in the scatter plot.

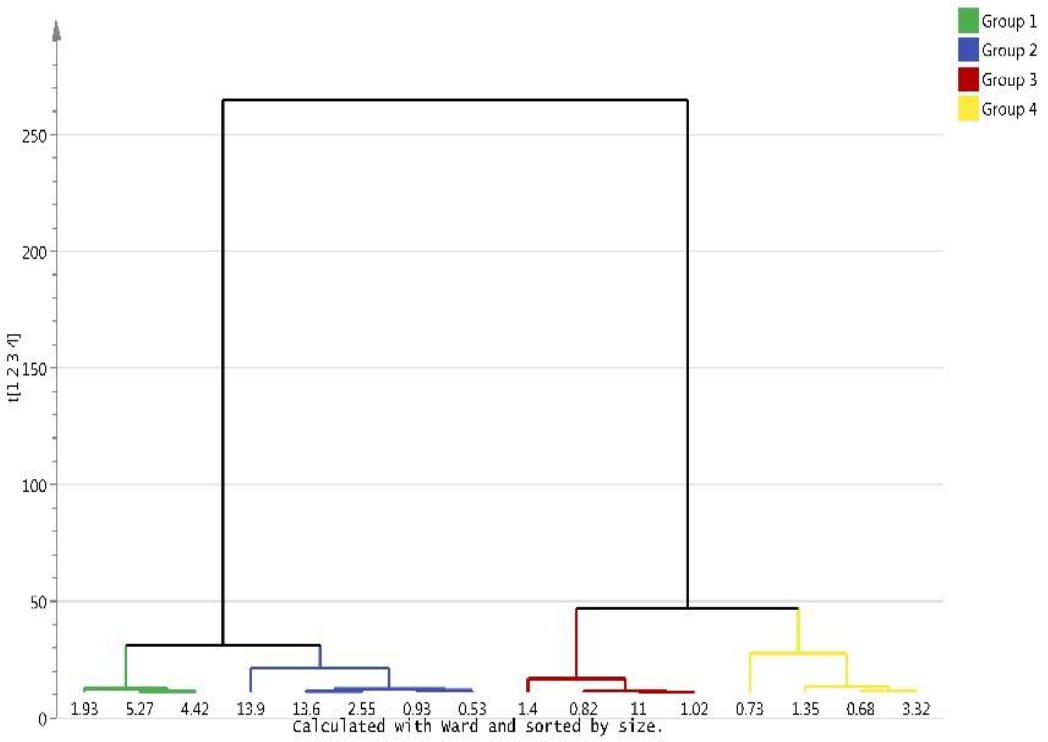


Figure 6.15 Hierarchical Clustering Analysis (HCA) plot for the 16 samples of 6-APA from PCA model of negative ionisation mode HILIC-MS.
The dendrogram shows observations clustered into four hierarchical groups; (green; 1.93, 5.27 and 4.42), (blue; 13.9, 13.6, 2.55, 0.93 and 0.53), (red; 1.4, 0.82, 11 and 1.02) and (yellow; 0.73, 1.35, 0.68 and 3.32) by using the Ward clustering method. X-axis represents the samples and y-axis shows the variability index. The values of the variability index among the samples. The larger values show big differences between group variability and the lower values show more similarity.

The negative ions formed in the mass spectral data were not investigated in further detail due to the low likelihood of finding any statistically significant correlations to impurities.

The percentage of the major impurities (those >0.5%) from the analysis of organic impurities of the clarity samples by negative ionisation mode HILIC-MS was calculated based on the level of both 6-APA and the ion at m/z 247 which is a methanolysis product of 6-APA (Figure 6.16). The major impurity and adduct peaks were $[M-H]^-$ m/z 431 (6-APA dimer), m/z 495 (dimer of m/z 247), m/z 463 (complex of m/z 247 and 6-APA), m/z 189 (D-4-carboxy-5,5-dimethyl-2-aminomethylthiazolidine (CDAT)) and m/z 261 (ethanolysis of 6-APA) whose proposed structure is discussed in more detail in section 3.2.

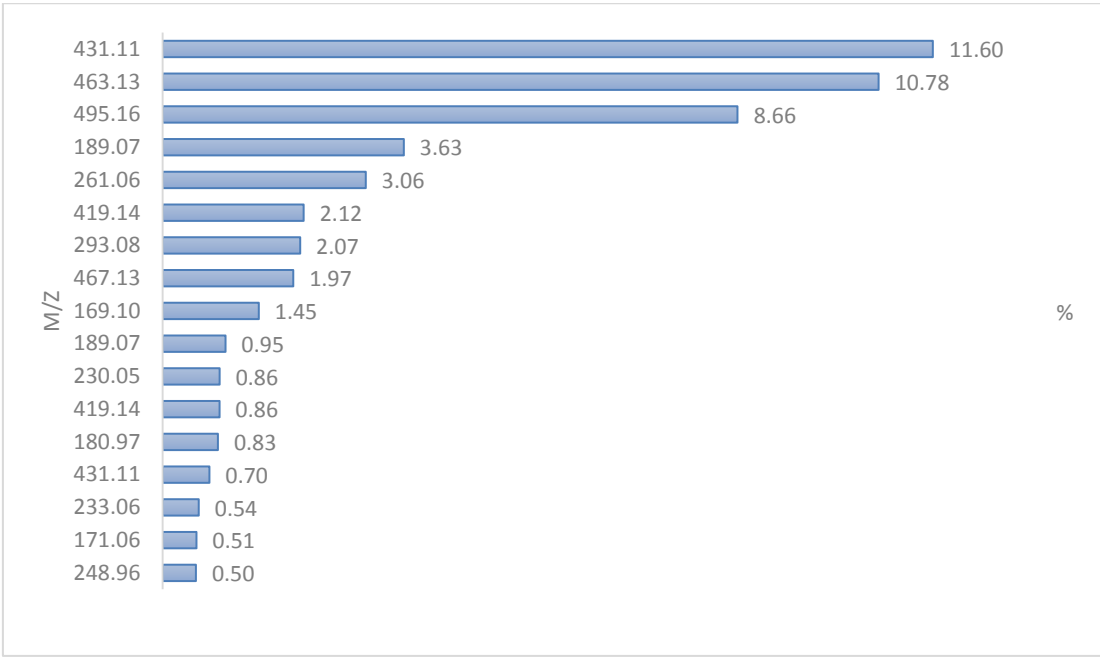


Figure 6.16 the percentage of the major Impurities (>0.5%) from the analysis of organic impurities of the clarity samples by negative ionisation mode HILIC-MS.
The x-axis is m/z of the major peaks and y-axis is level of impurity (relative amounts % related to the main peaks 6-APA + m/z 247). The value of (% of impurity) was calculated by taking the mean of the peaks in all samples.

6.3.4.2 Positive ion mode mass spectrometry

Positive ion mode refers to the protonated organic species which were detected in the mass spectrometer in positive ion mode. A PCA model was generated to discriminate the different profiles between the samples. The discrimination is dependent on the differentiation by the impurities formed in the samples. The total PCs with cross validation ($R^2X(cum)$= 0.98, $Q^2(cum)$= 0.864) showed a good level of model fit to the data with good reproducibility.

A scores plot is used to display the pattern of observations (samples) due to the differences and similarities expressed by their variables (impurities). It was obtained by PCA and shows the clusters of the samples (Figure 6.17). The HCA plot displays some clustering according to the different clarity values of the samples in addition the loadings plot which indicates the scattering of the impurities related to their influence in the samples (Figure 6.18).

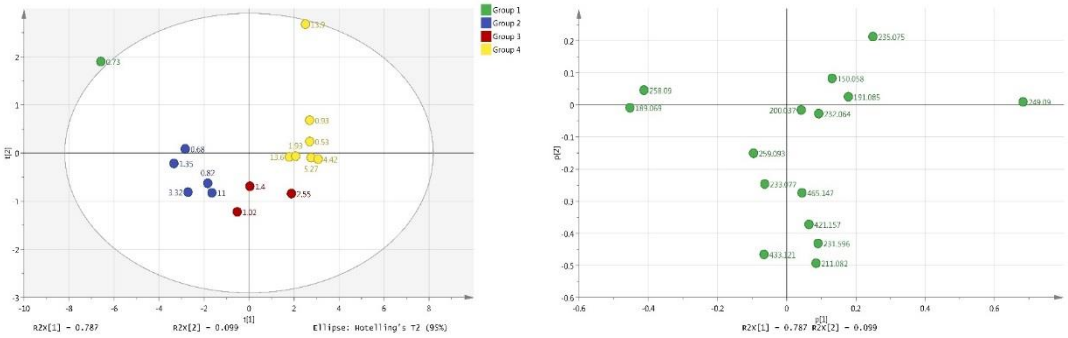


Figure 6.17 PCA scores scatter plot (t1 vs. t2) (left) loadings plot (p1 vs. p2) (right) of PCA analysis of organic impurities of the clarity samples by positive ionisation mode HILIC-MS.

PC1 and PC2 obtained by using Pareto scale. Three PCs described 98% of the whole variation data (PC1: 78.7%, PC2: 9.9% and PC3: 9.4%). Scores plot shows the discrimination between the samples according to the impurity profile of each sample. Loading plot points out the impurities that govern the position of each sample in the scatter plot.

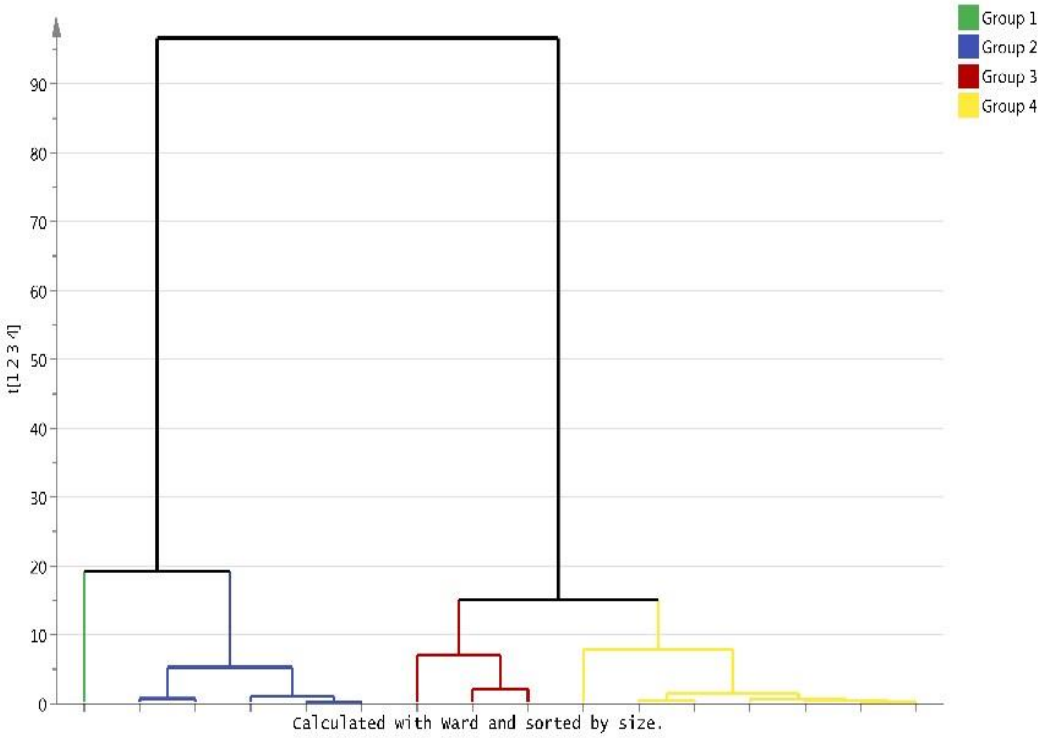


Figure 6.18 Hierarchical Clustering Analysis (HCA) plot for the 16 samples of 6-APA from PCA model of positive ionisation mode HILIC-MS.

The dendrogram shows observations clustered into four hierarchical groups; (green; 0.73), (blue; 0.68, 1.35, 0.82, 11, 3.32 and 3.32), (red; 1.4, 2.55 and 1.02) and (yellow; 13.9, 0.93, 0.53, 1.93, 13.6, 4.42 and 5.27) by using Ward clustering method. X-axis represents the samples and y-axis shows the variability index. The values of the variability index among the samples. The larger values show big differences between group variability and the lower values show more similarity.

Based on the data shown in both the PCA and HCA plots in the upper figures 6.17 and 6.18, the clustering of the samples according to their impurities shows poor

discrimination related to clarity values. It revealed that organic impurities detected in positive ion mode have no significant correlation to the level of clarity of the sample. This is similar finding from the data obtained in the negative ion mode ESI-MS in the previous section.

Figure 6.19 shows the major impurities (>0.5%) were related to the main peaks 6-APA and ion at m/z 249 (methanolysis of 6-APA). Impurities at m/z 235 (Penicic acid), m/z 433 (6-APA dimer), m/z 191 (CDAT) and m/z 189 with empirical formula ($C_7H_{13}O_2N_2S$) and m/z 258 whose proposed structures are discussed in more detail in the next sections.

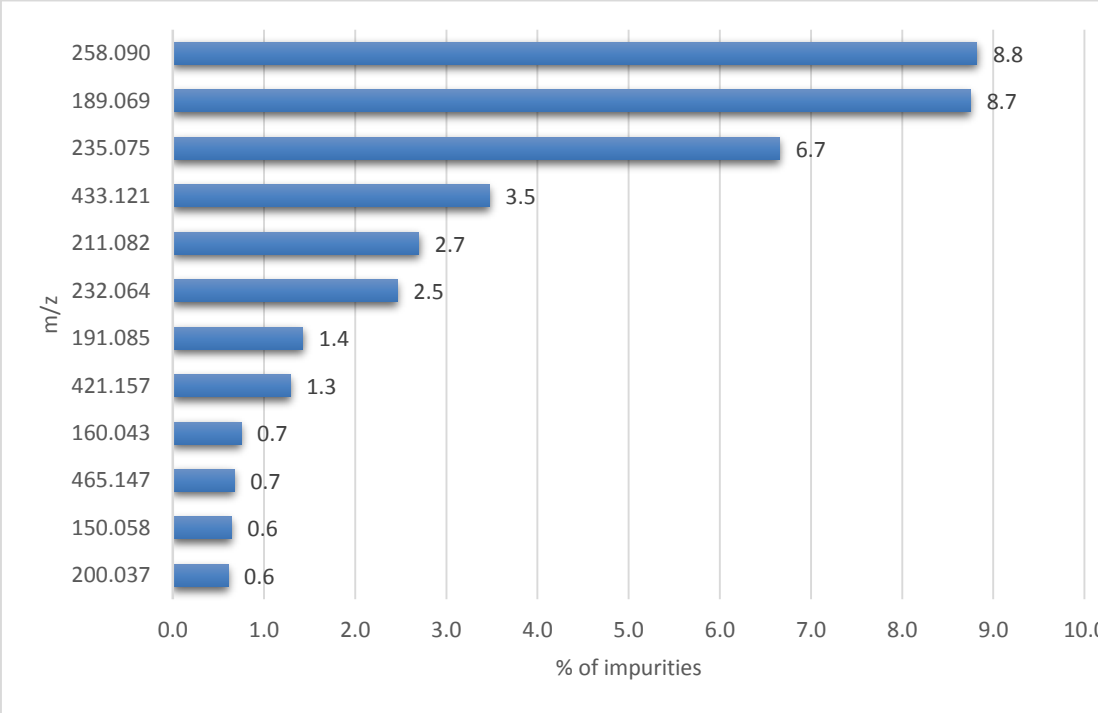


Figure 6.19 the percentage of Impurity from the analysis of organic impurities of the clarity samples by positive mode HPLC-MS.
The x-axis is m/z of the major peaks and y-axis is level of impurity (relative amounts % related to the main peaks 6-APA + m/z 249). The value of (% of impurity) was calculated by taking the mean of the peaks in all samples.

A statistical analysis of all this data showed that no statistically valid correlation between impurity levels and high acid clarity value could be found. A strong fit has

been identified in the model, but if the model does have a predictive power it is based on something subtler than relative level of impurities.

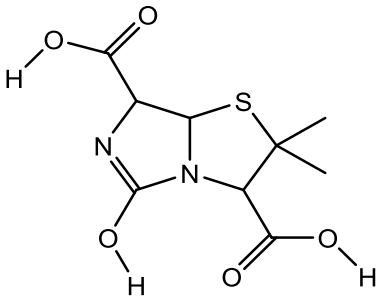
The high acid clarity value certainly does not correlate to a single impurity, and from the data analysis, within experimental and analytical error, we are unable to identify a specific group of impurities whose levels rise or fall in correlation with acid clarity value.

The next section of this report will discuss and propose structures for the most common impurities that have been identified.

6.4 Identification of the Impurities

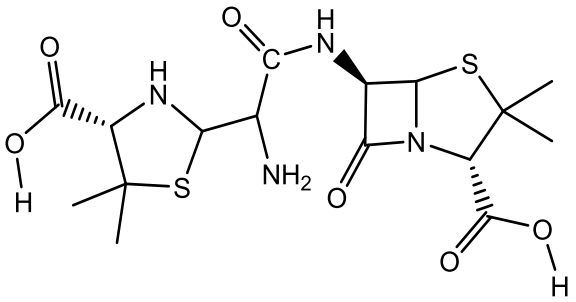
6.4.1 Known Organic Impurities and Degradants

Some of the molecules that were detected to some extent in the HILIC-MS analysis are impurities known to GSK during the 6-APA manufacturer process. The impurity with a molecular weight of 260.26 Da and the chemical formula ($C_9H_{12}N_2O_5S$) is related to 8-Hydroxypenillic acid (8-HPA). The impurity with a molecular weight 432.51 Da with the chemical formula ($C_{16}H_{24}N_4O_6S$) is related to 6-aminopenicillanic acid dimer. The impurity with a molecular weight 234.27 Da with the chemical formula ($C_8H_{14}N_2O_4S$) is related to penicic acid (the hydrolysis product of 6-APA). The impurity with m/z 248.08 Da with the chemical formula ($C_9H_{16}N_2O_4S$) is related to methanolysis product of 6-APA. The impurity with a molecular weight 190.08 Da with molecular formula $C_7H_{14}N_2O_2S$ is assigned to CDAT and the impurity with a molecular weight 233.08 Da with chemical formula ($C_8H_{15}N_3O_3S$) is related to the ammonia ring opened product. The structures of known impurities are shown in the figure 6.20.



