



Swansea University  
Prifysgol Abertawe



## Swansea University E-Theses

---

# Investigation of biological matrices for novel biomarkers by modern mass spectrometric methods.

Godfrey, Amy Ruth

### How to cite:

---

Godfrey, Amy Ruth (2008) *Investigation of biological matrices for novel biomarkers by modern mass spectrometric methods.* thesis, Swansea University.

<http://cronfa.swan.ac.uk/Record/cronfa42718>

### Use policy:

---

This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence: copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder. Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

Please link to the metadata record in the Swansea University repository, Cronfa (link given in the citation reference above.)

<http://www.swansea.ac.uk/library/researchsupport/ris-support/>

**INVESTIGATION OF BIOLOGICAL MATRICES  
FOR NOVEL BIOMARKERS BY MODERN MASS  
SPECTROMETRIC METHODS.**

**By**

**Amy Ruth Godfrey BSc (Hons)**

**A thesis submitted in fulfilment of the requirements for the Degree of Doctor of  
Philosophy in Swansea University**

ProQuest Number: 10807487

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10807487

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346





## DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed.....(candidate)

Date..... 13 MARCH 2008

## STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references.

Signed.....(candidate)

Date..... 13 MARCH 2008

## STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed.....(candidate)

Date..... 13 MARCH 2008

## SUMMARY

The primary objective was to introduce novel or develop existing techniques for the identification of new biomarkers within a range of biological matrices by modern mass spectrometric methods. Samples interrogated were hemodialysis concentrate, whole tissue sections and whole blood, with each having inherent challenges for use with mass spectrometry. Hence, published research has focused on other biological matrices or modes of detection for achieving the relevant aim. This current work overcame these issues by improving sample preparation including, the use of existing protocols for completely novel applications.

Haemodialysate solution has proved most fruitful for identifying new candidate biomarkers. We have reproducibly detected 15 known and 6 novel uremic solutes within hemodialysate, a biological matrix previously deemed unsuitable for liquid chromatography/electrospray ionisation-mass spectrometry (LC/ESI-MS). This work included a validation of the novel methodology with stability and reproducibility investigations to test robustness. This highlighted a previously unrecorded thermally labile nature of some uremic solutes within the dialysate solution. A putative structural assignment has been made for 4 novel uremic solutes named, 5-(amino-1,2,-dihydroxy-ethyl)-3-nitrosooxy-[1,2,4]trioxine-3,6-diol, 2-(5,6-diamino-6-diazenyl-cyclohex-1-enyl)-2-hydroxy-acetimidic acid, *N*-[2-(7-hydroxy-3-methyl-ocatahydroimidazo[1,5- $\alpha$ ]pyridine-6-yl)-2-oxo-acetyl]-guanidine, and 3-(6-hydroxy-cyclohexa-1,3-dienyl)-2-imino-3-oxopropionaldehyde. We have also identified that the chemical nature of solutes will dictate their removal during dialysis treatment and highly polar conventional biomarkers, urea and creatinine, are not representative of non-polar analyte excretion. This allows us to knowledgeably suggest recommendations to improve future treatment modalities.

The mass spectrometric analysis of whole tissue sections, in particular those that are paraffin embedded, pose a new range of challenges. Current MALDI matrices are unable to penetrate deep within tissue limiting their use to the tissue surface only. We have evaluated a range of novel dansylated MALDI matrices for this purpose that is detectable by fluorescence spectroscopy to aid in locating the matrix compound following application. Each dansylated MALDI matrix showed better penetration into the tissue sections, yet maintaining fluorescence detection, when compared to standard matrices CHCA, sinapinic acid and DHB. Of these novel matrices dansylhydrazine proved most successful in ionising proteins and peptides by forming a protonated molecule and related adducts. These additional mass shifted peaks, when included in a tryptic peptide database search, can improve the probability of the original protein/peptide identification. We now have the potential to obtain a total image of frozen tissue by using CHCA and dansylhydrazine in combination to ionise proteins/peptides at the surface or at depth, respectively. Further work is required for the preparation protocols with paraffin embedded sections for this total imaging principle to be applied.

Finally we have illustrated the advantages of discovering novel haemoglobin variants in blood with a new ion mobility time-of-flight mass spectrometer, the Synapt HDMS system (Waters, MA, USA). We have identified a new variant that co-elutes with glycated haemoglobin peaks present in chromatograms used for conventional blood screening. Ion mobility technology and data extraction enhances the clarity of the results regarding multiple charging and variant characteristics. This enabled the exact determination of the amino acid substitution or mutation for the variant, with its assignment to a haemoglobin chain and the specific location within the chain.

## ACKNOWLEDGEMENTS

*'Success is the ability to go from failure to failure without losing your enthusiasm!'*

Winston Churchill

This is why a PhD will always live up to the promise that you will be trained as a doctor of philosophy regardless of your field. During these challenging times it is the calibre of people around you that are vital in helping you through, and it is therefore part their project in addition to your own. Firstly, I would like to express my gratitude to Dr Edward Browne of GlaxoSmithKlein, whom without wise words of advice during my industrial placement I would not have applied to work with Prof. Gareth ('the Boss') Brenton. Gareth, a very special thank you! Your encouragement and patience has been vital in nurturing my ambition to aim that little bit higher and have a career in academic research. You're a star! Secondly, Prof. Russ Newton, a very knowledgeable supervisor who has given some invaluable advice and pushed me to succeed. Special thanks to our collaborator Dr Peter Willshaw, always willing to answer any questions regarding the haemodialysis work whenever I knocked on his door. A debt of gratitude to Dr Ed Dudley, with whom I worked prior to my PhD. Always ready to help, then and now, and one of the essential people in keeping the BAMS laboratory running, in spite of the many students! To my colleagues in BAMS, EPSRC and School of Medicine, in particular Ange, Sarah, Liz and Ally – you were fundamental in preserving my sanity!