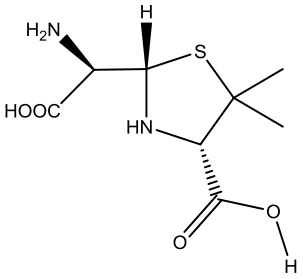
A

8-HPA



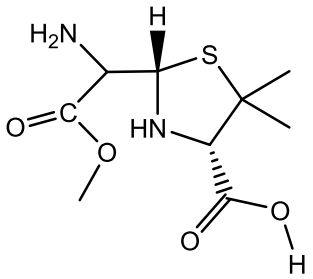
B

6-aminopenicillinc acid dimer



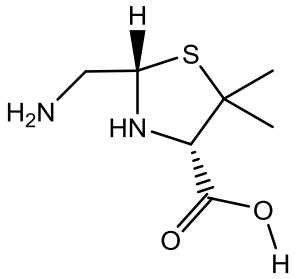
C

Penicic acid (Hydrolysis product)



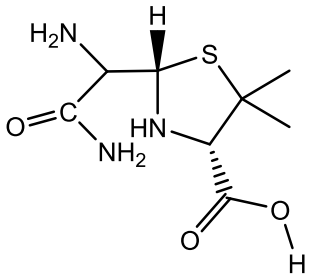
D

Methanolysis product



E

CDAT



F

Ammonia ring opened product

Figure 6.20 the structure of the known impurities, (A) 6-hydroxypenicillic acid, (B) 6-aminopenicillinc acid dimer, (C) Penicic acid (Hydrolysis product), (D) Methanolysis product, (E) CDAT, and (F) Ammonia ring opened product.

6.4.2 Structure Elucidation of Postulated Unknown Organic Impurities and Degradants

The main unknown impurities seem to be best fitted against compounds where ring opening via alcoholysis or hydrolysis has occurred, although other impurities could also fit the mass spectra.

A range of the main impurities and postulated impurities were selected based on data from the protonated ESI-LC/MS analysis, and were investigated to propose possible structures using MS^2 analysis. LC-MS/MS was generated using the method conditions described in section 2.4. MS^2 spectra were used for detecting each molecular ion fragment to provide valuable information used in structural elucidation.

The HRMS data proposed that the impurity at ($[MH]^+$, m/z 189.0692) had an elemental composition of $C_7H_{13}N_2O_2S$. The EIC trace for this impurity is shown in figure 6.21. The precursor ion showed fragment peaks at m/z 172.042 and m/z 160.042 with chemical formulae of ($C_7H_{10}NO_2S$) and ($C_6H_{10}NO_2S$) related to the loss of amine and CH_3N moiety, respectively. Other product ions at m/z 143.063, with the elemental composition $C_6H_{11}N_2S$, showed loss of a carboxylic acid group and m/z 128.053 with the chemical formula of ($C_6H_{10}NS$) is related to additional loss of the amino group. The fragment ion at m/z 69.044 with the chemical formula ($C_3H_5N_2$) is a thiazolidine ring opened product.

Figure 6.22 presents the proposed structure and fragmentation pathway of the impurity at m/z 189 which is related to the loss of carbonyl (C=O) group of the β - lactam ring.

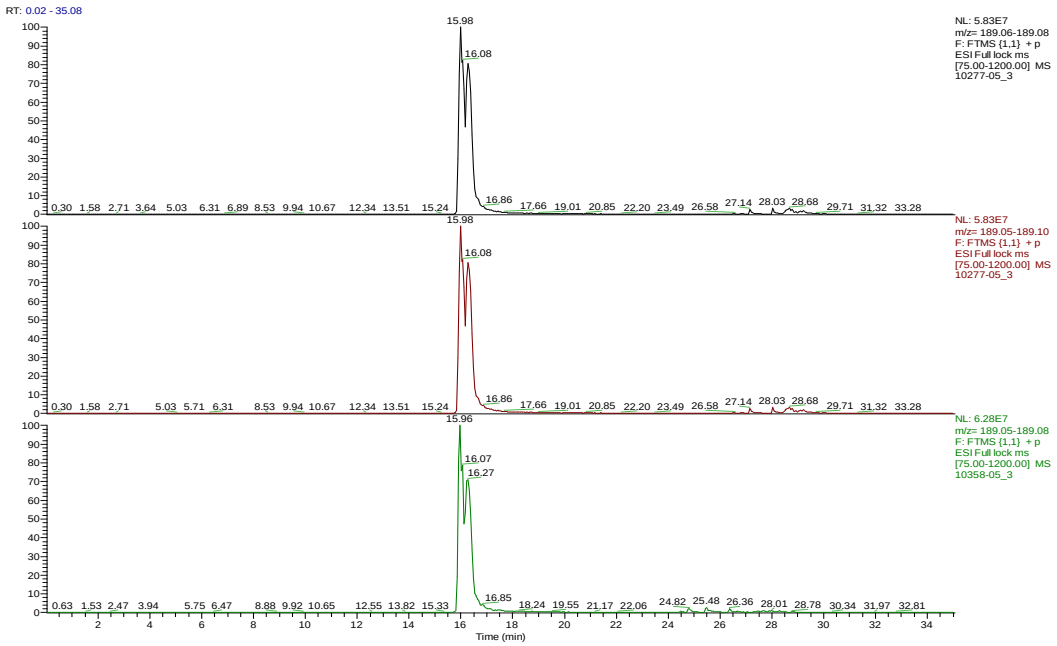


Figure 6.21 EICs for impurity at ([MH]+, m/z 189.0692) in samples with clarity index at 13.6, 5.27 and 0.53.

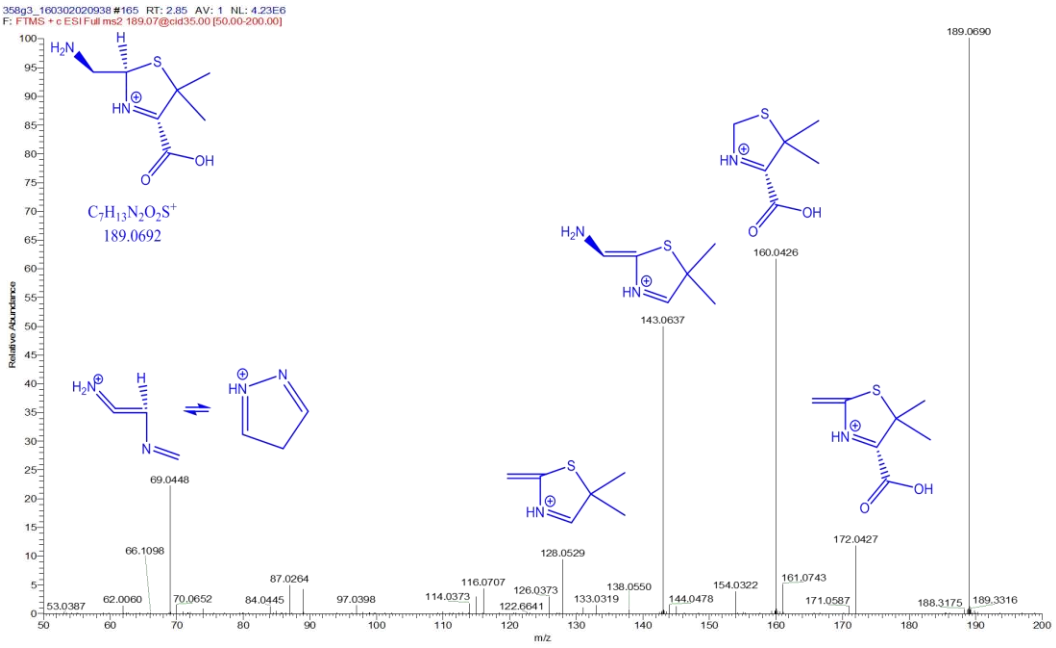


Figure 6.22 MS^2 spectra and proposed structure of impurity at ([M+H]$^+$, m/z 189.0690) and its fragments.

The protonated molecular ion at m/z 235.0747 with the predicted molecular formula $C_8H_{15}N_2O_4S$ is penicic acid which is a hydrolysed form of 6-APA. The product ions at m/z 218.048 and m/z 174.05 have the proposed chemical formulae of $C_8H_{15}NO_4S$ and $C_7H_{12}NO_2S$ and are related to deamination for the former, and deamination and decarboxylation for the latter compound, from the precursor ion. The MS^2 data provided fragmentation data for the product ion at m/z 191.084, which is related to the loss of the carboxylic acid group and has the elemental composition ($C_7H_{15}N_2O_2S$). Fragment ions at m/z 160.042 and m/z 146.063 are related to further loss of CH_5N and CH_3NO moieties from m/z 191.084, respectively. Another fragment at m/z 128.053 is related to decarboxylation of the most abundant fragment at m/z 174.058. Figure 6.23 shows the EIC trace for penicic acid impurity. The proposed structure and fragmentation pathway of the impurity m/z 235.0747 (penicic acid) are shown in the figure 6.24.

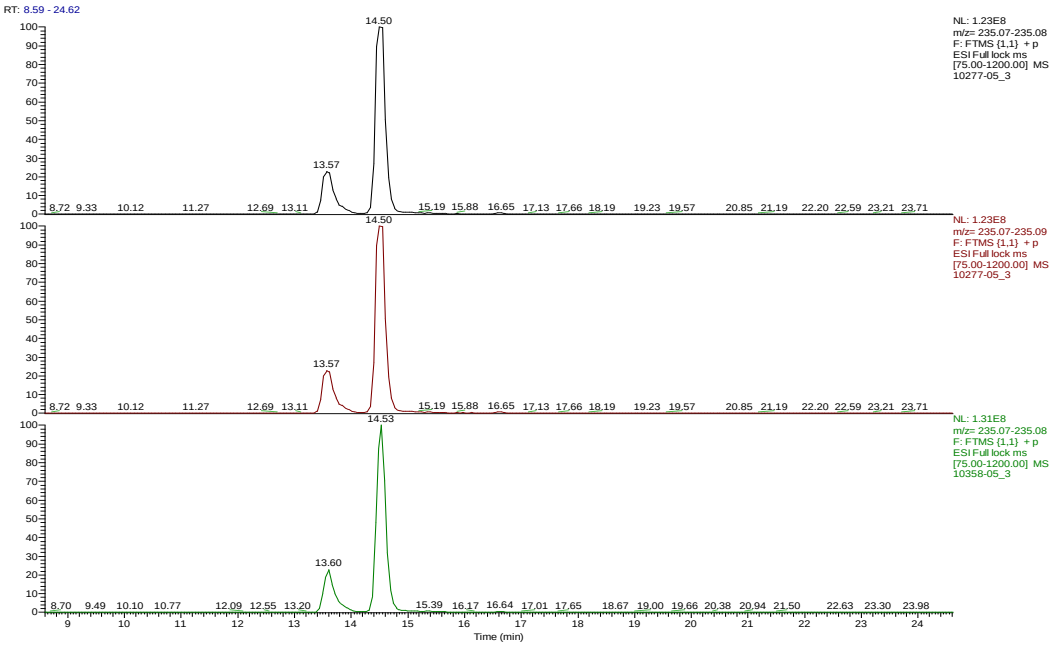


Figure 6.23 EICs for impurity at ([MH]$^+$, m/z 235.0747) in samples with clarity index at 13.6,5.27 and 0.53.

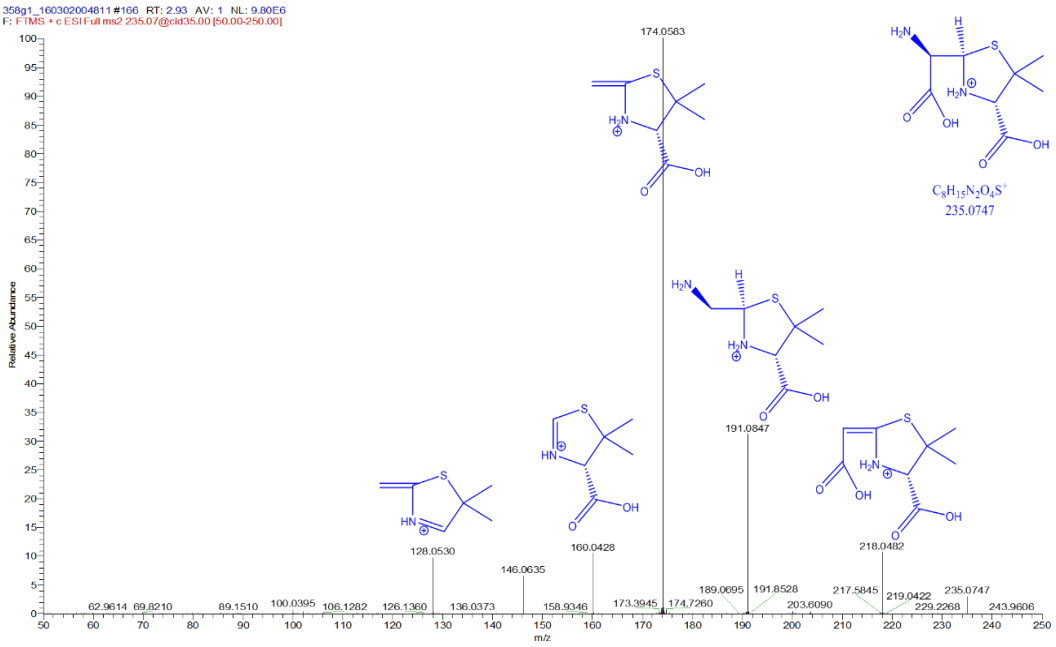


Figure 6.24 MS^2 spectra and proposed structure of impurity at ([M+H]$^+$, m/z 235.0747) and its fragments.

The HR-MS/MS data revealed that impurity m/z 249.0904 has an elemental composition of $C_9H_{17}N_2O_4S$ with an extra methanol group on the 6-APA. The other impurity at m/z 232.0638 of 6-APA is also shown as a fragment ion of the impurity at

m/z 249. It is correlated to the chemical formula ($C_9H_{14}NO_4S$) which is linked to loss of amino group from the impurity at m/z 249.090. The fragment peak at m/z 200.037 is shown in both MS^2 spectra of the impurities displayed by the loss of a methanol group from the ion at m/z 232.063. The other fragments produced from the impurity at m/z 232.063 are m/z 218.047 and m/z 186.058 which are the loss of methyl and carbonyl groups from the precursor ion (i.e. m/z at 232.063), respectively. The fragment ion at m/z 172.042 is related to further loss of methylene from the ion at m/z 186.058.

Figures 6.25 and 6.26 show the EIC trace for methanolysis impurities with m/z 249.090 and m/z 232.063, respectively. The two extracted ion traces give peaks with the same retention time thus m/z 232.063 probably results from the source fragmentation of m/z 249.090. The impurity arises from the methanolysis of 6-APA (Figure 6.27).

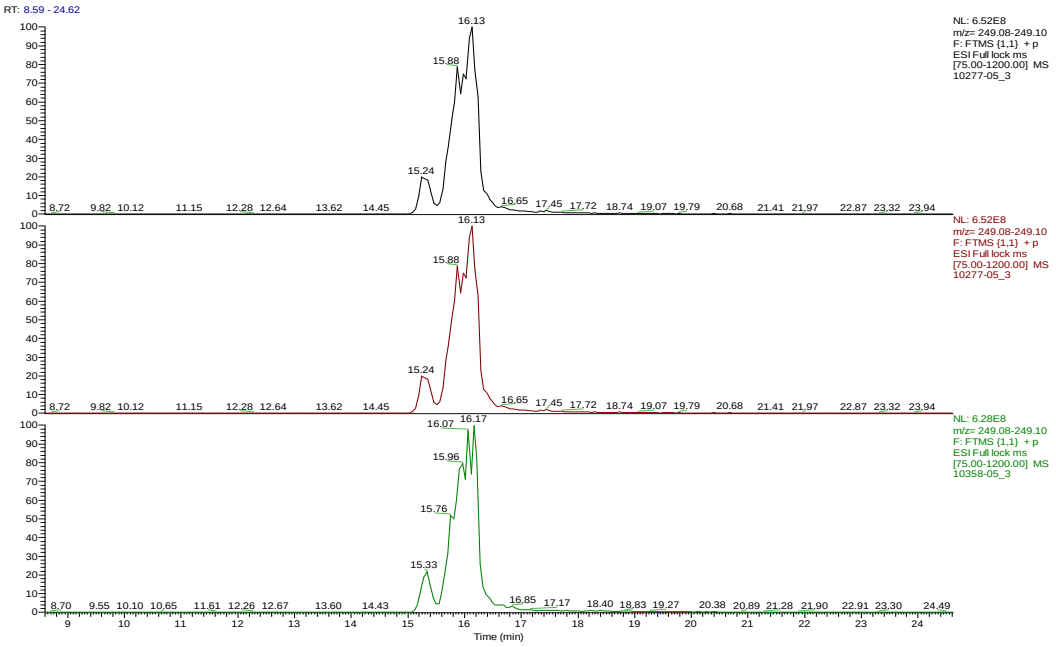


Figure 6.25 EICs for impurity at ([MH]$^+$, m/z 249.0904) in samples with clarity index at 13.6,5.27 and 0.53.

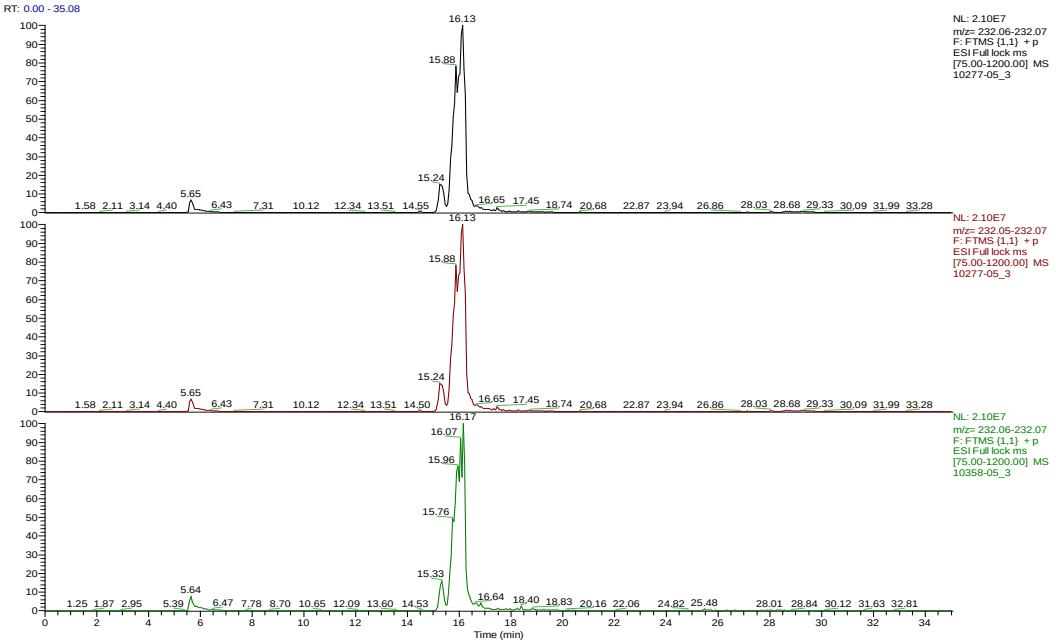


Figure 6.26 EICs for impurity at ([MH]$^+$, m/z 232.0638) in samples with clarity index at 13.6,5.27 and 0.53.

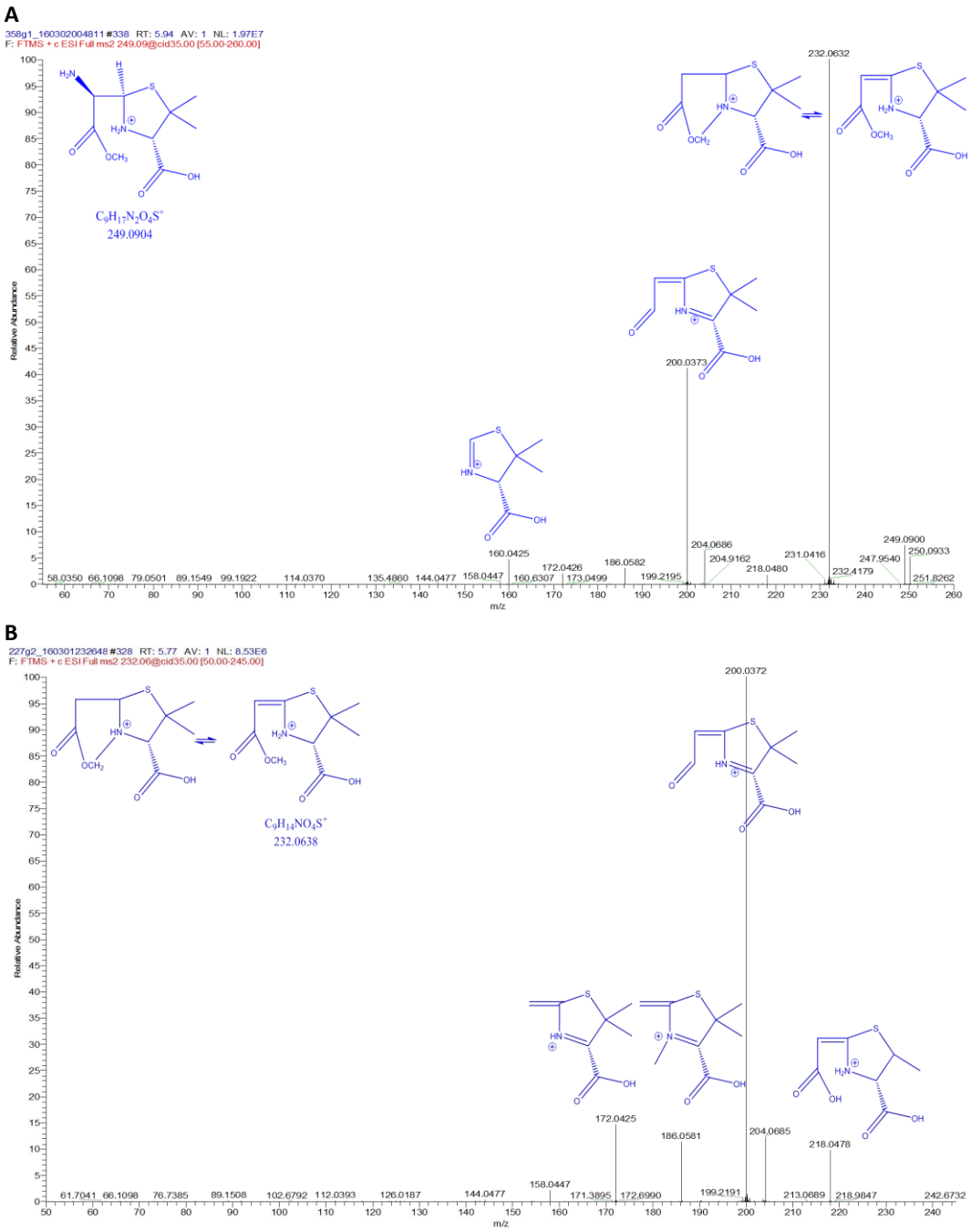


Figure 6.27 MS^2 spectra and proposed structure of impurities at ($[M+H]^+$, m/z 249.09) (A) at ($[M+H]^+$, m/z 232.06) (B) and their fragments.

The MS^2 data revealed that the impurity at m/z 261.0904 has an elemental composition of $C_{10}H_{17}N_2O_4S$. The ion fragmented to produce an ion at m/z 215.048 which is correlated to $C_8H_{11}N_2O_3S$ which results from ethanol group loss. The EIC trace for this impurity is shown in figure 6.28. The proposed structure of the impurity is related to the ethanolysis of 6-APA (Figure 6.29).

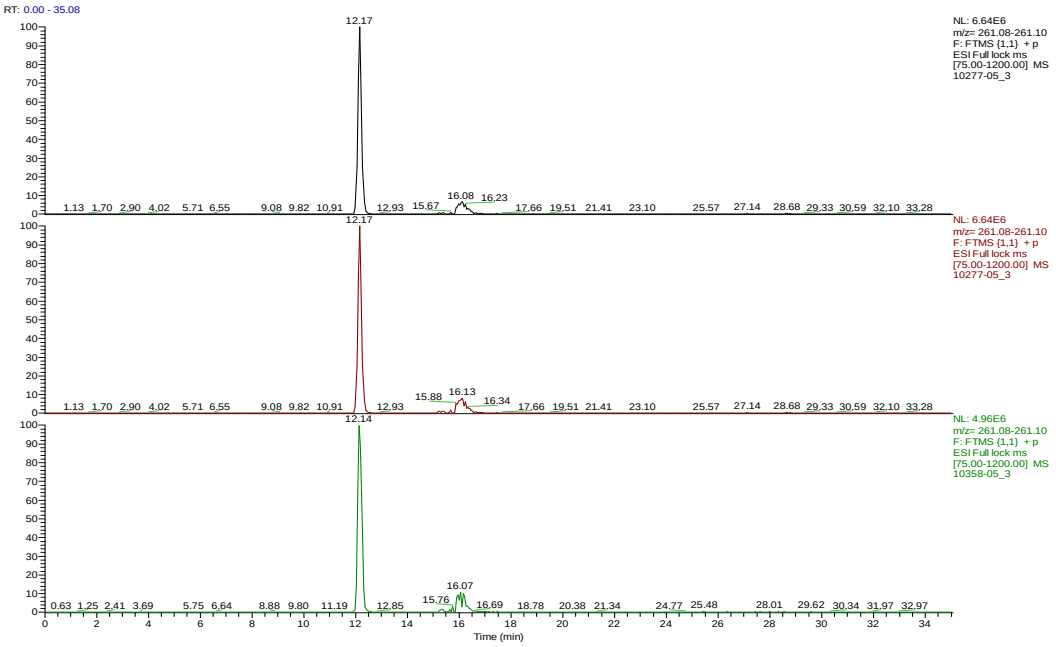


Figure 6.28 EICs for impurity at ([MH]$^+$, m/z 261.0904) in samples with clarity index at 13.6,5.27 and 0.53.

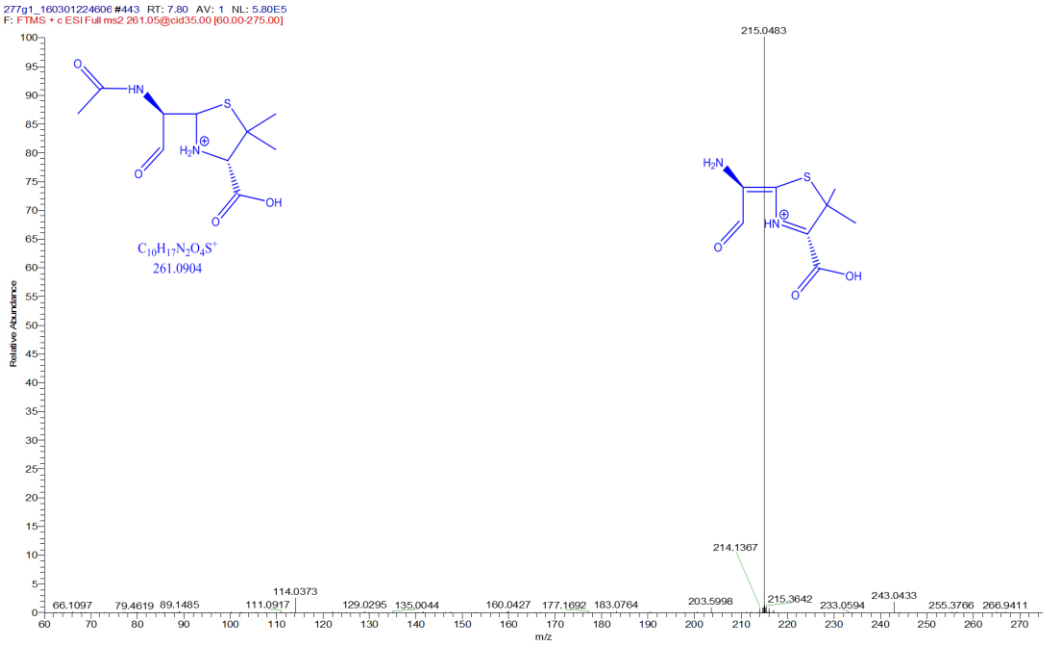


Figure 6.29 MS^2 spectra and proposed structure of impurity at ([M+H]$^+$, m/z 261.0904) and its fragments.

Figure 6.30 shows the EIC trace for dimer of 6-APA impurity at m/z 433.1210. This impurity has the formula $C_{16}H_{25}N_4O_6S$. The MS^2 data provided the fragment ions at m/z 416.092, 274.085, 200.037 and 160.042 of which have elemental compositions of $C_{16}H_{25}N_4O_5S$, $C_{10}H_{16}N_3O_4S$, $C_8H_{10}NO_3S$ and $C_6H_{10}NO_2S$, respectively (Figure 6.31).

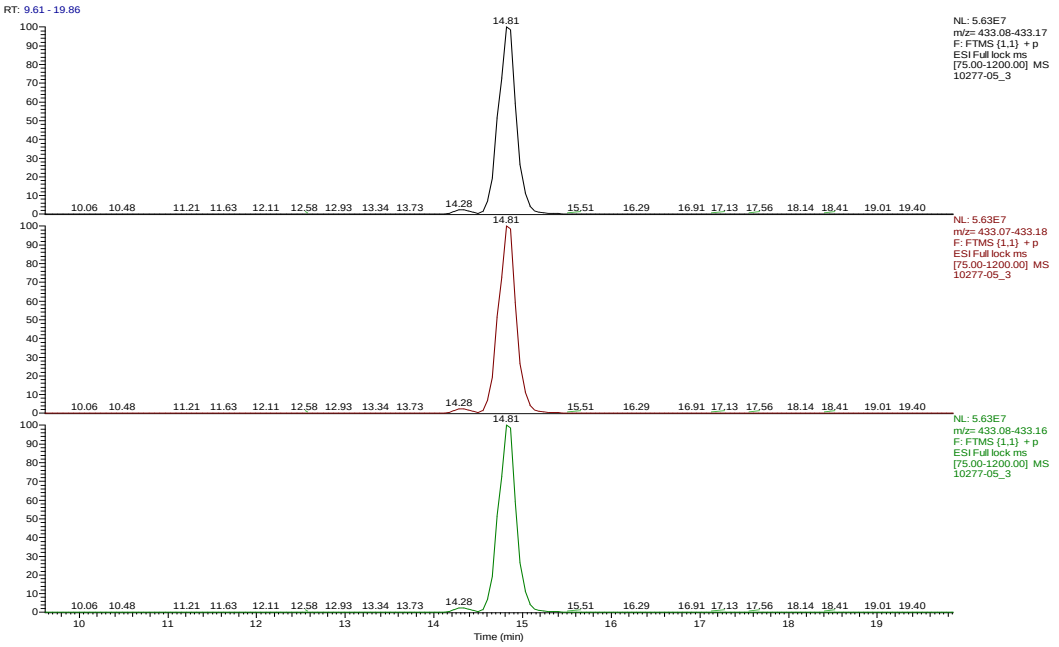


Figure 6.30 EICs for impurity at ([MH]$^+$, m/z 433.1210) in samples with clarity index at 13.6,5.27 and 0.53.

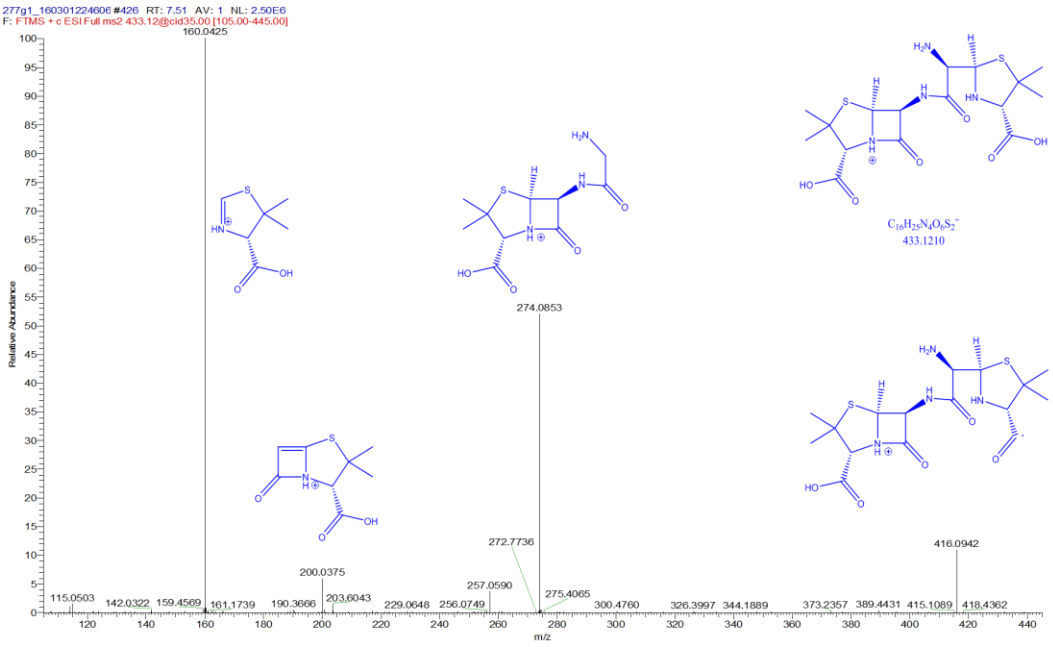


Figure 6.31 MS^2 spectra and proposed structure of impurity at ([M+H]$^+$, m/z 433.1210) and its fragments.

The LC-MS/MS spectra and accurate mass reveal the structure of one of the unknown impurities as a [MH]$^+$ ion with a m/z of 200.0376. The EIC trace for this impurity is shown in figure 6.32. Collision induced dissociation (CID) of this precursor ion generated daughter ions at m/z 172.04, m/z 154.03, m/z 126.04 and m/z 144.05. The

abundant fragment ion at m/z 154.03 had a chemical formula of C_7H_8NOS, resulting from the loss of carboxylic acid group from the precursor ion. The ion at m/z 172 was generated from loss of carbonyl (C=O) group. The fragment ions at m/z 144 and m/z 126 have the chemical formulae of $C_6H_{10}NOS$ and C_6H_8NS, respectively. This leads to a proposed structure of the impurity with m/z 200.0376 as being deaminated 6-APA as shown in Figure 6.33.