I would also like to thank staff and patients at the Renal Unit and Pathology Department of Morriston and Singleton Hospitals. In particular, credit should be given to Dr Ash Mikhail and Lisa Bastin (Clinical Research Unit), as they were essential in keeping the patients upbeat for supplying me with samples! Further gratitude and appreciation to Prof. Jim Scrivens and Dr Jon Williams of Warwick University and Dr Brian Greene of Waters, for their cooperation and kindness has allowed me to use their very nifty Synapt HDMS mass spectrometer. Also, recognition should be given to our collaborators at ThermoFisher Scientific, Gary Woffendin and Helen Welchman - without your help it would have been unlikely that I'd identified one novel toxin let alone four.

Finally, to my family and friends, your support and encouragement has meant a great deal throughout the whole of my academic career. A special mention to my parents, John and Menna, my brother, Andrew, and my aunty and uncle, Val and Mike; your love and unwavering support in everything I do has given me the confidence and strength to achieve much more than I ever imagined – and it's only the beginning! Last but certainly not least, Anthony. You have been there when things have seemed almost unachievable. You always know exactly what to say and do, and provide clarity and a grin to days when I'm most despondent. Your one in a million love. Cheers!

## ABBREVIATIONS

<b>°C</b>	Degrees centigrade
<b>Å</b>	Angstroms
<b>A</b>	Absorption
<b>ACN</b>	Acetonitrile
<b>AGC</b>	Automatic gain control
<b>AMP</b>	Adenosine monophosphate
<b>APCI</b>	Atmospheric pressure chemical ionisation
<b>API</b>	Atmospheric pressure ionsiation
<b>BSA</b>	Bovine serum albumin
<b>C-8</b>	8-carbon chain
<b>C-18</b>	18-carbon chain
<b>CE</b>	Collision energy
<b>CID</b>	Collision induced dissociation
<b>CKD</b>	Chronic kidney disease
<b>cm</b>	Centimetres
<b>Da</b>	Daltons
<b>dc</b>	Direct current (U)
<b>DC</b>	Dansylcadaverine
<b>DCA</b>	Dansyl-DL- $\alpha$ -aminocaprylic acid
<b>DDA</b>	Data dependent analysis
<b>DH</b>	Dansylhydrazine
<b>DHB</b>	2,5-dihydroxbenzoic acid
<b>DNA</b>	Deoxyribonucleic acid
<b>DUA</b>	11-(dansylamino)undecanoic acid
<b>EI</b>	Electron impact
<b>ESI</b>	Electrospray ionisation
<b>ESRD</b>	End-stage renal disease
<b>eV</b>	Electron volts
<b>F</b>	Fluorescence
<b>FAB</b>	Fast atom bombardment
<b>FT-ICR-MS</b>	Fourier transform ion cyclotron resonance mass spectrometer

<b>g</b>	Gram
<b>GC/MS</b>	Gas chromatography/mass spectrometry
<b>Hb</b>	Haemoglobin
<b>Hb-A0</b>	Normal haemoglobin
<b>HbX</b>	Haemoglobin variant
<b>HCl</b>	Hydrochloric acid
<b>HCOOH</b>	Formic acid
<b>HD</b>	Haemodialysis
<b>HETP</b>	Column height equivalent to one chromatographic theoretical plate
<b>HPLC</b>	High performance liquid chromatography
<b>IC</b>	Internal conversion
<b>ICAT</b>	Isotope coded affinity tag
<b>i.d.</b>	Internal diameter
<b>IEF</b>	Isoelectric focussing
<b>ISC</b>	Intersystem crossing
<b>k'</b>	Capacity factor
<b>K</b>	Chromatographic partition coefficient
<b>kDa</b>	kiloDalton
<b>kV</b>	kilovolts
<b>LC</b>	Liquid chromatography
<b>LC/MS</b>	Liquid chromatography/mass spectrometry
<b>LC/MS/MS</b>	Liquid chromatography/tandem mass spectrometry
<b>LCM</b>	Laser capture microdissection
<b>LDI</b>	Laser desorption ionisation
<b>logP</b>	Partition coefficient
<b>LSIMS</b>	Liquid secondary ion mass spectrometry
<b>m/z</b>	Mass-to-charge ratio
<b>[M+H]<sup>+</sup></b>	Protonated molecule
<b>MALDI</b>	Matrix-assisted laser desorption/ionisation
<b>mg</b>	Milligrams
<b>min.</b>	Minute
<b>MIMS</b>	Matrix-assisted laser desorption/ionisation (MALDI) imaging mass spectrometry

<b>mL</b>	Millilitres
<b>mM</b>	Millimolar
<b>mm</b>	Millimetres
<b>MRM</b>	Multiple reaction monitoring
<b>ms</b>	Milliseconds
<b>MS</b>	Mass spectrometry
<b>MS/MS</b>	Tandem mass spectrometry
<b>MS<sup>n</sup></b>	Repeated tandem mass spectrometry (n = number of repetitions)
<b>MW</b>	Molecular weight
<b>MWCO</b>	Molecular weight cut-off
<b>N</b>	Column efficiency
<b>N<sub>2</sub></b>	Diatomic nitrogen
<b>Nd<sup>+</sup>/YAG</b>	Neodymium-doped yttrium aluminium garnet laser
<b>nm</b>	Nanometre
<b>NMR</b>	Nuclear magnetic resonance
<b>OFN</b>	Oxygen free nitrogen
<b>P</b>	Phosphorescence
<b>PA</b>	Proton affinity
<b>PBS</b>	Phosphate buffered saline
<b>PD</b>	Peritoneal dialysis
<b>pH</b>	Activity or concentration of hydrogen ions [H <sup>+</sup> ] in solution
<b>pI</b>	Isoelectric point
<b>PSD</b>	Post-source decay
<b>Q-ToF</b>	Quadrupole-time-of-flight
<b>R</b>	Resolving power
<b>rf</b>	Radio frequency (V)
<b>mRNA</b>	Messenger ribonucleic acid
<b>RP</b>	Reverse-phase
<b>RP-HPLC</b>	Reverse-phase high performance liquid chromatography
<b>rpm</b>	Revolutions per minute
<b>RT or t<sub>R</sub></b>	Retention time

<b>RRT</b>	Relative retention time
<b>S<sub>0</sub></b>	Ground singlet state
<b>S<sub>1</sub></b>	Excited singlet state
<b>SA</b>	Sinapinic acid
<b>SCX</b>	Strong cation exchange column
<b>SDS</b>	Sodium dodecyl sulphate
<b>SDS PAGE</b>	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
<b>SEC</b>	Size exclusion chromatography
<b>SIM</b>	Single ion monitoring
<b>SNP</b>	Single nucleotide polymorphism
<b>SOPs</b>	Standard operating procedures
<b>SPE</b>	Solid phase extraction
<b>SRIG</b>	Stacked ring ion guide
<b>T<sub>0</sub></b>	Ground triplet state
<b>T<sub>1</sub></b>	Excited triplet state
<b>TFA</b>	Trifluoroacetic acid
<b>Th</b>	Thomson
<b>ToF</b>	Time-of-flight
<b>ToF-ToF</b>	Tandem time-of-flight
<b>Tris-HCl</b>	Tris(hydroxymethyl)aminomethane acidified buffer
<b>TWIG</b>	Travelling wave ion guide
<b>UV</b>	Ultraviolet radiation
<b>V</b>	Volts
<b>Vm<sup>-1</sup></b>	Volts per metre
<b>v/v</b>	Volume per volume
<b>WHO</b>	World health organisation
<b>w/v</b>	Weight per volume
<b>α</b>	Relative retention or selectivity or separation factor
<b>CHCA</b>	α-cyano-4-hydroxycinnamic acid
<b>3-HPA</b>	3-hydroxypicolinic
<b>μg</b>	Micrograms
<b>μL</b>	Microlitres
<b>μA</b>	Microamps



# CONTENTS

<b>Declaration</b>	Page I
<b>Summary</b>	II
<b>Acknowledgements</b>	III-IV
<b>Abbreviations</b>	V-VIII

## **1. CHAPTER 1 - Introduction to Haemodialysis and Biomarker Investigations**

1.1 <u>Haemodialysis</u>	Page 2
1.2 <u>Peritoneal Dialysis</u>	4
1.3 <u>Haemodialysis Biomarkers</u>	4
1.4 <u>Stages of Biomarker Development</u>	6
1.4.1 Understanding the origin and process of the disease	6
1.4.2 Biomarker brief: what information does it offer?	6
1.4.3 Biological matrices used for sampling	7
1.4.4 Strategies for determining the target biomarker	7
1.4.5 The patient population and sample groups	8
1.4.6 The analytical methodology	8
1.4.7 Requirement of feasibility studies and biomarker screening	8
1.4.8 Qualitative investigations to identify candidate biomarkers	9
1.4.9 Development of clinical assay for initial diagnosis	9
1.4.10 Identification of other possible applications of the biomarker	9
1.4.11 Application of the identified biomarker and in evaluating the clinical data	10
1.5. <u>Conventional Analytical Techniques for Biomarker Investigations</u>	10
1.5.1. Surface enhanced laser desorption/ionisation	10
1.5.2. Microarray technology	11
1.5.3. Protein arrays	12
1.5.4. Tissue microarrays	12
1.5.5. 2-D gel electrophoresis	12
1.5.6. ICAT techniques	14
1.5.7. Single nucleotide polymorphism techniques	15
1.5.8. LC/MS	15

1.6. <u>Biological Matrix Selection, Preparation and Analysis Techniques</u>	Page 15
1.7. <u>Separation Science</u>	17
1.7.1 Solid phase extraction	17
1.7.2. Size exclusion chromatography	19
1.7.3. Liquid-liquid chromatography	20
1.7.4. Ion-exchange chromatography	21
1.7.5. Reverse phase liquid chromatography	21
1.7.5.1 Partition coefficient	22
1.7.5.2 Resolution	23
1.7.5.2.1 Capacity factor	25
1.7.5.2.2 Relative retention	26
1.7.5.2.3 Column efficiency	26
1.8 <u>Mass Spectrometric Analysis</u>	28
1.8.1. Origins of mass spectrometry	28
1.8.2. An introduction to mass spectrometry	29
1.8.3. Sample ionisation and modern interfaces	31
1.8.3.1. Electrospray ionisation	34
1.8.3.2. Matrix-assisted laser desorption/ionisation (MALDI)	36
1.8.4. Ion separation and analysis	37
1.8.4.1 Quadrupole mass filter	39
1.8.4.2 Quadrupole ion trap mass analysers	43
1.8.4.3 Time-of-flight mass analysers	45
1.8.4.4 Orbitrap mass analyser	47
1.9. <u>Project Brief and Hypothesis</u>	48
References	49 - 51
<b>2. <u>Chapter 2 – Materials and Instrumentation</u></b>	
2.1 <u>Chemicals</u>	52
2.2 <u>Instrumentation</u>	
2.2.1 Chromatographic equipment	53
2.2.1.1 UV screening crude pooled samples	53
2.2.1.2 Size exclusion	53
2.2.1.3 Reverse phase HPLC	53