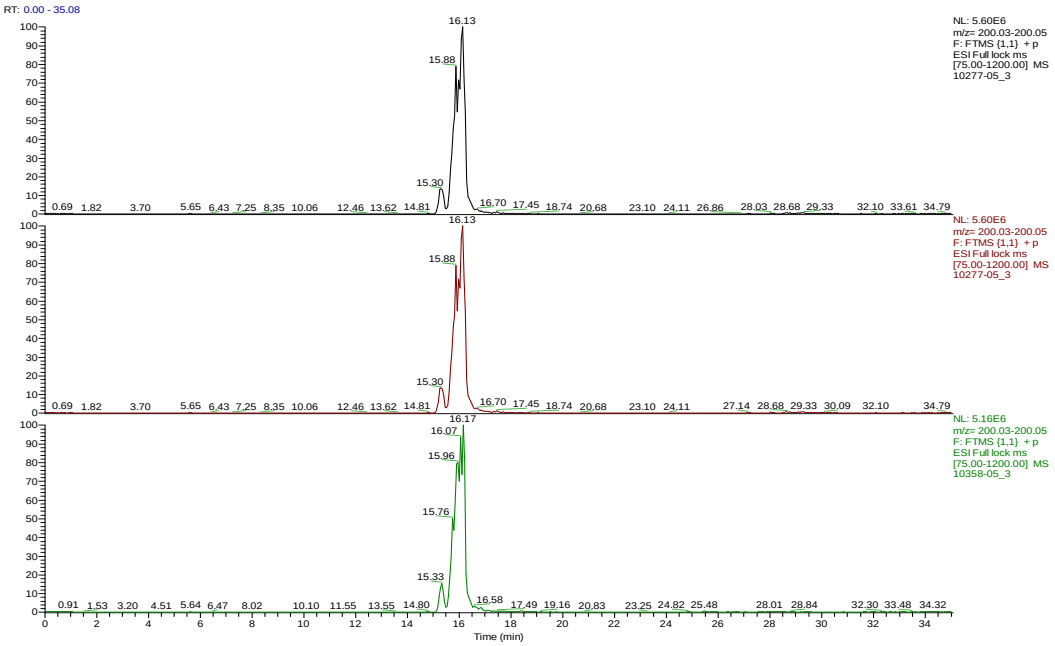


Figure 6.32 EICs for impurity at ([MH]$^+$, m/z 200.0376) in samples with clarity index at 13.6,5.27 and 0.53.

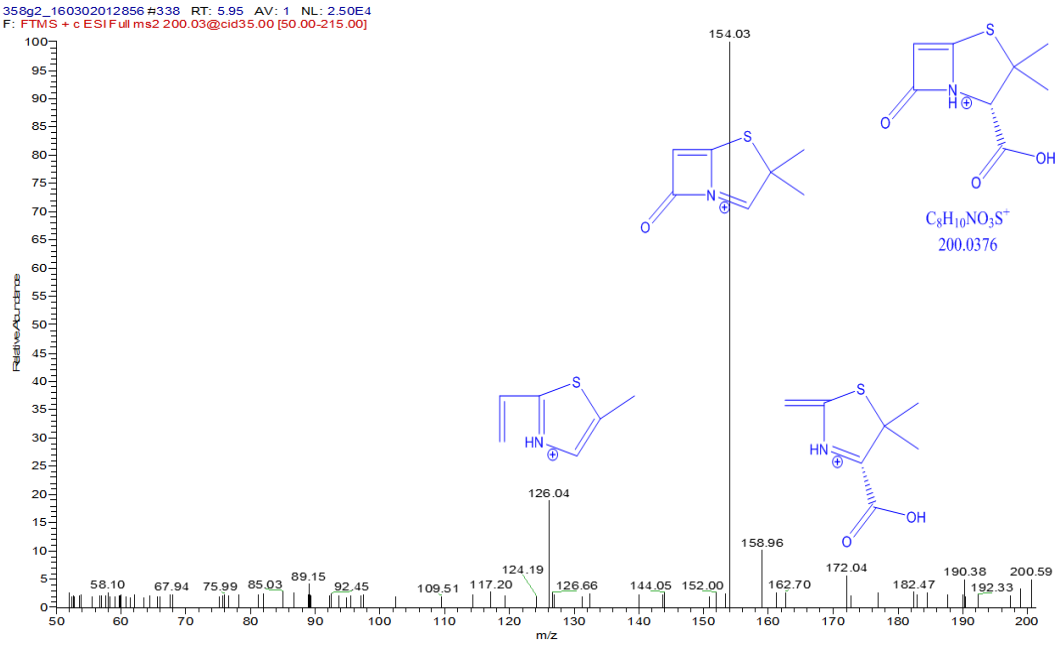
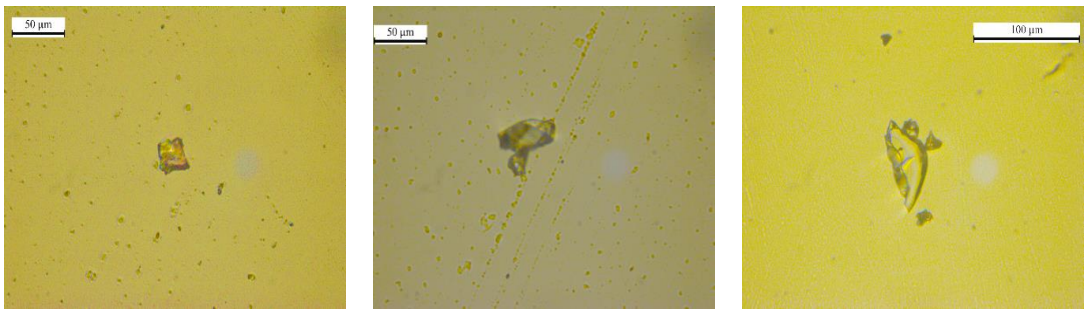


Figure 6.33 MS^2 spectra and proposed structure of the deaminated 6-APA impurity at ($[M+H]^+$, m/z 200.0376) and its fragments.

From the LC-MS and LCMS/MS data analysis, we can confirm that the largest single organic impurity in the powdered samples is the methanolysis product where the β -lactam ring has been opened by methanol.

Optical imaging was done to examine the acid clarity solution in several samples under a high laser scanning confocal microscopy (Leica SP-5 and Leica DM6000) to see if there were any obvious reasons for the difference in the clarity values.

Three samples with clarity index of 0.53, 5.27 and 13.6 were investigated under laser scanning confocal microscopy at magnifications of 50x as shown in the figure 6.34.



Acid clarity value: 0.53 Acid clarity value: 5.27 Acid clarity value: 13.6

Figure 6.34 The optical imaging the acid clarity solution under a scanning electron microscope of the acid clarity solution of three samples with clarity index 13.6, 5.27 and 0.53 using magnification power (x50)

It was reasoned that the particulates were likely to contain high levels of organic, lipophilic materials so they were filtered through a 2 micron nylon filter with the hypothesis that the organic impurities would be retained on the filter and could be analysed by dissolving the residue in an organic solvent.

The residue retained on the filter was eluted from the filter with ethanol and analysed on RP-LC/MS (C18 column) using MS/MS method conditions described in section 2.4. The percentage of the main impurities from LC-MS analysis of acid clarity solution of three samples with clarity index 13.6, 5.27 and 0.53 are shown in the figure 6.35. The main impurity was at m/z 258.0907 which is an unknown impurity. The other impurities were identified in the previous section.

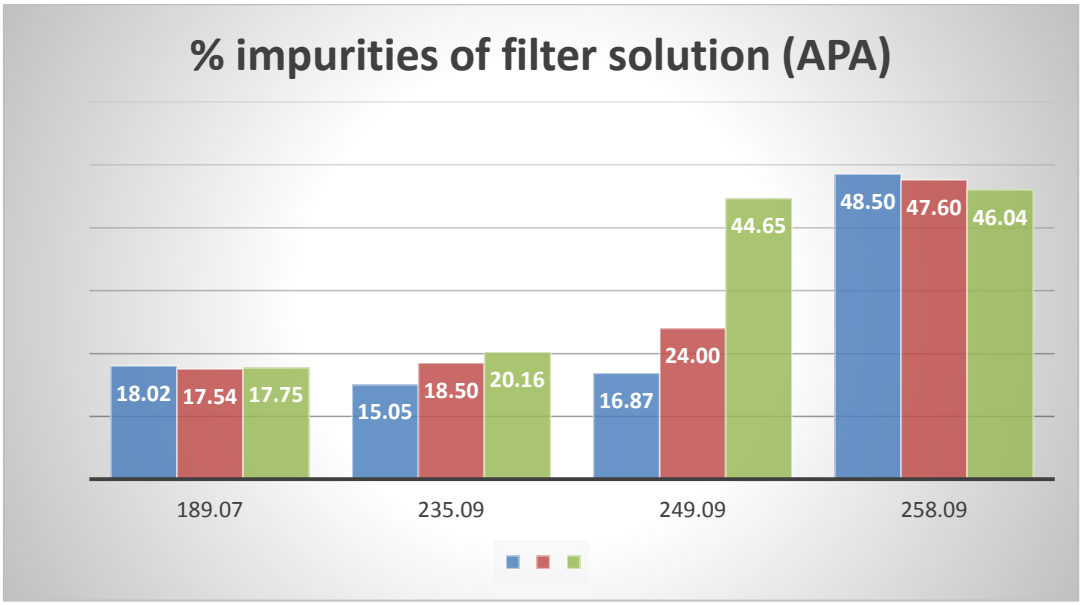


Figure 6.35 % of the main impurities of residue samples from LC-MS analysis of acid clarity solution of three samples with clarity index 13.6 (Green), 5.27 (Red) and 0.53 (Blue).

The chromatogram below (Figure 6.36) shows peak from the MS^2 analysis for the unknown m/z ion of 258.0907 at (RT=12.4min). The protonated ion at m/z 258.0907 had the chemical formula $C_{10}H_{16}N_3O_3S$, the ion fragmented to product ions at m/z 241.06, 160.04 and m/z 141.01 which correlated to $C_{10}H_{13}N_2O_3S$, $C_6H_{10}NO_2S$ and $C_5H_5N_2OS$, respectively. The most abundant fragment at m/z 241.06 is related to the loss of the elements of ammonia from the amine group in the 6-position and a proton from position-5 inserting a 5,6-double bond. The product ion at m/z 160.04 is related to the loss of $C_4H_6N_2O$ moiety from the open ring of the precursor ion (Figure 6.37). These data suggested an interaction between 6-aminopenicillanic acid and the mobile phase acetonitrile with ring-opening of the beta-lactam forming a vinylamide via the 7-carbonyl.

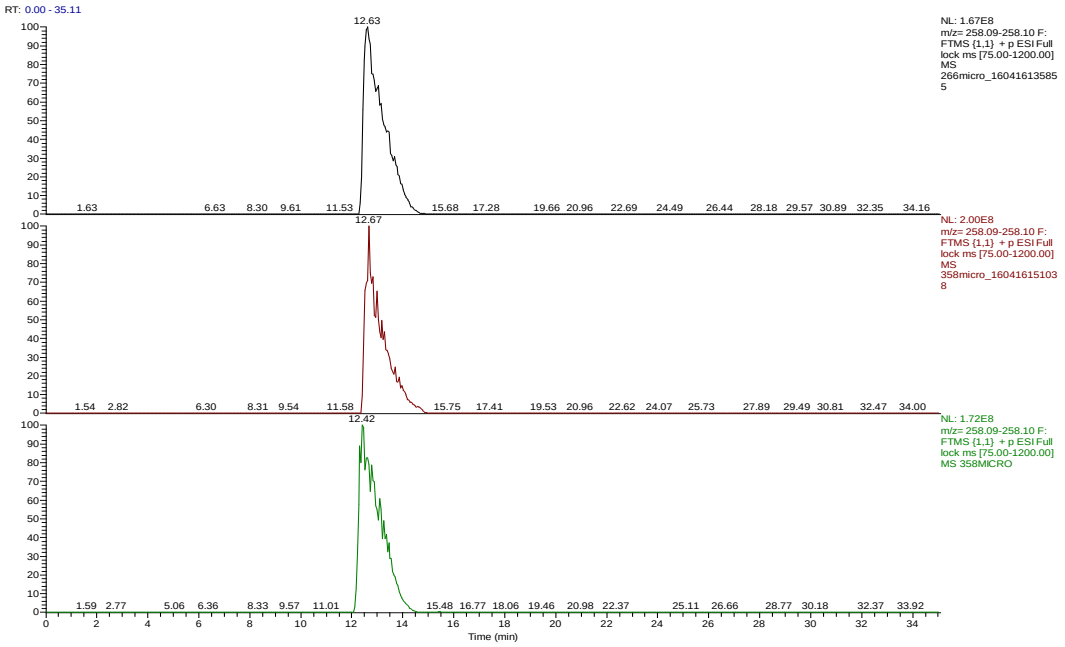


Figure 6.36 EICs for impurity at ([M+H]$^+$, m/z 258.09) in samples with clarity index at 13.6, 5.27 and 0.53.

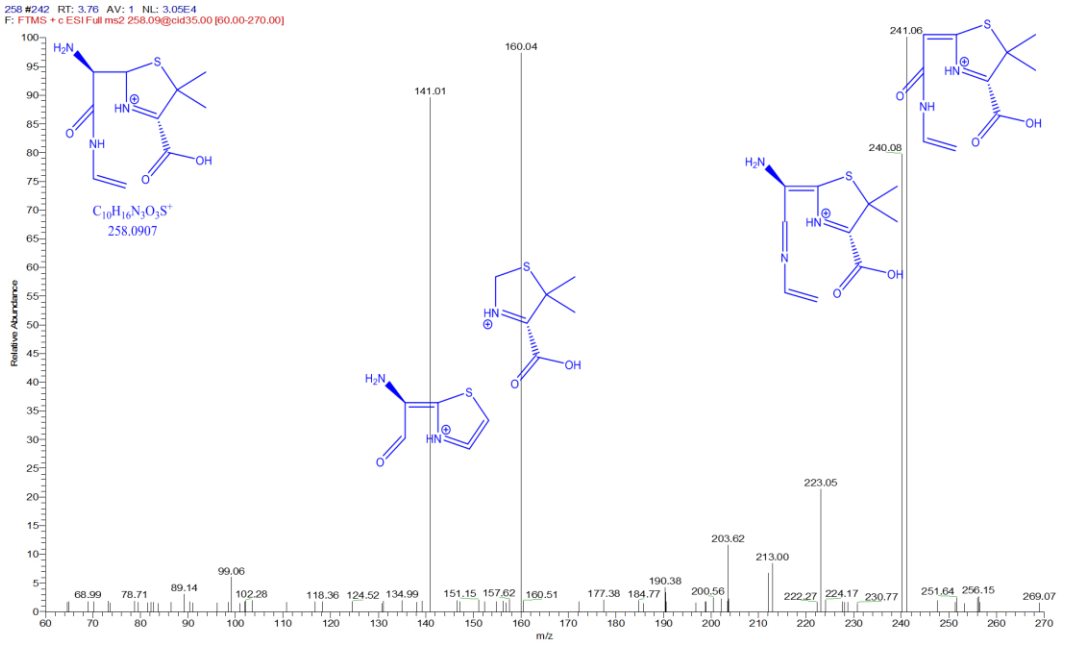
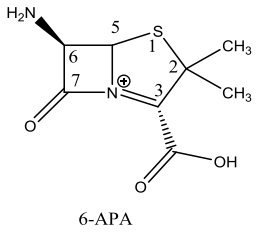


Figure 6.37 MS^2 spectra and proposed structure of impurity at ([M+H]$^+$, m/z 258.09) and its fragments.

The NMR data (Figure S6.2, appendix) revealed that the major component in the sample was actually 6-aminopenicillanic acid which has a molecular weight of 216 amu and not a vinylamide with m/z 258.0907 as suggested by the MS data. Rather,

the 'molecular weight' suggested in the LC-MS must be a positively charged cluster ion composed of the following components; [6-aminopenicillanic acid + acetonitrile + H]$^+$. For one thing, there were no vinylic protons or carbons relating to the vinylamide proposed by the MS analysis. In the HMBC spectrum two geminal methyl singlets at δ1.22 (^{13}C 27.2ppm) and 1.56 (^{13}C 26.8ppm) correlated only with a quaternary carbon at 59.6ppm, a CH at 72.0ppm and to one another's carbons. The carbon at 72.0ppm bore a proton at δ3.75 (HSQC spectrum, Figure S6.2, appendix) which correlated with a carbonyl at 170.2ppm. These chemical shifts were typical of C-2, C-3 and attachments of a penicillanic acid type of compound. Further, C-3 (59.6ppm) had a correlation to a deshielded proton doublet at δ4.95 (H-5, *J*=7.6) and the latter proton coupled in the COSY spectrum to the proton doublet at δ3.96 (H-6). H-5 also had correlations in the HMBC spectrum to C-2 (72.0ppm) and C-7 (169.0ppm). These data were identical to published NMR spectral data for 6-aminopenicillanic acid (Table 6.1) (Crea et al., 2012). It is possible that the turbidity might be due to formation of a small amount of a less soluble polymorph of 6-APA and that this is what is collected on the filter.

Table 6.1 ^{1}H and ^{13}C NMR Spectra of 6-APA and a trace of some impurities of residue samples of acid clarity solution.



6-APA

Atom no.	$^\delta$H (ppm)	Integrity	Multiplicity	*J* (Hz)	$^\delta$C (ppm)	Multiplicity
2					59.5	CH
5	4.94	1H	d	7.6	64	CH
6	3.96	1H	m		57.6	CH
7					169	C
2-CH_3	1.22	3H	s		27.2	CH_3
2-CH_3	1.56	3H	s		26.8	CH_3
3-COOH					170.2	C

6.5 Analysis of Very High Acid Clarity Sample

Three samples were supplied by GSK with two “normal” clarity range samples and one with very high acid clarity value (i.e. high turbidity). It was hoped that the deliberate degradation of the process by running with lots of pH acidity spikes would lead to high levels of impurities that would enable a correlation with the high acid clarity values and impurities to be identified.

6.5.1 Optical Imaging

The optical imaging was done by CMAC group to examine several samples under a high power optical microscope to see if there were any obvious reasons for the high clarity values. Possible causes for the high acid clarity values (corresponds to hazy solutions) could be the presence of solids, a second liquid phase or gas bubbles; it was hoped that optical microscopic could help identify which of these phenomena were occurring.

Solutions were prepared for three samples with clarity index values of 1.49, 1.71 and 89 according to the GSK standard procedure of acid clarity solution, and investigated under polarised mode of microscope at magnifications of 2.5, 5, 10, 20, 50 and 100 (Figure 6.38).