2.2.2 Chromatographic conditions	Page 53
2.2.2.1 Size exclusion	53
2.2.2.2 Reverse phase HPLC	54
2.2.2.2.1 Injection programme	55
2.2.3 Mass spectrometers	55
2.2.3.1. Ion trap	55
2.2.3.2. Orbitrap	55
2.2.3.3. Voyager DE-STR MALDI-ToF	56
2.2.4. Mass spectrometric conditions	57
2.2.4.1. Ion trap methods	57
2.2.4.1.1. Full mass scan	57
2.2.4.1.2. Fragmentation MS/MS methods	58
2.2.4.1.3. Data dependent MS <sup>n</sup>	59
2.2.4.2. Orbitrap methods	60
2.2.4.2.1. Full mass and data dependent scan: Uremic analytes (candidate biomarkers) 1, 2 and 3	60
2.2.4.2.2. Full mass and data dependent scan: Uremic analytes (candidate biomarkers) 4, 5 and 6	61
2.2.4.2.3. Mass Targeted Fragmentation Scan: Uremic analytes (candidate biomarkers) 1, 2, and 3	62
2.2.4.2.4. Mass Targeted Fragmentation Scan: Uremic analytes (candidate biomarkers) 4, 5, 6	63
2.2.4.3. Voyager DE-STR MALDI-ToF methods	64
2.2.4.3.1. Mass Range 50-2000Da	64
2.2.4.3.2. Mass Range 1k-7kDa	64
2.2.4.3.3. Mass Range 2k-20kDa	65
2.2.4.3.4. Mass Range 5k-100kDa	65
2.2.5. Lab Apparatus	66
2.3. <u>Solutions</u>	66
2.3.1. SEC test solutions	66
2.3.1.1 $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA)	66
2.3.1.2 Adenosine monophosphate (AMP)	67
2.3.1.3 Angiotensin I	67
2.3.2 Known uremic toxin standard solutions	67
<b>3. <u>Chapter 3 – Validation and Stability Investigations of Novel Methodology for Screening Hemodialysate Using LC/MS</u></b>	
3.1. <u>Introduction</u>	68

3.2. <u>Method Validation 1: Validation of Analytical Protocols for Crude Pooled Samples</u>	Page 69
3.2.1. UV analysis of crude sample	69
3.2.2. Size exclusion chromatography	70
3.2.2.1. UV absorbance	70
3.2.2.2. pH and conductivity data	79
3.2.2. Reverse phase chromatography	84
3.2.3.1. Stability Investigations	85
3.2.3.2. Reproducibility study of the reverse phase chromatography	87
3.3. <u>Method Validation 2: Identification of known uremic toxins</u>	90
3.3.1. Creatinine	91
3.3.2. Hypoxanthine	92
3.3.3. Pseudouridine	94
3.3.4. Advanced Glycation End Products	98
3.3.5. Dimethylglycine	102
3.3.6. Indole-3-acetic acid	104
3.3.7. $\beta$ -guanidinopropionic acid	107
3.3.8. N-acetylarginine	109
3.3.9. Hippuric acid	112
3.3.10. N-acetyltryptophan	115
3.3.11. Uric acid	117
3.3.12. <i>Myo</i> -inositol	120
3.3.13. N <sup>6</sup> -methyladenosine	123
3.3.14. Xanthosine	125
3.3.15. Hydroquinone	128
3.4. <u>Method Validation Summary</u>	130
References	134 - 137
<b>4. <u>Chapter 4 – Identification and Evaluation of Novel Uremic Analytes for Assessing Haemodialysis Adequacy in Patient Specific Samples</u></b>	
4.1. <u>Introduction</u>	138
4.2. <u>Comparative LC/MS and LC/MS/MS Investigations of Pooled Hemodialysate</u>	139

4.2.1. Selecting candidate biomarkers	Page 139
4.2.2. UV characterisation of novel candidate biomarkers	140
4.2.3. Advanced mass spectrometric structural elucidation investigations	141
4.2.3.1. Accurate mass	141
4.2.3.2. Structure investigations using MS/MS	142
4.2.4. Discussion of results	145
4.2.4.1. Candidate biomarker 1: <i>m/z</i> 241.0296	145
4.2.4.2. Candidate biomarker 2: <i>m/z</i> 214.1298	151
4.2.4.3. Candidate biomarker 3: <i>m/z</i> 275.0478	157
4.2.4.4. Candidate biomarker 4: <i>m/z</i> 270.1566	161
4.2.4.5. Candidate biomarker 5: <i>m/z</i> 180.0650	167
4.2.4.6. Candidate biomarker 6: <i>m/z</i> 335.0545	173
4.3. <u>Excretion Kinetics of Urea and Creatinine with Novel Uremic Biomarkers</u>	177
4.3.1. Excretion during haemodialysis	179
4.3.1.1. Urea	179
4.3.1.2. Creatinine	181
4.3.1.3. Candidate Biomarker 1: <i>m/z</i> 241.0296	183
4.3.1.4. Candidate Biomarker 2: <i>m/z</i> 214.1298	185
4.3.1.5. Candidate Biomarker 3: <i>m/z</i> 275.0478	187
4.3.1.6. Candidate Biomarker 4: <i>m/z</i> 270.1566	190
4.3.1.7. Candidate Biomarker 5: <i>m/z</i> 180.0650	192
4.3.1.8. Candidate Biomarker 6: <i>m/z</i> 335.0545	194
4.3.2. Suitability of use as biomarker(s)	196
4.3.3. Effect of exercise on dialysis excretion and its future in treatment	197
4.4. <u>Summary of Identifying New Uremic Analytes and Their Application</u>	197
References	200
<b>5. <u>Chapter 5 – Characterisation of Novel MALDI matrices for MALDI Tissue Imaging</u></b>	
5.1. <u>Introducing MALDI Imaging Mass Spectrometry</u>	201

5.1.1. Tissue preparation	Page 202
5.1.1.1. Tissue blotting on polymer membranes	202
5.1.1.2. Tissue imaging using laser capture micordissection (LCM)	203
5.1.1.3. Direct MALDI Imaging	203
5.1.2. Tissue analysis by MALDI mass spectrometry	205
5.1.2.1. Matrix-assisted laser desorption/ionisation time-of-flight/mass spectrometry (MALDI- ToF-MS)	206
5.1.2.2. Fundamentals of MALDI matrices for identifying new matrices for imaging mass spectrometry	208
5.2. <u>Investigating Novel Matrices for Tissue Imaging</u>	212
5.2.1. Discussion of novel matrices	213
5.2.1.1. Fluorescence microscopic study of novel matrices	215
5.2.1.2. Application of novel compounds as MALDI matrices	217
5.2.2. Investigating sample preparation protocols of novel matrices	218
5.2.2.1. Practical considerations of solvent composition	218
5.2.2.2. Testing novel matrices for protein and peptide analysis	219
5.2.2.2.1. Protein and peptide mixes	219
5.2.2.2.2. Solvent composition	224
5.2.2.2.3. Matrix concentration	226
5.2.2.2.4. Ionisation mode for novel matrices	227
5.2.2.2.5. Investigation of optimum sample preparation conditions for novel matrices	227
5.2.2.2.5.1. Fluorescent microscopic results	227
5.2.2.2.5.2. Mass spectrometric suitability of novel matrices	234
5.3 <u>Comparative studies of Novel and Current Matrices for Protein and Peptide Analysis</u>	236
5.4 <u>Application of novel matrices onto tissue surfaces</u>	239
5.4.1 Preparation protocol for tissue sections	239
5.4.1.1 Paraffin sections	239
5.4.1.2 De-paraffinised sections	240
5.4.1.3 Frozen sections	240
5.4.2 Mass spectrometric results	241
5.4.2.1 Frozen tissue sections	241
5.4.2.1.1 Dried droplet	241
5.4.2.1.2 Matrix layering	242

5.4.2.1.3	Drying temperature effects on fresh tissue sections	Page 247
5.4.2.1.4	'Sandwich' preparative method using conventional matrices	247
5.4.2.2	Paraffinised and de-paraffinised tissue sections	252
5.4.2.2.1	"5X" AntigenPlus retrieval buffers	253
5.4.2.2.1.1	Preparation of pH 6 and 10	253
5.4.2.2.1.2	Preparation of pH 7.4	253
5.4.2.2.1.3	Results	254
5.4.2.2.2	Comparison of immunohistochemical testing and mass spectrometric data	254
5.4.2.2.2.1	Results: testing for CD-45 protein	255
5.4.2.2.2.2	Results: testing for Progesterone receptor (precursor)	256
5.4.2.2.2.3	Results: testing for Calcitonin (precursor)	258
5.4.2.2.2.4	Results: testing for S100 (precursor)	260
5.4.2.2.2.5	Discussion: immunohistochemical vs mass spectral imaging	262
5.5.	<u>Conclusions</u>	263
5.5.1.	Future Work	264
References		266-268

**6. Chapter 6 – Identification of a Novel Haemoglobin Variant by Ion Mobility Spectrometry Coupled to Time-Of-Flight Mass Spectrometry**

6.1.	<u>Introduction</u>	269
6.1.1.	Current preparative methods for identifying haemoglobin variants	271
6.1.1.1.	Electrophoresis	271
6.1.1.2.	High performance liquid chromatography (HPLC)	271
6.1.1.3.	DNA Sequencing	272
6.1.1.4.	Liquid chromatography-mass spectrometry (LC-MS)	273
6.2.	<u>The Analytical Problem!</u>	273
6.3.	<u>Application of Novel Analytical Mass Spectrometric Methodology for Identifying Novel Variant</u>	275
6.3.1.	Sample preparation	275
6.3.1.1.	Intact haemoglobin chains	275

6.3.1.2. Tryptic digestion	Page 276
6.3.2. Ion mobility mass spectrometry analysis	276
6.3.2.1. Ion mobility mass spectrometry	276
6.3.2.1.1. Fundamentals of ion mobility	277
6.3.2.1.2. Degree of ion mobility separation or resolution	280
6.3.2.1.3. Sensitivity in ion mobility experiments	283
6.3.2.1.4. Instrumental developments in ion mobility mass spectrometry	284
6.3.2.2. Instrument parameters	290
6.3.2.2.1. Synapt HDMS mass spectrometry system	290
6.3.2.3. Identification of a novel haemoglobin variant	291
6.3.2.3.1. Analysis of intact haemoglobin chains – discovering a mutant	292
6.3.2.3.2. Analysis of tryptic haemoglobin peptides	296
6.3.2.3.3. Pre-ion mobility MS/MS investigations of tryptic peptides	299
6.4. <u>Conclusion</u>	301
References	303-304
<b>7. <u>Summary: Investigation of Clinically Significant Biomolecules in Various Biomatrices.</u></b>	305-308
References	309
<b><u>Appendices</u></b>	See attached CD
<b>Appendix 1</b>	
<b>Appendix 3</b>	
<b>Appendix 4</b>	



## **CHAPTER 1:**

### **Introduction to Haemodialysis and Biomarker Investigations**

The current obesity epidemic throughout the western world has resulted in a considerable increase in the condition of Type II diabetes mellitus. Recently, the World Health Organisation (WHO) has predicted that global prevalence of Type II will rise from 175 million patients in 2003 to over 350 million by 2030. One of the major consequences of this disorder is renal failure in the form of chronic kidney disease (CKD), which eventually progresses to end-stage renal disease (ESRD). The global population with ESRD has already surpassed one million and is expected, in the United States alone, to have an incidence rate of over 400,000 a year by 2030<sup>[1]</sup>. The clinical condition resulting from renal failure is known as 'Uremia' or the 'Uremic Syndrome', and literally translates to 'urine in the blood'<sup>[2]</sup>. It is characterised by a toxic state that is attributed to the retention and accumulation of solutes usually excreted by the healthy kidney and their affect on the normal biochemical pathways. Once diagnosed with renal insufficiency or uremia the patient has two options for survival, a continuous treatment of dialysis, or a kidney replacement from a suitable donor. However, most patients are treated using dialysis since there is currently a shortage of donor kidneys. The WHO predicts that there will be an increase in the requirement of dialysis treatment with which the economic impact will be staggering. In 2001 the United States spent over \$22.83 billion on their ESRD program which only included immediate health care costs and excluded any long term expenditure<sup>[1]</sup>. This impending burden on the international economy has created a global interest in improving existing renal treatments, and to identify the underlying physiological mechanism specific to the type of renal disease.