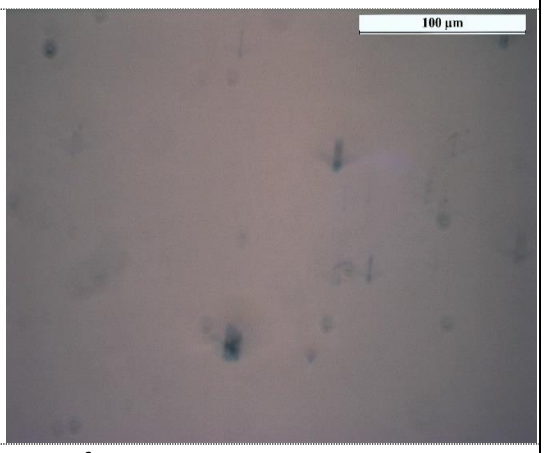
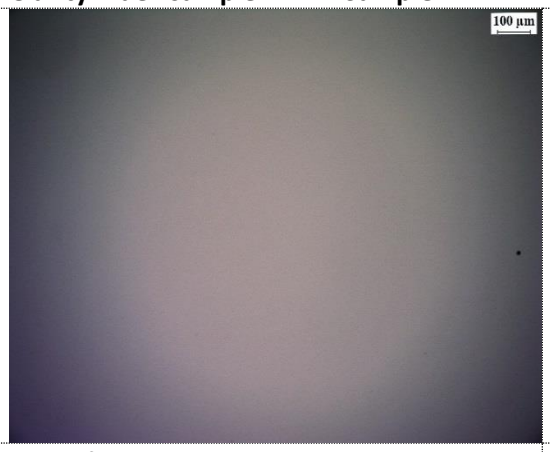
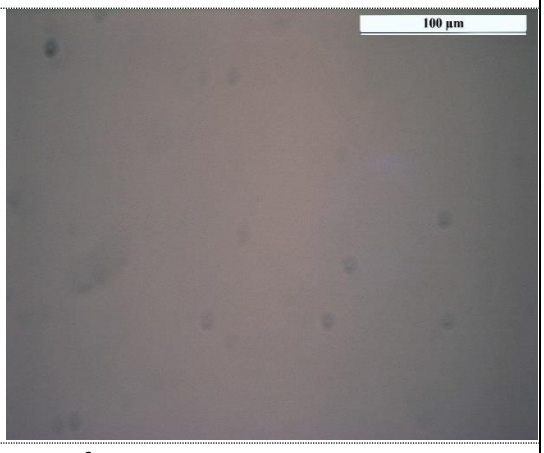
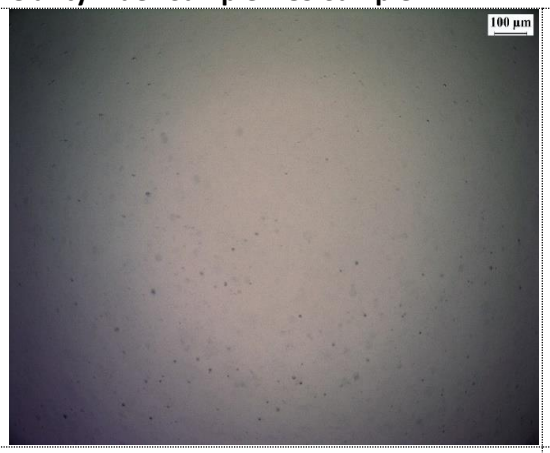
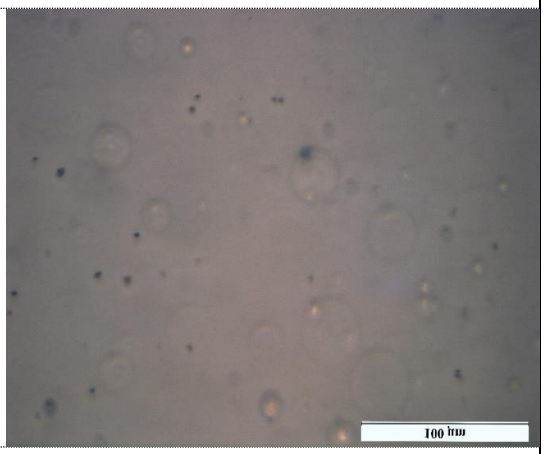
Clarity index = 1.49 Sample	
	
Magnification = 10	Magnification = 50
Clarity Index sample = 1.71 Sample	
	
Magnification = 10	Magnification = 50
Clarity index Sample = 89 Sample	
	
Magnification = 10	Magnification = 50

Figure 6.38 the optical imaging the acid clarity solution under a polarised mode of microscope of acid clarity solution of three samples with clarity index 1.49, 1.71 and 89 (provided by CMAC).
NOTE: dark spots in these images are believed to be dirt on slides.

The results for the samples with clarity index values of 1.49 and 1.71 images indicates small amounts of bubbles or droplets at higher resolution. In the sample with a clarity index value of 89, two phases were observed and this solution had a large amount of small bubbles, droplets or a second liquid phase. There were no visible dirt particles on the slides. According to CMAC, based on previous experience, it is most likely that we are observing a second liquid phase and it is droplets that are observed. Attempts to separate the two phases by centrifugation and ultra-centrifugation were unsuccessful and a discontinuous phase could not be run directly through the HPLC-MS for analysis.

6.5.2 Chemometric analysis of typical and very high acid clarity samples

The three samples clarity index of 1.49, 1.71 and 89 were analysed using HPLC-HRMS in positive ionized mode to investigate the change in the impurity pattern of very high acid clarity sample. The HILIC-MS method was also used for analysing the samples. PCA and OPLS-DA were used to specify the significant impurities that had an effect on the model. The total PCs show cross validation ($R^2X(cum)= 0.995$, $Q^2(cum)= 0.904$) which are considered an excellent level of model fits to the data and reproducibility.

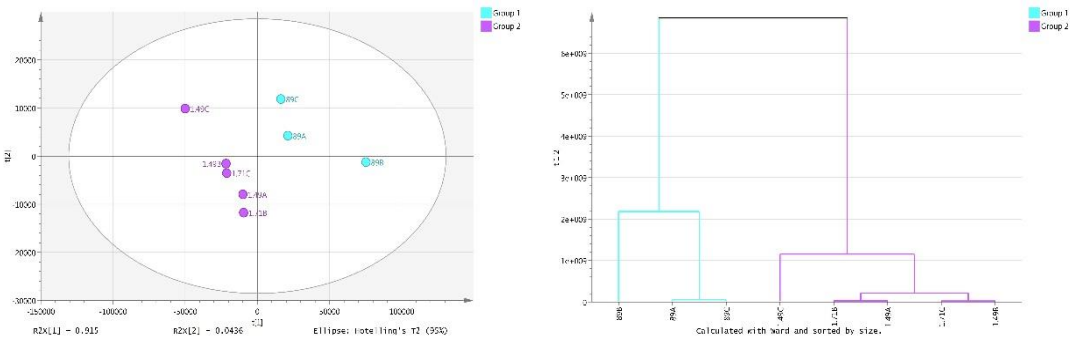


Figure 6.39 PCA scores scatter plot (t1 vs. t2) (left) HCA plot (p1 vs. p2) (right) of PCA analysis of 3 samples with clarity index 1.49, 1.71, and 89 of 6-APA from PCA model of positive ionisation mode HPLC-MS.

PC1 and PC2 obtained by using Pareto scale. Three PCs described 99.5% of the whole variation data (PC1: 91.5%, PC2: 4.36% and PC3: 3.62%). Scores plot shows the discrimination between the samples according to the impurity profile of each sample. Loading plot points out the impurities that govern the position of each sample in the scatter plot. The dendrogram shows observations clustered into two hierarchical groups; (green; 89) and (blue; 1.71and 1.49) by using the Ward clustering method. X-axis represents the samples and y-axis shows the variability index. The values of the variability index among the samples. The larger values show big differences between group variability and the lower values show more similarity.

According to the above data (figure 6.39), the sample with a clarity index 89 (group 1) shows a slight change in the impurity profile. The two samples with lower clarity values at 1.49 and 1.71 (group 2) show no significant difference related to the impurity pattern. OPLS-DA was applied based on the clustering of HCA which discriminates the samples into two groups, one related to the sample with clarity index of 1.49 and 1.71 and the other with the sample with clarity index of 89.

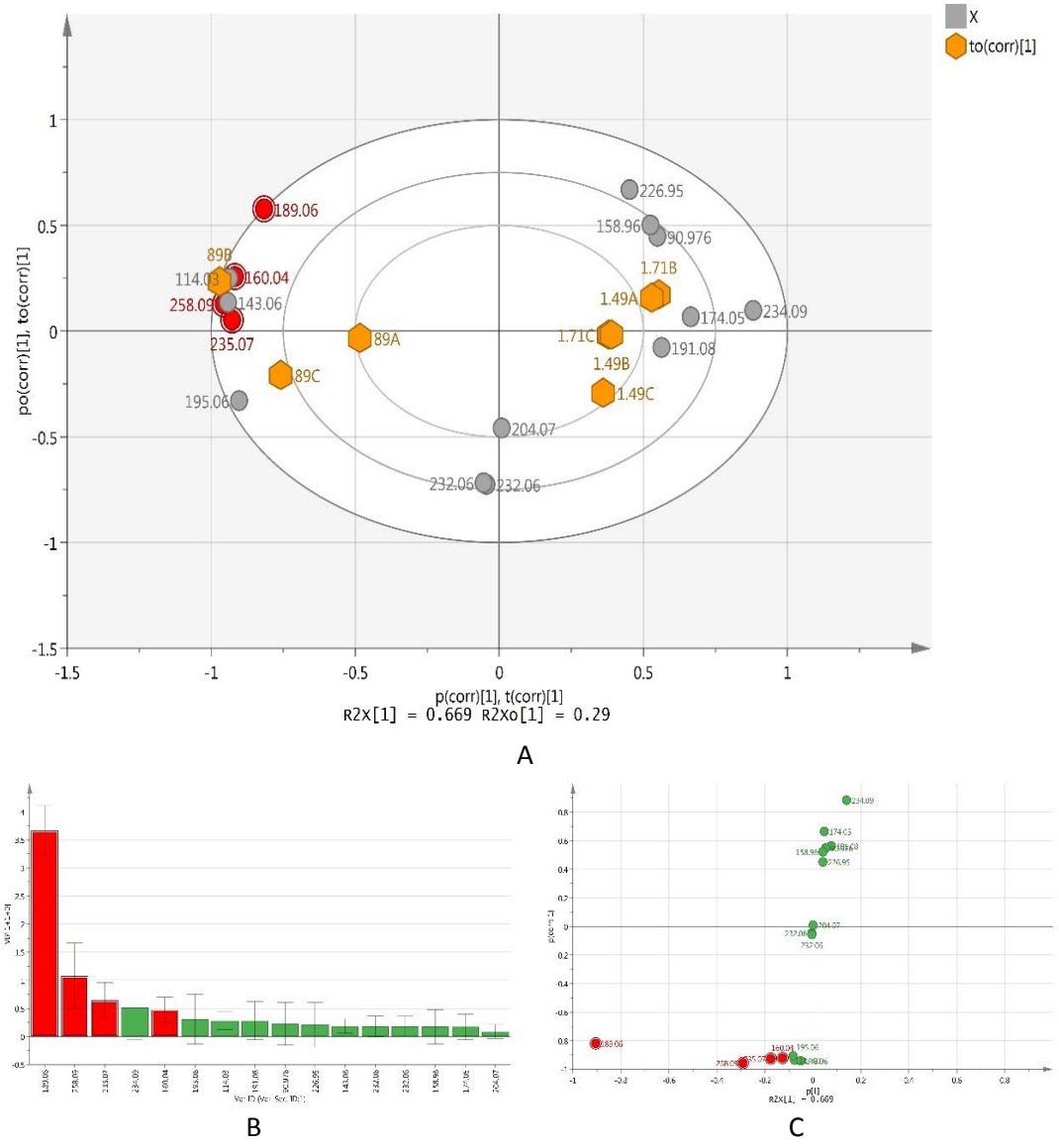


Figure 6.40 SIMCA analysis of very high acid clarity samples
using OPLS-DA (A) Bioplot (scores vs loading) shows distribution of impurities over the samples. (The orange hexagon shape is related to the samples and circles are related to the impurities). (B) VIP plot is sorted from high to low, and shows confidence intervals for the VIP values at the 95% level. (C) S-plot shows ions distributed according to its correlation (x-axis) and magnitude of effect (y-axis) of the groups; (the down red points are related to group 1).

The intention with the Biplot is to co-chart scores and loadings together for their simultaneous display and interpretation. Hence, this plot displays similarities and dissimilarities between samples and allows us to interpret the samples in terms of the variables (Figure 6.40A). The VIP plot is sorted from high to low, and shows confidence intervals for the VIP values at the 95% level. VIP-values larger than 0.5 indicate "important" X-variables. The impurities were labelled by red points show the significant variables (Figure 6.40B). The selective important variables in the S-plot represent the most relevant impurities which contribute in the variation among the groups. These potential impurities were investigated to confirm differences in their average intensity levels by considering the p-values (< 0.05) and 95% CIs. Although there are some species that are slightly related to group 2 (which are shown in the upper right points in S-plot), these are non-significant. The red points on the plot show a significant correlation to the group 1 (Figure 6.40C).

This detailed analysis of the data allows us to identify the relevant m/z ions that correlate to specific items in the VIP plot and these are shown in the table 6.2 below.

Table 6.2 Significant m/z of species Identified in SIMCA which make the difference in sample with clarity index =89.

$[MH]^+$	Molecular formula	Probability	Average (group 1)	Average (group 2)	Std. dev. (group 1)	Std. dev. (group 2)	% change 1 vs 2
189.069	$C_7H_{13}N_2O_2S$	1.84E-02	8.85E+09	6.89E+09	1.15E+09	6.22E+08	1.22
258.091	$C_{10}H_{16}O_3N_3S$	9.67E-04	6.03E+08	4.31E+08	4.70E+07	3.48E+07	1.28
160.043	$C_6H_{10}NO_2S$	2.38E-03	1.76E+08	1.41E+08	9.21E+06	9.49E+06	1.19
235.075	$C_8H_{15}N_2O_4S$	8.87E-03	1.54E+08	9.44E+07	3.14E+07	1.42E+07	1.38

The table above shows that there are 4 ions with significantly different levels between the very high acid clarity sample (group 1) and the more normal level clarity samples (group 2). The increase in the levels of these impurities could be related to the formation of droplets.

6.6 Conclusion

This work described the investigations into the causes of high acid clarity values (i.e. high turbidity) for 6-APA powder manufactured by GSK. Positive ion species, negative ion species, low boiling point organics and neutral species were thoroughly investigated in 16 plant samples with varying levels of acid clarity values. Different techniques were applied including positive and negative mode ion chromatography for cationic and anionic analysis, headspace-GC, HILIC-HRMS, RPLC-MS^2 and NMR. PCA and OPLS-DA were used as statistical methods to determine the pattern of impurities and monitoring the change in the acid clarity value samples.

The key focus of this work was to establish the cause of the formation of a second liquid phase composed of small droplets when the turbidity of the 6-APA was determined. These droplets led to a hazy appearance and low visual clarity of the sample solutions.

The occurrence of high acid clarity values could not be statistically linked to any specific impurity or set of impurities; but major anionic, cationic and organic impurities were identified with possible structures for these being postulated. In addition, several ions were detected that where significantly higher, or lower, in samples with very high clarity values as compared to other standard clarity value samples. It was not possible to identify a single impurity or group of impurities that could be directly correlated to the formation of the second liquid phase causing the high turbidity readings.

The cause of the acid clarity samples has been identified at the production plant as due to pH spikes during the manufacturing process, so a high level of acid catalysed impurity was a likely cause of the high acid clarity values.

Despite extensive chromatographic and chemometric analysis of the results, no specific impurity, or combination of impurities could be identified that correlate in a statistically significant way to a high acid clarity level. We have postulated structures for several of the main impurities that have not been previously identified. A final

theory is that a small amount of an acid insoluble polymorph of 6-APA might be formed as a result of the spikes in acidity.

Chapter 7
General conclusion and Future work

7 General conclusion and Future work

7.1 General conclusion

Alpha Lipoic acid is a potent free radical scavenger and antioxidant. The study investigated a method that able to distinguish the impurity profile of lipoic acid raw material from those of three of its purified samples in order to determine the success of the crystallisation process in removing impurities without introducing further impurities.

The method used was LC/MS^n and multivariate analysis of the acquired chromatographic and mass spectrometric data, to characterise unknown impurities and assess the success of the crystallisation process in removing impurities. In addition, this was performed to provide strong supporting evidence for the identity of the impurities and to understand how the side products might be formed from the synthesis pathway of LA.

This study has shown that the impurity profile of the LA raw material is very similar to, but quite distinguishable from, those of the three purified materials. The observed inter material variability in the impurity profiles might be an effect of the purification process. It is very clear that the raw material is well separated from all the three purified materials, implying that the purification process causes significant reduction in the levels of certain impurities, particularly trimers of LA (with 5 or 6 S atoms) and

LA with extra fatty acid chains with 5 or 6 C atoms. Thus these impurities probably interfere with the co-crystallisation of LA with nicotinamide and their removal through purification improves the crystallisation process.

Chlorpheniramine maleate is a first generation histamine H1 (antihistaminic) receptor antagonist. In this work, Four chlorpheniramine maleate drug substances were investigated for the purpose of impurity profiling using both RP-HPLC and HILIC techniques hyphenated to HRMS and MS^2. The data were investigated to give information on the proposed structure of impurities and degradation pathways of the samples.

Drug impurity analysis achieved by LC/MS offers semi-quantitative results about the concentration of impurities by comparing the peak areas of the impurities with the peak area of API isotope peak. Several new impurities products were detected and identified and. All samples were successfully classified by SIMCA-P in accordance to their impurity profiles associated with their manufacturer/source synthetic route. Thus it is clearly possible to classify such generic materials according to their impurity profiles and this would make possible to spot fake branded products using this type of technique on the basis of the pattern of impurities in the API.