Dialysis is the artificial process by which bodily fluids, such as blood, are filtered to remove toxic waste and excess volumes of water. A major disadvantage of dialysis is that it is always unsuccessful in curing the patient of their renal dysfunction as it is not capable of reproducing normal renal function in its entirety. The glomerulus of a conventional healthy kidney is capable of clearing solutes of up to approximately 58,000Da in molecular weight, while solutes of a higher mass are metabolized by the tubular system of the kidney by secretory and reabsorptive processes<sup>[3]</sup>. At present neither forms of dialysis can accomplish this complete kidney function and hence, treatment of patients with compromised renal function, based solely on dialysis, will eventually result in a reduced life expectancy. Dialysis can be sectioned into two techniques, haemodialysis (HD), and peritoneal dialysis (PD). The primary characteristic that differentiates between these two forms of dialysis is that HD utilises an externally situated membrane for filtration, whilst PD uses an existing membrane present within the body.

### 1.1 Haemodialysis

Haemodialysis involves the passage of the patient's blood through a dialysis machine in which it is filtered to remove any toxins or excess water. The 'clean' dialysed blood is then returned to the patient via the machine. The dialysis machine therefore, has three main functions:

1. to pump blood and monitor blood flow,
2. to remove toxic substances and waste from the blood, and
3. to monitor blood pressure and rate of fluid clearance via the body.

The actual haemodialysis unit or 'dialyzer' is a large rectangular apparatus that contains thousands of small fibres of dialysis membrane. Conventional membranes

are cellulose based with a molecular weight cut-off point significantly lower than the capability of a healthy kidney. The development of high-flux membranes do offer a higher cut-off range but are only available at considerable cost. Both the dialysis fluid and the blood (via the patient) are directed into these fibres by separate inlets and filtration occurs. The waste fluid is then removed from the dialyzer and the freshly filtered blood passes back into the patient<sup>[4]</sup>. As with Peritoneal dialysis, the dialysis solution must be selected according to the patient's requirements and can be altered depending on their state.

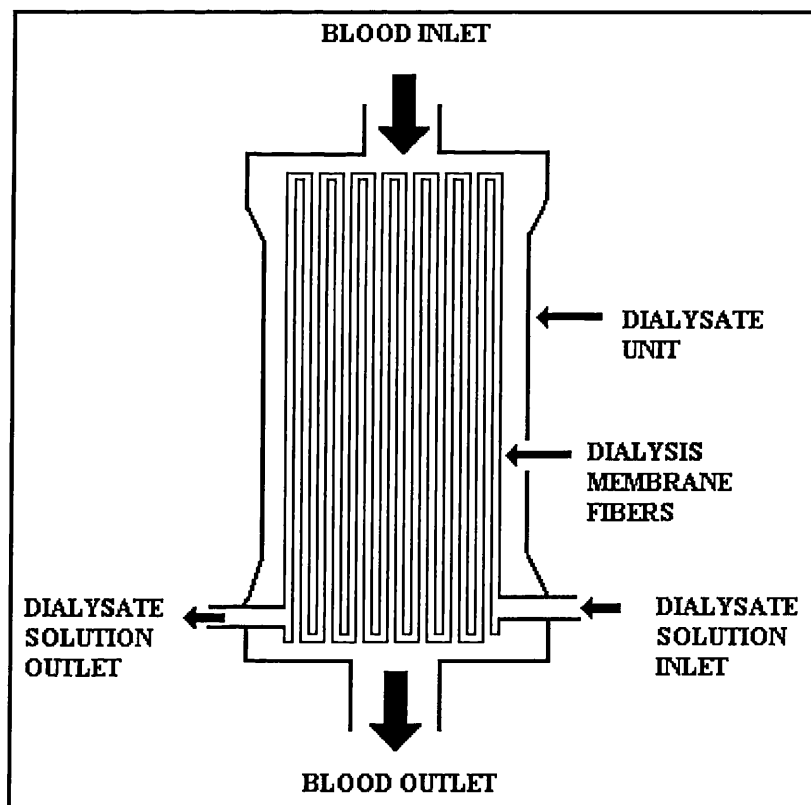


Figure 1.1: Schematic of the internal structure of a haemodialysis unit showing the extensive network of membrane fibers ensuring the ultrafiltration process.

## 1.2 Peritoneal Dialysis (PD)

This involves the use of the semi-permeable Peritoneal membrane to filter waste products and excess volumes of fluid via the blood. For example, a PD catheter is inserted into the abdomen of the patient through which dialysis fluid can flow into the peritoneal cavity. A small length of this tube is exposed out of the abdominal wall to which bags of dialysis fluid are attached. To optimise the PD technique for patient requirements, the constituents of the initial dialysis fluid can be altered. For example, a 'strong' PD fluid will contain high levels of glucose and can thus, remove more water from the blood than a 'weaker' fluid<sup>[4]</sup>. PD initially appears advantageous as it is a mobile technique with the apparatus carried by the patient. However, patients that have experienced major abdominal surgery may have some scarring of the peritoneal membrane, and can render it ineffective for dialysis. Other disadvantages include a lesser capability to filter toxins than a healthy kidney, and considerable pressure and responsibility on the patient to ensure treatment is successful.

## 1.3 Haemodialysis Biomarkers

The contents of the dialysate solution (after dialysis) should be relatively similar regardless of which technique is used. According to Vanholder and Ringoir<sup>[5]</sup> the solution is expected to contain a wide range of solutes that may or may not have potential toxicity. These dialysate substrates can be segregated according to size and if they are protein bound or non-protein bound. During the period of 1968 to 2002 publications have unveiled over 90 possible toxins associated with renal insufficiency, of which 68 are considered as low molecular weight toxins (<500Da), 12 have a molecular weight greater than 12,000Da, and 10 are middle molecular weight toxins

of between 500 and 12,000Da. Approximately a quarter of these are protein-bound and most of these particular toxins are low molecular weight solutes. Many of these solutes (shown in Appendix 1) may contribute to the elevated level of toxicity associated with uremia, and a significant number of these and their potential toxicities have been investigated.

One major function of identifying a toxin that is consistently elevated during renal dysfunction is that it be used as an indicator of the effectiveness of the dialysis process. It is important to distinguish between what is characterised as uremic toxins and biomarkers. Essentially a biomarker is indicative of the result of the toxic effect and may not necessarily be a toxin. This study noted that there were several insufficiencies with dialysis, one of which was that the main molecular biomarker used to measure the efficiency of dialysis, urea (in addition to creatinine on occasions), was realised to be unsuitable. The authors therefore provided a list of all known dialysate solutes (shown in Appendix 1) with potential toxicity and a comprehensive evaluation of a number of possible replacements for urea. Biomarkers have been succinctly described as *'a biologic characteristic that is measured and evaluated objectively as an indicator of normal biologic processes, pathogenic processes, or pharmacologic response to therapeutic intervention'*<sup>[6]</sup>. The most successful biomarkers can achieve highly accurate results by performing a relatively non-invasive sampling and simple analysis procedures. The 'ideal' biomarker for

assessing dialysis adequacy should therefore have the following characteristics<sup>[5]</sup>:

1. It should be retained in renal failure.
2. It should be eliminated by dialysis.
3. Have a proven toxicity.
4. Its generation and elimination should be representative for other (preferably toxic) solutes.
5. The concentration of the solute should be related to the clinical outcome.
6. It should be easily determined.

#### **1.4. Stages of Biomarker Development**

Throughout the process of biomarker development it is necessary that both the assay performance and diagnostic efficacy are validated, and should involve the following stages<sup>[6]</sup>:

##### **1.4.1. Understanding the origin and process of the disease**

If correct and reliable this information can save significant resources and time since it limits the probability of any mis-interpretation of results. Related diseases can offer a great source of additional information as they may share similar physiological pathways that can indicate a possible mechanism of the disease.

##### **1.4.2. Biomarker brief: what information does it offer?**

It is vital that the significance and limitations of the data gained from measuring such a biomarker are understood as it may be used to represent several different facets of clinical medicine. For example, biomarkers can aid in the characterisation of initial

detection or progression of the disease, and its severity after diagnosis before or after the application of relevant treatment.

#### 1.4.3. Biological matrices used for sampling

A successful biomarker assay should, if possible, involve a relatively non-invasive sampling procedure. Common sampling matrices are urine, plasma, serum and blood, although the analysis of the latter three is susceptible to interferences due to the presence of albumin. Since conventional haemodialysis membranes have a molecular weight cut-off point of approximately 10,000 Da, it is unlikely that it will be contaminated with albumin, making hemodialysate an advantageous sample to use. Generally the selection of matrix type will be a compromise between ease of sample collection and analysis, clinical significance, and stability versus specificity to renal ailments.

#### 1.4.4. Strategies for determining the target biomarker

This stage can be affected by several factors including, the type of disease to be monitored and its subclasses, the sample matrix, and the functions of the biomarker. The choice of sample matrix in particular is vital to the success of the eventual assay. It must be evaluated whether it is more beneficial for the performance of the assay to screen a sample of diseased tissue that is more likely to yield potential biomarkers, or to use a biological fluid that can be sampled more easily. A major disadvantage of using tissue samples is that some systematic responses to the disease may be missed, later resulting in an artificial negative result<sup>[7]</sup>. Secondly, the origin of these samples must be decided. Animal samples are generally easy to obtain while human samples offer greater accuracy of the final outcome as they are more representative of the

required result. Consideration should also be given as to which type of samples should be used in the analysis comparison. For example, are the circumstances of the investigation suited to the comparison of several different groups of samples or is it more appropriate to compare an individual diseased sample versus a normal 'healthy' sample. The latter tends to be the most common approach and can be very successful if the assay is to determine the presence or absence of disease.

#### 1.4.5. The patient population and sample groups

It is essential that the samples collected are obtained from meticulously considered sources so sample integrity and validity remains at a high standard. The selection of the sample population should involve phenotypic investigations of each patient with specific notes made regarding patient age, sex, race and current prescribed medicines.

#### 1.4.6. The analytical methodology

There are many analytical techniques suitable for both qualitative and quantitative studies of biological matrices. The eventual technique should be chosen with consideration given to factors such as the type of sample and any associated interferences, the number of samples expected to be analysed, and the object of the investigation. Further information regarding this topic will be discussed later in this chapter.

#### 1.4.7. Requirement and feasibility studies for biomarker screening

Prior to the commencement of the analysis a review must be completed to assess if the biomarker is still necessary and if the development plan is still feasible. In addition to these issues a statistical analysis should be performed to calculate the number of



patients required for the study validation and the clinical trial, in order to provide sufficient clinical information to derive meaningful conclusions.

#### 1.4.8. Qualitative investigations to identify candidate biomarkers

Following the initial screening experiments an ideal result would include a minimum of 50 candidate biomarkers, which must then be ordered according to their potential and importance. This would generally involve a 'rational design approach,' using the clinical characteristics of the investigated disease as a major factor in prioritising the candidate biomarker. An evaluation can then be performed to confirm the suitability of the possible biomarker and its sampling matrix.

#### 1.4.9. Development of clinical assay for initial diagnosis

The previously designed clinical assay must be optimised so that it is robust and can function with high reproducibility within different laboratories. Standard operating procedures (SOPs) must be devised and adhered to ensure the high standard of reproducible results that are rigorously evaluated by a set of quality control procedures. In addition to method optimisation, parameters such as analyte thermal stability in the biological matrix, minimum analysis volume and the lowest limit of detection should be deduced. On completion the assay can be subjected to conventional clinical trials with sufficient sample size to identify any influential factors such as age, race or gender.

#### 1.4.10. Identification of other possible applications of the biomarker

Upon confirmation of the biomarker being truly representative of the presence of the disease, its suitability of detecting it at the earliest stages must be investigated. This

can be accomplished by comparing samples from patients with a positive diagnosis for the disease and healthy controls of the same age. This study is intended to give some indication of any change in biomarker characteristics over time and if it follows the natural progression of the disease. Additional screening studies can determine important issues such as the stage in which the disease is detected, the incidence of the disease and the specificity of the clinical assay.

#### 1.4.11. Application of the identified biomarker and in evaluating the clinical data.

The intricate nature of pathophysiological mechanisms within the body has meant that it is unlikely that one biomarker alone will provide adequate specific and sensitive information regarding the target disease. Therefore the aim of using a single biomarker should be to provide additional information to existing clinical data and result in a more accurate diagnosis than the biomarker data alone.