The counterions are commonly present with polar ionisable drug substances and may significantly impact on their quality, safety and efficacy and on the manufacturing

process. Capillary high-pressure ion chromatography (HPIC) with suppressed conductivity detection was employed to achieve a high resolution separation, and optimized to obtain a high sensitivity of detection, for various inorganic anions and low-molecular weight organic acid impurities in selected salt forms of several active pharmaceutical ingredients (APIs) available commercially. The APIs investigated contained counterions of chloride (chlorpromazine and propranolol), maleate (chlorpheniramine), acetate (oxytocin, leu-enkephalin, chlorhexidine and lanreotide), citrate (alverine and orphenadrine) and mesylate (pergolide mesylate and MSA). The API standards that were chosen for the investigation contained different counterions in order to determine whether or not the chosen counterion had a bearing on the levels of anionic impurities. This technique was applied to achieve a high resolution separation of the counterions and was optimized to obtain a high sensitivity of detection.

We found that there were both qualitative and quantitative differences in anionic impurities among the different batches and forms of chlorpromazine HCl. We also observed that the level of anionic impurities was greater in pure enantiomers - which have a more complex processing history - than in corresponding racemic mixtures. The results showed that the acetate counterion is not displaced by anions of stronger acids during processing since, in general, the levels of impurities were no higher than those in the chlorides. The lowest levels of impurities in samples containing acetate were observed in chlorhexidine which is a strongly basic drug. The major trace anionic impurity in mesylate containing samples was nitrate at least 4% and 2% in both pergolide

mesylate and methane sulphonic acid (MSA) standard, respectively. This suggests that this impurity might be related to MSA and during crystallisation may associate more strongly with pergolide mesylate (the drug being studied) than MSA. This method was found to be robust enough for widespread use for fast determination and profiling of inorganic anions and organic acid impurities in APIs.

6-aminopenicillanic acid (6-APA) is the penicillin core structure. This work describes the investigations into the causes of high acid clarity values (i.e. high turbidity) for 6-APA powder. Positive ion species, negative ion species, low boiling point organics and neutral species were thoroughly investigated in 16 plant samples with varying levels of acid clarity values. Different techniques were applied including positive and negative mode ion chromatography for cationic and anionic analysis, headspace-GC-MS, HILIC-HRMS, RPLC-MS^2 and NMR. PCA and OPLS-DA were used as statistical methods to determine the pattern of impurities and monitoring the change in the acid clarity value samples.

The key focus of this work was to establish the cause of the formation of a second liquid phase composed of small droplets when the turbidity of the APA was determined. These droplets led to a hazy appearance and low visual clarity of the sample solutions.

The occurrence of high acid clarity values could not be statistically linked to any specific impurity or set of impurities; but major anionic, cationic and organic impurities were identified with possible structures for these being postulated. In addition, several ions were detected that where significantly higher, or lower, in samples with very high clarity

values as compared to other standard clarity value samples. It was not possible to identify a single impurity or group of impurities that could be directly correlated to the formation of the second liquid phase causing the high turbidity readings.

Despite extensive chromatographic and chemometric analysis of the results, no specific impurity, or combination of impurities could be identified that correlated in a statistically significant way to a high acid clarity level. We have postulated structures for several of the main impurities that have not been already identified. A final theory is that a small amount of an acid insoluble polymorph of 6-APA might be formed as a result of the spikes in acidity.

7.2 Further Work

Further experiments could be carried out on larger numbers of lipoic acid batches in order to correlate impurity profile with the ease of crystallisation.

It would be of interest to apply the impurity profiling methods to batches of formulated APIs in order to see if the final products from different manufacturers could be distinguished on the basis of impurity profile.

It is the author's opinion that analysis of the 6-aminopenicillanic acid impurities present has been completed to a detailed level and further investigation of this area is not likely to lead to identification of the root cause of high acid clarity issues. The only area we

could suggest to investigate further would be analysis of what species are present in the second liquid (droplet) phase that is formed as this may help identify additional causality.

8 References

ACEVSKA, J., STEFKOV, G., CVETKOVIKJ, I., PETKOVSKA, R., KULEVANOVA, S., CHO, J. & DIMITROVSKA, A. 2015. Fingerprinting of morphine using chromatographic purity profiling and multivariate data analysis. *Journal of pharmaceutical and biomedical analysis,* 109**,** 18-27.

AFANADOR, N., TRAN, T. & BUYDENS, L. 2013. Use of the bootstrap and permutation methods for a more robust variable importance in the projection metric for partial least squares regression. *Analytica chimica acta,* 768**,** 49-56.

AHUJA 2007. Assuring quality of drugs by monitoring impurities. *Advanced drug delivery reviews,* 59**,** 3-11.

AHUJA, S., & DONG, M. W 2005. Handbook of Pharmaceutical Analysis by HPLC. *Amsterdam: Elsevier Academic Press.*

ALEXANDER, A. J., HOOKER, T. F. & TOMASELLA, F. P. 2012. Evaluation of mobile phase gradient supercritical fluid chromatography for impurity profiling of pharmaceutical compounds. *J Pharm Biomed Anal,* 70**,** 77-86.

AMES, B. N. 2006. Stability of lipoic acid. Google Patents.

ARMENTA, S., ALCALA, M. & BLANCO, M. 2011. A review of recent, unconventional applications of ion mobility spectrometry (IMS). *Anal Chim Acta,* 703**,** 114-123.

B'HYMER, C. 2003. Residual Solvent Testing: A Review of Gas-Chromatographic and Alternative Techniques. *Pharmaceutical Research,* 20**,** 337-344.

BAERTSCHI, S. W. 2006. Analytical methodologies for discovering and profiling degradation-related impurities. *TrAC Trends in Analytical Chemistry,* 25**,** 758-767.

BANSAL, A. K., KUMAR, L. & AESHNA, A. 2008. Salt selection in drug development *Pharmaceutical Technology,* 3**,** 128–146.

BEEBE, K. R., PELL, R. J. & SEASHOLTZ, M. B. 1998. *Chemometrics: a practical guide*, Wiley-Interscience.

BHATTACHARYYA, L. & ROHRER, J. S. 2012. Ion-Exchange Chromatography. *Applications of ion chromatography in the analysis of pharmaceutical and biological products.* John Wiley & Sons.

BIEWENGA, G. P., HAENEN, G. R. & BAST, A. 1997. The pharmacology of the antioxidant lipoic acid. *General Pharmacology: The Vascular System,* 29**,** 315-331.

BLANCO, M., COELLO, J., ITURRIAGA, H., MASPOCH, S. & PÉREZ-MASEDA, C. 2000. Determination of polymorphic purity by near infrared spectrometry. *Analytica Chimica Acta,* 407**,** 247-254.

BROOKES, M. 1986. The synthesis of the enantiomers of lipoic acid (Unpublished doctoral dissertation). *University of Warwick, England*.

BUSZEWSKI, B. & NOGA, S. 2012. Hydrophilic interaction liquid chromatography (HILIC)--a powerful separation technique. *Anal Bioanal Chem,* 402**,** 231-47.

BYLESJÖ, M., RANTALAINEN, M., CLOAREC, O., NICHOLSON, J. K., HOLMES, E. & TRYGG, J. 2006. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics,* 20**,** 341-351.

CHAUVE, B., GUILLARME, D., CLEON, P. & VEUTHEY, J. L. 2010. Evaluation of various HILIC materials for the fast separation of polar compounds. *J Sep Sci,* 33**,** 752-64.

CHEN, J., JIANG, W., CAI, J., TAO, W., GAO, X. & JIANG, X. 2005. Quantification of lipoic acid in plasma by high-performance liquid chromatography–electrospray ionization mass spectrometry. *Journal of Chromatography B,* 824**,** 249-257.

CHNG, H. T., NEW, L. S., NEO, A. H., GOH, C. W., BROWNE, E. R. & CHAN, E. C. 2010. A sensitive LC/MS/MS bioanalysis assay of orally administered lipoic acid in rat blood and brain tissue. *Journal of pharmaceutical and biomedical analysis,* 51**,** 754-757.

CONBOY, J., HENION, J. D., MARTIN, M. & ZWEIGENBAUM, J. 1990. Ion chromatography/mass spectrometry for the determination of organic ammonium and sulfate compounds. *Analytical Chemistry,* 62**,** 800-807.

CREA, F., CUCINOTTA, D., DE STEFANO, C., MILEA, D., SAMMARTANO, S. & VIANELLI, G. 2012. Modeling solubility, acid–base properties and activity coefficients of amoxicillin, ampicillin and (+) 6-aminopenicillanic acid, in NaCl (aq) at different ionic strengths and temperatures. *European Journal of Pharmaceutical Sciences,* 47**,** 661-677.

DE BEER, T., BODSON, C., DEJAEGHER, B., WALCZAK, B., VERCRUYSSE, P., BURGGRAEVE, A., LEMOS, A., DELATTRE, L., HEYDEN, Y. V., REMON, J. P., VERVAET, C. & BAEYENS, W. R. 2008. Raman spectroscopy as a process analytical technology (PAT) tool for the in-line monitoring and understanding of a powder blending process. *J Pharm Biomed Anal,* 48**,** 772-9.

DE BEER, T., BURGGRAEVE, A., FONTEYNE, M., SAERENS, L., REMON, J. P. & VERVAET, C. 2011. Near infrared and Raman spectroscopy for the in-process monitoring of pharmaceutical production processes. *Int J Pharm,* 417**,** 32-47.

DE CARVALHO ROCHA, W. F. & NOGUEIRA, R. 2011. Use of multivariate statistical analysis to evaluate experimental results for certification of two pharmaceutical reference materials. *Accreditation and quality assurance,* 16**,** 523.

DROZDZEWSKA, K., KESTENS, V., HELD, A., ROEBBEN, G. & LINSINGER, T. 2007. DIFFERENTIAL SCANNING CALORIMETRY TO MEASURE THE PURITY OF POLYCYCLIC AROMATIC HYDROCARBONS. *Journal of Thermal Analysis and Calorimetry,* 88**,** 757–762.

DURRANI, A. I., SCHWARTZ, H., NAGL, M. & SONTAG, G. 2010. Determination of free α-lipoic acid in foodstuffs by HPLC coupled with CEAD and ESI-MS. *Food Chemistry,* 120**,** 1143-1148.

ECKERS, C., LAURES, A. M., GILES, K., MAJOR, H. & PRINGLE, S. 2007. Evaluating the utility of ion mobility separation in combination with high-pressure liquid chromatography/mass spectrometry to facilitate detection of trace impurities in formulated drug products. *Rapid Commun Mass Spectrom,* 21**,** 1255-63.

EDGAR C. NICOLAS , T. H. S. 1998. Active drug substance impurity profiling Part II. LC:MS:MS fingerprinting. *Journal of Pharmaceutical and Biomedical Analysis,* 16**,** 825-836.

EL-HAGRASY, A. S., DELGADO-LOPEZ, M. & DRENNEN, J. K., 3RD 2006. A Process Analytical Technology approach to near-infrared process control of pharmaceutical powder blending: Part II: Qualitative near-infrared models for prediction of blend homogeneity. *J Pharm Sci,* 95**,** 407-21.

ELDER, D. P., DELANEY, E., TEASDALE, A., EYLEY, S., REIF, V. D., JACQ, K., FACCHINE, K. L., OESTRICH, R. S., SANDRA, P. & DAVID, F. 2010. The utility of sulfonate salts in drug development. *Journal of pharmaceutical sciences,* 99**,** 2948-2961.

ELDER, D. P. & SNODIN, D. J. 2009. Drug substances presented as sulfonic acid salts: overview of utility, safety and regulation. *J Pharm Pharmacol,* 61**,** 269-78.

ERIKSSON, L., BYRNE, T., JOHANSSON, E., TRYGG, J. & VIKSTRÖM, C. 2013a. Hierarchical cluster analysis, HCA. *Multi-and megavariate data analysis basic principles and applications*

Umetrics Academy.

ERIKSSON, L., BYRNE, T., JOHANSSON, E., TRYGG, J. & VIKSTRÖM, C. 2013b. Identification of discriminating variables. *Multi-and megavariate data analysis basic principles and applications.* Umetrics Academy.

ERIKSSON, L., BYRNE, T., JOHANSSON, E., TRYGG, J. & VIKSTRÖM, C. 2013c. Orthogonal PLS (OPLS). *Multi-and megavariate data analysis basic principles and applications.*: Umetrics Academy.

ERIKSSON, L., BYRNE, T., JOHANSSON, E., TRYGG, J. & VIKSTRÖM, C. 2013d. Principal component analysis. *Multi-and megavariate data analysis basic principles and applications.* Umetrics Academy.

ERIKSSON, L., BYRNE, T., JOHANSSON, E., TRYGG, J. & VIKSTRÖM, C. 2013e. Transformation and expansion. *Multi-and megavariate data analysis basic principles and applications.* UMETRICS ed. Sweden: Umetrics Academy.

F. A. CHMILENKO, I. V. KOROBOVA & DANILENKO, L. N. 2000. Determination of a polymer surfactant by potentiometry with an ionselective electrode using inorganic or organic anions as counterions of the electrodeactive compound. *Journal of Analytical Chemistry,* 55**,** 1058-1062.

FERNANDEZ, F. M., CODY, R. B., GREEN, M. D., HAMPTON, C. Y., MCGREADY, R., SENGALOUNDETH, S., WHITE, N. J. & NEWTON, P. N. 2006. Characterization of solid counterfeit drug samples by desorption electrospray ionization and direct-analysis-in-real-time coupled to time-of-flight mass spectrometry. *ChemMedChem,* 1**,** 702-5.

FERREIRA, A. P. & TOBYN, M. 2015. Multivariate analysis in the pharmaceutical industry: enabling process understanding and improvement in the PAT and QbD era. *Pharmaceutical development and technology,* 20**,** 513-527.

FREED, A. L., KALE, U., ANDO, H., ROSSI, D. T. & KINGSMILL, C. A. 2004. Improving the detection of degradants and impurities in pharmaceutical drug products by applying mass spectral and chromatographic searching. *J Pharm Biomed Anal,* 35**,** 727-38.

FRITZ, J. S. 2000. Recent developments in the separation of inorganic and small organic ions by capillary electrophoresis. *Journal of Chromatography A,* 884**,** 261–275.

GALINDO-PRIETO, B., ERIKSSON, L. & TRYGG, J. 2015. Variable influence on projection (VIP) for OPLS models and its applicability in multivariate time series analysis. *Chemometrics and Intelligent Laboratory Systems,* 146**,** 297-304.

GIRON, D. & GOMBRONN, C. 1995. PLACE OF DSC PURITY ANALYSIS IN PHARMACEUTICAL DEVELOPMENT. *Journal of Thermal Analysis and Calorimetry - J THERM ANAL CALORIM , ,* 44**,** 217-251.

GÖRÖG, S. 2006. The importance and the challenges of impurity profiling in modern pharmaceutical analysis. *TrAC Trends in Analytical Chemistry,* 25**,** 755-757.

GRODOWSKA, K. & PARCZEWSKI, A. 2010. Organic solvents in the pharmaceutical industry. *Acta Pol Pharm,* 67**,** 3-12.

HADDAD, P. R., JACKSON, P. E. & SHAW, M. J. 2003. Developments in suppressor technology for inorganic ion analysis by ion chromatography using conductivity detection. *Journal of Chromatography A,* 1000**,** 725-742.

ICH-Q3A R2 (2002). Impurities in New Drugs Substances Q3A (R2). *The International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use*.

ICH-Q3A(R2) 25 October 2006. Impurities in New Drug Substances. ICH Harmonised Tripartite Guideline.

ICH GUIDELINE Q3D, I. H. Guideline for Elemental Impurities q3d. *Current Step,* 4.

ICH GUIDELINE Q3C, I. H. T. 2005. Impurities: Guideline for residual solvents Q3C (R5). *Current Step,* 4**,** 509.

JOLLIFFE, I. 2002. *Principal component analysis*, Wiley Online Library.

KANTIANI, L., FARRÉ, M., BARCELÓ, D. & BARCELÓ, D. 2009. Analytical methodologies for the detection of β-lactam antibiotics in milk and feed samples. *TrAC Trends in Analytical Chemistry,* 28**,** 729-744.

KATAOKA, H. 1998. Chromatographic analysis of lipoic acid and related compounds. *Journal of Chromatography B: Biomedical Sciences and Applications,* 717**,** 247-262.

KOVALESKI, J., KRAUT, B., MATTIUZ, A., GIANGIULIO, M., BROBST, G., CAGNO, W., KULKARNI, P. & RAUCH, T. 2007. Impurities in generic pharmaceutical development. *Adv Drug Deliv Rev,* 59**,** 56-63.

KREFT, K., KOZAMERNIK, B. & URLEB, U. 1999. Qualitative determination of polyvinylpyrrolidone type by near-infrared spectrometry. *International Journal of Pharmaceutics,* 177**,** 1-6.

KRISHNAN, C. V., GARNETT, M. & ANTONAWICH, F. Free radicals in neurodegenerative diseases: modulation by palladium α-lipoic acid complex. *Neurodegenerative diseases-processes, prevention, protection and monitoring***,** 89-126.

LAURIDSEN, C., LEONARD, S. W., GRIFFIN, D. A., LIEBLER, D. C., MCCLURE, T. D. & TRABER, M. G. 2001. Quantitative Analysis by Liquid Chromatography–Tandem Mass Spectrometry of Deuterium-Labeled and Unlabeled Vitamin E in Biological Samples. *Analytical Biochemistry,* 289**,** 89-95.

LI, M., WANG, X., CHEN, B., LIN, M., BUEVICH, A. V., CHAN, T. M. & RUSTUM, A. M. 2009. Use of liquid chromatography/tandem mass spectrometric molecular fingerprinting for the rapid structural identification of pharmaceutical impurities. *Rapid Commun Mass Spectrom,* 23**,** 3533-42.

MARSHALL, A., GRAUL, R., MORGAN, M. Y. & SHERLOCK, S. 1982. Treatment of alcohol-related liver disease with thioctic acid: a six month randomised double-blind trial. *Gut,* 23**,** 1088-1093.

MATHKAR, S., KUMAR, S., BYSTOL, A., OLAWOORE, K., MIN, D., MARKOVICH, R. & RUSTUM, A. 2009. The use of differential scanning calorimetry for the purity verification of pharmaceutical reference standards. *J Pharm Biomed Anal,* 49**,** 627-31.

MICHALSKI, A., DAMOC, E., HAUSCHILD, J.-P., LANGE, O., WIEGHAUS, A., MAKAROV, A., NAGARAJ, N., COX, J., MANN, M. & HORNING, S. 2011. Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Molecular & Cellular Proteomics,* 10**,** M111. 011015.

MOLDOVEANU, S. & DAVID, V. 2013. Essentials in Modern HPLC Separations. Amsterdam: Elsevier.

MULLIGAN, K. J., BRUEGGEMEYER, T. W., CROCKETT, D. F. & SCHEPMAN, J. B. 1996. Analysis of organic volatile impurities as a forensic tool for the examination of bulk pharmaceuticals. *Journal of Chromatography B: Biomedical Sciences and Applications,* 686**,** 85-95.

MULLIN, J. W. 2001. *Crystallization,* Oxford Butterworth-Heinemann,.

NAGY, Z. K., CHEW, J. W., FUJIWARA, M. & BRAATZ, R. D. 2008. Comparative performance of concentration and temperature controlled batch crystallizations. *Journal of Process Control,* 18**,** 399-407.

NICOLAS, E. C. & SCHOLZ, T. H. 1998. Active drug substance impurity profiling: Part II. LC/MS/MS fingerprinting. *Journal of pharmaceutical and biomedical analysis,* 16**,** 825-836.

NICHOLS, D. E. 2006. Synthesis of Essential Drugs By Ruben Vardanyan and Victor Hruby. Elsevier, Amsterdam, The Netherlands. 2006. xvi+ 617 pp. 17× 24.5 cm. ISBN 10 0-444-52166-6. $240.00. ACS Publications.

NIKOLIĆ, R. S., KRSTIĆ, N. S., NIKOLIĆ, G. M., KOCIĆ, G. M., CAKIĆ, M. D. & ANĐELKOVIĆ, D. H. 2014. Molecular mechanisms of beneficial effects of lipoic acid in copper intoxicated rats assessment by FTIR and ESI-MS. *Polyhedron,* 80**,** 223-227.

OTTENS, M., LEBRETON, B., ZOMERDIJK, M., RIJKERS, M., BRUINSMA, O. & VAN DER WIELEN, L. 2001. Crystallization kinetics of ampicillin. *Industrial & engineering chemistry research,* 40**,** 4821-4827.

PACKER, L., ROY, S. & SEN, C. K. 1996. α-Lipoic Acid: A Metabolic Antioxidant and Potential Redox Modulator of Transcription. *In:* HELMUT, S. (ed.) *Advances in Pharmacology.* Academic Press.

PACKER, L., WITT, E. H. & TRITSCHLER, H. J. 1995. Alpha-lipoic acid as a biological antioxidant. *Free Radical Biology and Medicine,* 19**,** 227-250.

PAGE, P. C. B., RAYNER, C. M. & SUTHERLAND, I. O. 1990. An enantioselective synthesis of R-(+)-α-lipoic acid. *Journal of the Chemical Society, Perkin Transactions 1***,** 1615-1618.

PARMAR, A., KUMAR, H., MARWAHA, S. & KENNEDY, J. 2000. Advances in enzymatic transformation of penicillins to 6-aminopenicillanic acid (6-APA). *Biotechnology advances,* 18**,** 289-301.

PLUMB, R. S., JONES, M. D., RAINVILLE , P. D. & NICHOLSON, J. K. 2008. A Rapid Simple Approach to Screening Pharmaceutical Products Using Ultra-Performance LC Coupled to Time-of-Flight Mass Spectrometry and Pattern Recognition. *Journal of Chromatographic Science,* 46**,** 193-198.

PLUSKAL, T., CASTILLO, S., VILLAR-BRIONES, A. & OREŠIČ, M. 2010. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics,* 11**,** 395.

RAMAKRISHNAN, N., WOLFE, W. W. & CATRAVAS, G. N. 1992. Radioprotection of hematopoietic tissues in mice by lipoic acid. *Radiation research,* 130**,** 360-365.

RELJANOVIC, M., REICHEL, G., RETT, K., LOBISCH, M., SCHUETTE, K., MÖLLER, W., TRITSCHLER, H.-J. & MEHNERT, H. 1999. Treatment of diabetic polyneuropathy with the antioxidant thioctic acid (α-lipoic acid): a two year multicenter randomized double-blind placebo-controlled trial (ALADIN II). *Free radical research,* 31**,** 171-179.

ROCHELEAU, M.-J. 2008. Analytical Methods for Determination of Counter-ions in Pharmaceutical Salts. *Current Pharmaceutical Analysis,* 4**,** 25-32.

ROGGO, Y., CHALUS, P., MAURER, L., LEMA-MARTINEZ, C., EDMOND, A. & JENT, N. 2007. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *J Pharm Biomed Anal,* 44**,** 683-700.

ROLINSON, G. & GEDDES, A. 2007. The 50th anniversary of the discovery of 6-aminopenicillanic acid (6-APA). *International journal of Antimicrobial agents,* 29**,** 3-8.

ROMERO-TORRES, S., WIKSTRÖM, H., GRANT, E. R. & TAYLOR, L. S. 2007. Monitoring of mannitol phase behavior during freeze-drying using non-invasive Raman spectroscopy. *PDA J. Pharm. Sci. Technol,* 61**,** 131-145.

ROY, J. 2002. Pharmaceutical Impurities- A Mini-Review. *AAPS PharmSciTech,* 3**,** article 6.

SACRE, P. Y., DECONINCK, E., DASZYKOWSKI, M., COURSELLE, P., VANCAUWENBERGHE, R., CHIAP, P., CROMMEN, J. & DE BEER, J. O. 2011. Impurity fingerprints for the identification of counterfeit medicines--a feasibility study. *Anal Chim Acta,* 701**,** 224-31.

SANTOS PEREIRA, L. N. D., DA SILVA, I. S., ARAÚJO, T. P., TANAKA, A. A. & ANGNES, L. 2016. Fast quantification of α-lipoic acid in biological samples and dietary supplements using batch injection analysis with amperometric detection. *Talanta,* 154**,** 249-254.

SATINDER AHUJA, M. D. 2005. Handbook of Pharmaceutical Analysis by HPLC.

SATOH, S., SHINDOH, M., MIN, J. Z., TOYO'OKA, T., FUKUSHIMA, T. & INAGAKI, S. 2008. Selective and sensitive determination of lipoyllysine (protein-bound α-lipoic acid) in biological specimens by high-performance liquid chromatography with fluorescence detection. *Analytica chimica acta,* 618**,** 210-217.

SATOH, S., TOYO'OKA, T., FUKUSHIMA, T. & INAGAKI, S. 2007. Simultaneous determination of α-lipoic acid and its reduced form by high-performance liquid chromatography with fluorescence detection. *Journal of Chromatography B,* 854**,** 109-115.

SHEWALE, J. & SIVARAMAN, H. 1989. Penicillin acylase: enzyme production and its application in the manufacture of 6-APA. *Process Biochemistry,* 24**,** 146-154.

SNOW, N. H. & SLACK, G. C. 2002. Head-space analysis in modern gas chromatography. *TrAC Trends in Analytical Chemistry,* 21**,** 608-617.

SRIRAM, D. & YOGEESWARI, P. 2010. *Medicinal chemistry,* New Delhi, Dorling Kindersley (India).

STAHL, P. H. & WERMUTH, C. G. 2002. Handbook of pharmaceutical salts: properties, selection and use. *Chemistry International,* 24**,** 21.

STEPANOVA, S. & KASICKA, V. 2014. Determination of impurities and counterions of pharmaceuticals by capillary electromigration methods. *J Sep Sci,* 37**,** 2039-55.

SZÁNTAY, C., BÉNI, Z., BALOGH, G. & GÁTI, T. 2006. The changing role of NMR spectroscopy in off-line impurity identification: A conceptual view. *TrAC Trends in Analytical Chemistry,* 25**,** 806-820.

TRYGG, J. & WOLD, S. 2002. Orthogonal projections to latent structures (O-PLS). *Journal of chemometrics,* 16**,** 119-128.

VAN DEN BERG, R. A., HOEFSLOOT, H. C., WESTERHUIS, J. A., SMILDE, A. K. & VAN DER WERF, M. J. 2006. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC genomics,* 7**,** 142.

VARDANYAN, R. & HRUBY, V. 2006. *Synthesis of essential drugs*, Elsevier.

VERBEECK, R. K., KANFER, I. & WALKER, R. B. 2006. Generic substitution: the use of medicinal products containing different salts and implications for safety and efficacy. *Eur J Pharm Sci,* 28**,** 1-6.

WADE DEMOND, R. A. K., JAMES L. ITALIEN, DAVID LOKENSGARD, G. WEILERSBACHER, KEITH HERMAN 2000. Orthogonal HPLC methods for quantitating related substances and degradation products of pramlintide. *AAPS PharmSciTech,* 1**,** 50-59.

WAGNER, A. F., WALTON, E., BOXER, G. E., PRUSS, M. P., HOLLY, F. W. & FOLKERS, K. 1956. Properties and derivatives of α-lipoic acid. *Journal of the American Chemical Society,* 78**,** 5079-5081.

WANG, L., WANG, Z., XU, J.-H., BAO, D. & QI, H. 2006. An eco-friendly and sustainable process for enzymatic hydrolysis of penicillin G in cloud point system. *Bioprocess and biosystems engineering,* 29**,** 157-162.

WANG, X., LI, W. & RASMUSSEN, H. T. 2005. Orthogonal method development using hydrophilic interaction chromatography and reversed-phase high-performance liquid chromatography for the determination of pharmaceuticals and impurities. *Journal of Chromatography A,* 1083**,** 58-62.

WARD JR, J. H. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association,* 58**,** 236-244.

WATSON, D. 2012. Pharmaceutical Analysis: A Textbook for Pharmacy Students and Pharmaceutical Chemists. 2012: Elsevier.

WEGMAN, M. A., JANSSEN, M. H., VAN RANTWIJK, F. & SHELDON, R. A. 2001. Towards Biocatalytic Synthesis of β-Lactam Antibiotics. *Advanced Synthesis & Catalysis,* 343**,** 559-576.

WEISS, J. 2004. *Handbook of Ion Chromatography,* Weinheim, WileyVCH.

WEISSENBURGER, H. & VAN DER HOEVEN, M. 1970. An efficient nonenzymatic conversion of benzylpenicillin to 6-aminopenicillanic acid. *Recueil des Travaux Chimiques des Pays-Bas,* 89**,** 1081-1084.

WESTERHUIS, J. A., HOEFSLOOT, H. C., SMIT, S., VIS, D. J., SMILDE, A. K., VAN VELZEN, E. J., VAN DUIJNHOVEN, J. P. & VAN DORSTEN, F. A. 2008. Assessment of PLSDA cross validation. *Metabolomics,* 4**,** 81-89.

WHEELOCK, Å. M. & WHEELOCK, C. E. 2013. Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine. *Molecular BioSystems,* 9**,** 2589-2596.

WHO. 1999. *Counterfeit Drugs* [Online].

WIKLUND, S., JOHANSSON, E., SJÖSTRÖM, L., MELLEROWICZ, E. J., EDLUND, U., SHOCKCOR, J. P., GOTTFRIES, J., MORITZ, T. & TRYGG, J. 2008. Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical chemistry,* 80**,** 115-122.

WIKSTROM, H., MARSAC, P. J. & TAYLOR, L. S. 2005. In-line monitoring of hydrate formation during wet granulation using Raman spectroscopy. *J Pharm Sci,* 94**,** 209-19.

WILIAMS, R. C., BOUCHER, R., BROWN, J., SCULL, J. R., WALKER, J. & PAOLINI, D. 1997. Analysis of acetate counter ion and inorganic impurities in pharmaceutical drug substances by capillary ion electrophoresis with conductivity detection. *Journal of pharmaceutical and biomedical analysis,* 16**,** 469-479.

WILLIAMS, R. C. & BOUCHER, R. J. 2000. Analysis of potassium counter ion and inorganic cation impurities in pharmaceutical drug substance by capillary electrophoresis with conductivity detection. *Journal of pharmaceutical and biomedical analysis,* 22**,** 115-122.

WORLEY, B. & POWERS, R. 2013. Multivariate analysis in metabolomics. *Current Metabolomics,* 1**,** 92-107.

XIAOHUI, F., YI, W. & YIYU, C. 2006. LC/MS fingerprinting of Shenmai injection: a novel approach to quality control of herbal medicines. *J Pharm Biomed Anal,* 40**,** 591-7.

YOSHIDA, T. 2004. Peptide separation by Hydrophilic-Interaction Chromatography: a review. *J Biochem Biophys Methods,* 60**,** 265-80.

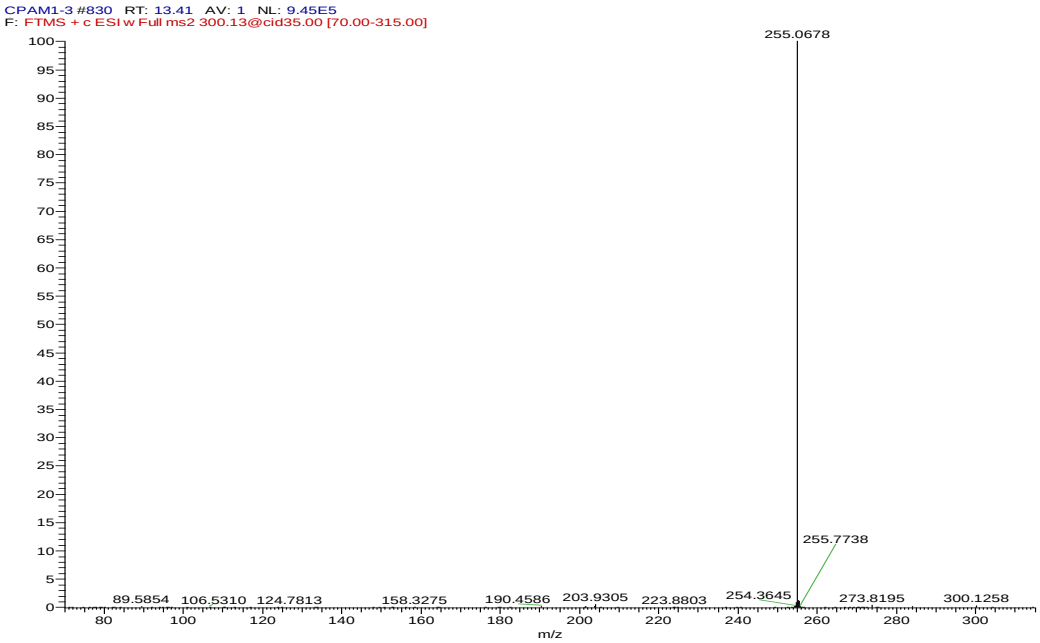
YUAN, Y. Z., ZHANG, M., FAN, X. L., WANG, G. H., HU, C. Q., JIN, S. H., VAN SCHEPDAEL, A., HOOGMARTENS, J. & ADAMS, E. 2012. Impurity profiling of etimicin sulfate by liquid chromatography ion-trap mass spectrometry. *J Pharm Biomed Anal,* 70**,** 212-23.

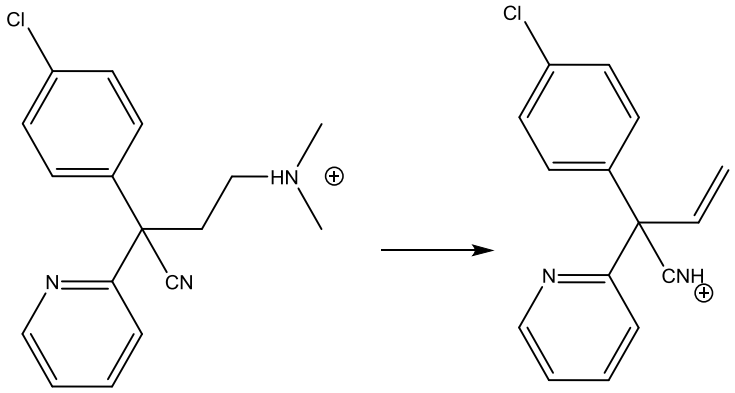
ZHAO, L., RAVAL, V., BRIGGS, N. E., BHARDWAJ, R. M., MCGLONE, T., OSWALD, I. D. & FLORENCE, A. J. 2014. From discovery to scale-up: α-lipoic acid: nicotinamide co-crystals in a continuous oscillatory baffled crystalliser. *CrystEngComm,* 16**,** 5769-5780.

ZHOU, Z. 2016. Non-target impurity profiling of marketplace Cetirizine using high-resolution mass spectrometry and multivariate data analysis. *Rapid Communications in Mass Spectrometry,* 30**,** 1941-1950.

ZUBAREV, R. A. & MAKAROV, A. 2013. Orbitrap mass spectrometry. *Anal Chem,* 85**,** 5288-96.

Appendices for Chapter 4.





$C_{17}H_{19}ClN_3^+$
Exact Mass: 300.1262

$C_{15}H_{12}ClN_2^+$
Exact Mass: 255.0684

Figure S4.1. MS^2 spectra and proposed structure of impurity D at ([M+H]$^+$, m/z 300) and its fragment. The HR-MS/MS data confirmed that impurities m/z 300.126 has elemental composition of $C_{17}H_{19}N_3Cl$. It is corresponded as impurity which generates other fragment ions at m/z 255.07. It results from the neutral loss of (C_2H_5N) from the precursor ion related to fragment ion come from alpha-cleavages. This impurity could be produced by adding a (CN) group to chlorpheniramine.

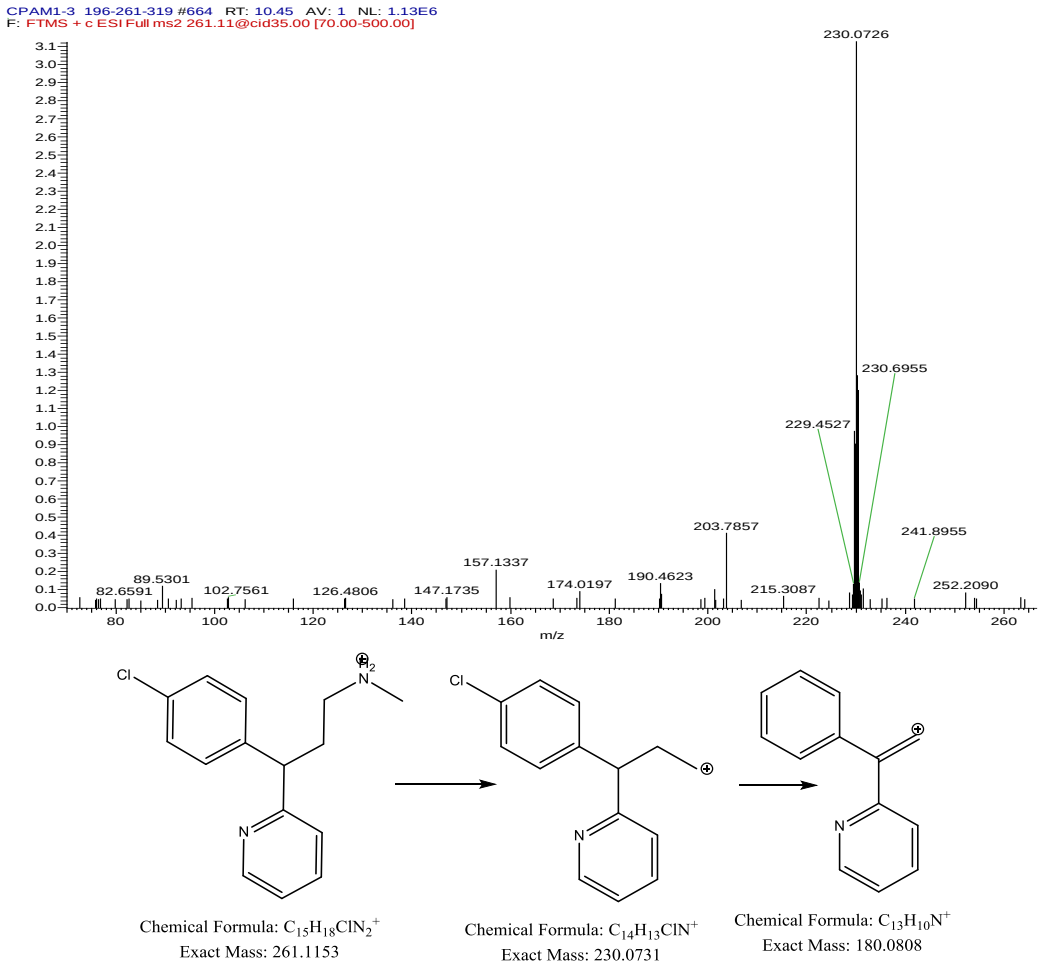


Figure S4.2. MS² spectra and proposed structure for impurity C ([M+H]⁺, m/z 261.1153) and its fragments.

The impurity has protonated ion $[M+H]^+$ 261.115 with elemental composition ($C_{15}H_{18}N_2Cl$) and a mass error less than 0.56 ppm is the impurity C according to the chlorpheniramine maleate monograph (Europe 2017). It is a result of loss of the methyl group of chlorpheniramine. It fragmented to yield product ion at m/z 230.07 correlated to the formula of $C_{14}H_{13}NCl$, as consequence of loss of the methylamine group of impurity C.

Appendices for Chapter 6.

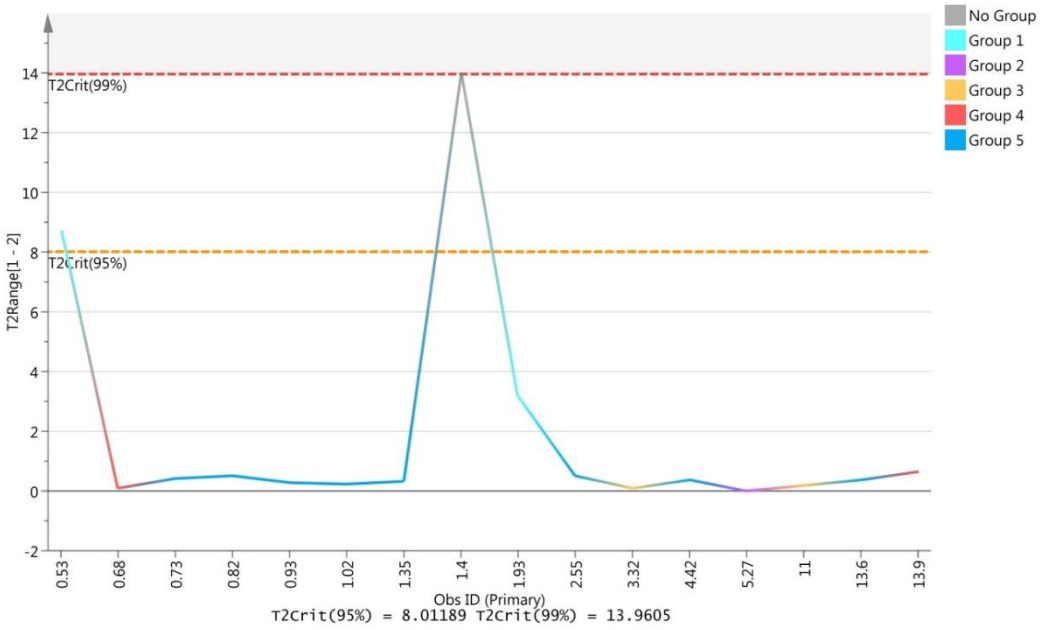
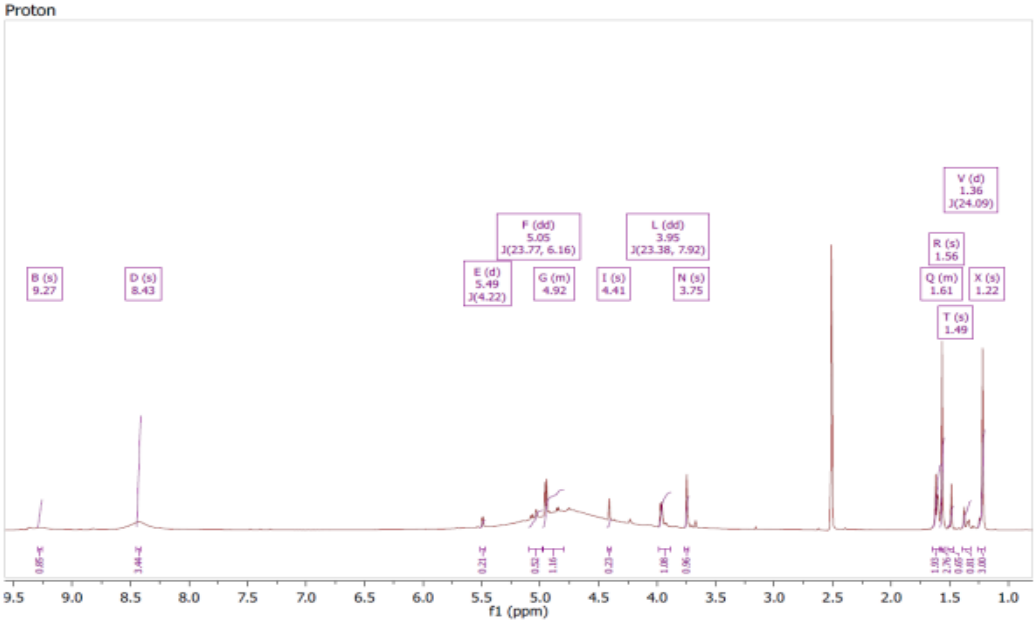
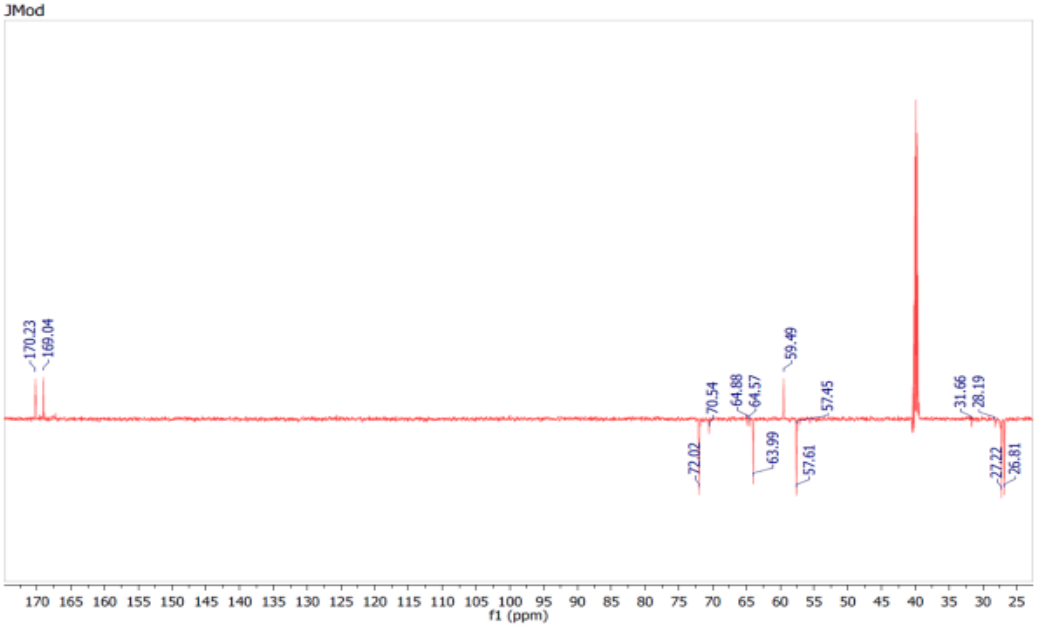


Figure S6.1 the hotelling's T2 plot shows sample with clarity value 1.4 has reached the T2Crit(99%) confidence limit line which considered a serious outlier.

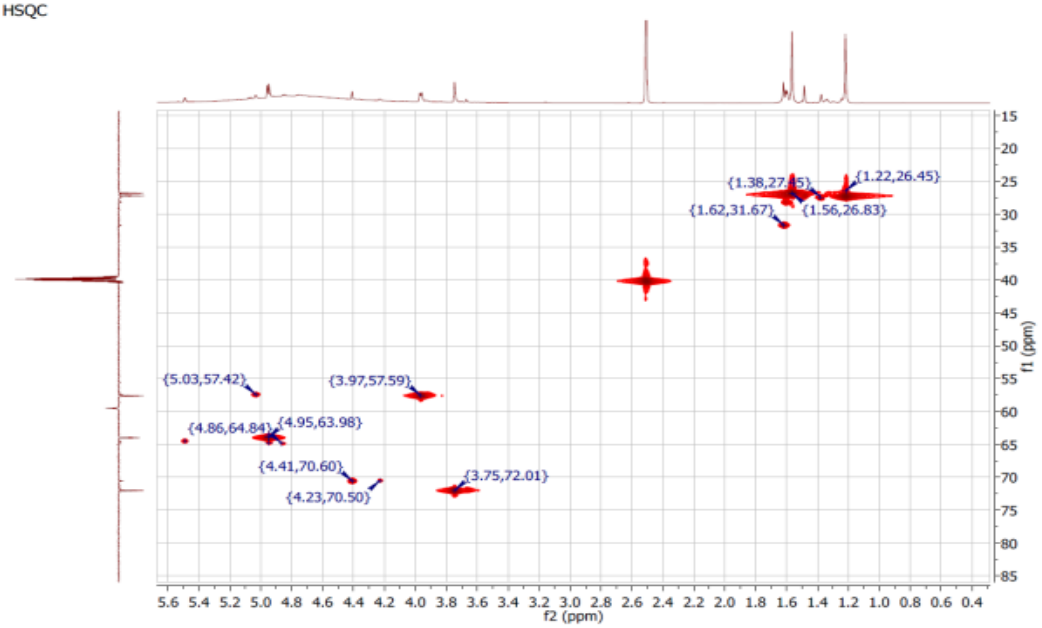
Proton


B (s)
9.27
D (s)
8.43
E (d)
5.49
J(4.22)
F (dd)
5.05
J(23.77, 6.16)
G (m)
4.92
I (s)
4.41
L (dd)
3.95
J(23.38, 7.92)
N (s)
3.75
V (d)
1.36
J(24.09)
R (s)
1.56
Q (m)
1.61
X (s)
1.22
T (s)
1.49
f1 (ppm)

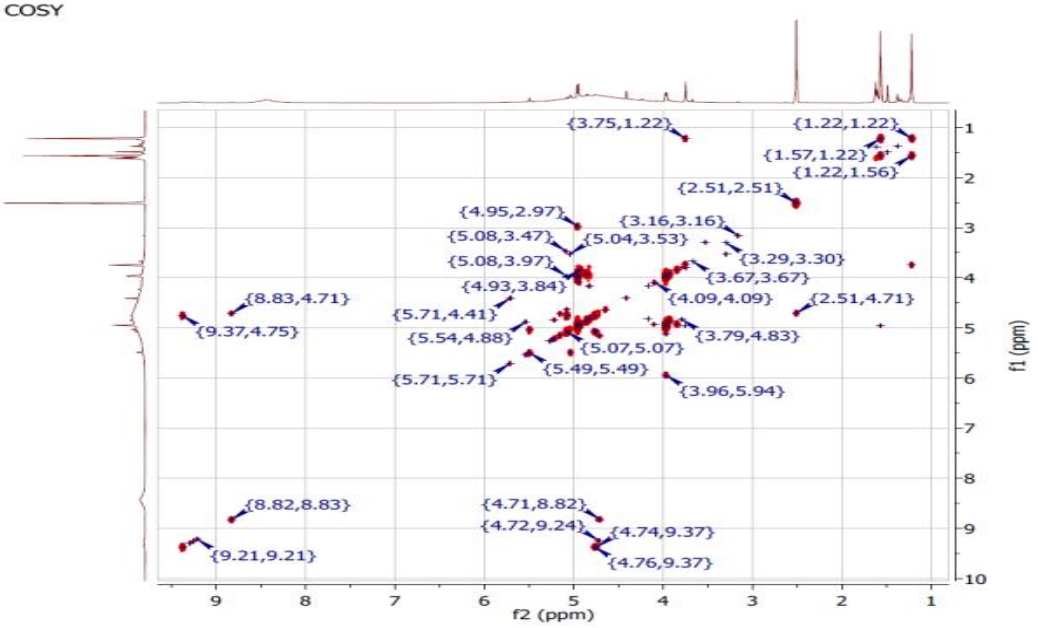
JMod


170.23
169.04
72.02
70.54
64.88
64.57
63.99
59.49
57.61
57.45
31.66
28.19
27.22
26.81
f1 (ppm)

HSQC


{1.38,27.45}
{1.22,26.45}
{1.62,31.67}
{1.56,26.83}
{5.03,57.42}
{3.97,57.59}
{4.86,64.84}
{4.95,63.98}
{4.41,70.60}
{4.23,70.50}
{3.75,72.01}
f2 (ppm)
f1 (ppm)

COSY


{3.75,1.22}
{1.22,1.22}
{1.57,1.22}
{1.22,1.56}
{2.51,2.51}
{4.95,2.97}
{3.16,3.16}
{5.08,3.47}
{5.04,3.53}
{3.29,3.30}
{5.08,3.97}
{3.67,3.67}
{4.93,3.84}
{4.09,4.09}
{2.51,4.71}
{8.83,4.71}
{9.37,4.75}
{5.71,4.41}
{5.54,4.88}
{3.79,4.83}
{5.07,5.07}
{5.49,5.49}
{5.71,5.71}
{3.96,5.94}
{8.82,8.83}
{9.21,9.21}
{4.71,8.82}
{4.72,9.24}
{4.74,9.37}
{4.76,9.37}
f2 (ppm)
f1 (ppm)

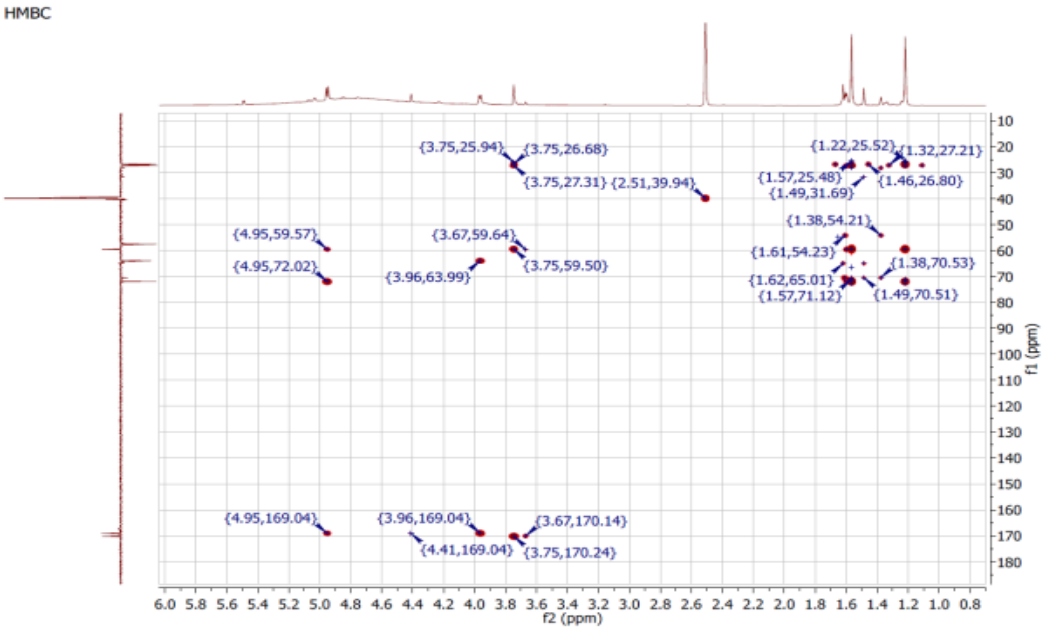


Figure S6.2 NMR spectra of isolated material of filter from very high acid clarity sample including ^{1}H, *JMod, HSQC, COSY and HMBC*.