### 1.5. Conventional Analytical Techniques for Biomarker Investigations

There are many techniques employed for biomarker investigations, and each has advantages and disadvantages. Ultimately, those that are eventually employed should provide the most accurate result in the minimum analysis time for the type of biological matrices chosen.

#### 1.5.1. Surface-enhanced laser desorption/ionisation (SELDI)

This is a relatively new technique that has become one of several frequently used analytical procedures for protein mapping or protein related biomarker experiments.

It essentially utilises stainless steel or aluminium-based supports, or chips that have a

variety of chemical or biological surfaces. Chemically adhered surfaces include hydrophilic, hydrophobic, pre-activated, normal-phase, immobilized metal affinity, and cationic or anionic groups, while biologically based surfaces can range from antibody, to antigen binding fragments, DNA, enzyme, or receptor type groups. A wide selection of binding agents can enable the differential capture of proteins based on the intrinsic properties of the proteins themselves. A series of washes are applied to remove non-specifically or weakly bound proteins and is then followed by the ionisation of the bound proteins by a laser for analysis by mass spectrometry. The major advantage of this technique is the ability to detect low molecular weight peptides and proteins that can be missed by other techniques synonymous with protein analysis such as 2-D gel electrophoresis. Although, the reliability of its peptide pattern results when comparing a 'healthy' to a diseased state is still surrounded by some controversy<sup>[8, 9]</sup>.

#### 1.5.2. Microarray technology

This is generally a screening procedure used for detecting abundance changes of mRNA within a large number of samples. It essentially involves spotting a series of DNA target sequences onto a 'carrier' in the form of a glass slide, silica chip or membrane. These sequences are used to hybridize a selection of nucleic acid probes associated with the diseased tissue such as cancerous tumour. There are several disadvantages regarding the use of microarray techniques although the following have a particular importance for biomarker investigations. Microarrays measure the mRNA and not actual protein abundance. Hence, the time point at which the analysis is carried out is critical to the result as changes in RNA transcription are both rapid and occur for only a brief period of time. Therefore, the most common issue

encountered with this technique is the challenging acquisition of sufficient high quality samples at the early stages of the disease to develop a biomarker for diagnosing this stage<sup>[10]</sup>.

### 1.5.3. Protein arrays

This is essentially a development of the microarray technique where in place of DNA target sequences it utilises proteins. Conventional protein arrays print antibodies onto the array to hybridize the target protein which are tagged with a fluorescent compound and detected by fluorescence spectrophotometry. However, problems are often encountered involving the degree of antibody specificity and the high difficulty of maintaining the assays within the dynamic range of the technique. These in addition to high cost of equipment and the undefined stability of the array often results in the protein array being overlooked in favour for other applicable techniques<sup>[10]</sup>.

### 1.5.4. Tissue microarrays

This is generally used to confirm protein array and microarray results during biomarker development. It can be used for simple hybridisation experiments by placing between 50-1000 cores of different tissues on a microscope slide or to gather additional immunohistochemical information. For biomarker studies this technique is often used to compare the tissue expression of a particular biomolecule for a target illness between diseased and 'healthy' tissue<sup>[11]</sup>.

### 1.5.5. 2-D gel electrophoresis

This is a technique for the separation and identification of proteins by a process of displacement in two dimensions. These are positioned at 90° to each other to allow

separation over a larger surface area and hence, the enhancement of the resolution of each component. It can function in two different modes, to identify the global protein expression (i.e. all the proteins in the sample), or to compare two or more protein samples for the identification of any variations (i.e. differential expression). The first stage, isoelectric focussing (IEF), separates proteins or peptides according to their pI values. The subsequent step however, involves use of a surfactant, sodium dodecyl sulphate (SDS), to further separate the proteins according to molecular weight. This process is carried out within a polyacrylamide gel and occurs due to the application of an electric field. The migration distance of the protein species, represented by  $R_f$ , can be used to estimate the mass of the protein as it is negatively proportional to the log function of the molecular weight.

In order for the IEF stage to be successful a pH gradient must be initiated by use of polyacrylamide gel additives such as ampholytes and immobilines. Ampholytes are a mixture of mobile amphoteric species that have a range of pI values that must be calibrated prior to sample application. Immobilines however, are additives that are stationary within the acrylamide gel, and unlike ampholytes do not require pre-focussing or calibration. Gradients may also be set up through a mixture of both of these additives in which an immobiline gel is used and ampholytes are included within the buffer.

These gels are synonymous with difficulties involving comparative studies and data processing, and are mainly employed for differential screening in biomarker investigations. These studies involve tagging specific protein samples with different fluorescent dyes which are run simultaneously on a particular gel<sup>[12]</sup>. The capability of identifying variations between samples has been enhanced through utilising two-colour imaging and the development of new software packages. These protein spots

can also be analysed by mass spectrometry but additional sample preparation steps are required. The spots must first be removed from the polyacrylamide gel and then broken down into 'bite-sized' chunks using the enzyme trypsin to make the analysis less cumbersome. These peptide chunks are generally analysed using either a time-of-flight (ToF) or an ion trap mass analyzer using matrix-assisted laser desorption/ionisation (MALDI) or electrospray ionisation (ESI) techniques. The resulting masses are then subjected to a database search, containing all known peptide sequences and their corresponding proteins. Each of the unknown peptide sequences are given a list of possible identities and ranked according to criteria including the protein size and the number of matching peptide sequences<sup>[13]</sup>.

#### 1.5.6. Isotope coded affinity tag (ICAT) techniques

This involves differential labelling of the free cysteine residues of the protein using a 'heavy' oxygen isotope. The protein sample is digested using trypsin with subsequent separation using liquid chromatography (LC), and analysis by tandem mass spectrometry (MS/MS). This technique has a greater detection range in terms of both molecular weight and isoelectric point (pI) of the proteins when compared to 2-D gel techniques. Analyte separation is not carried out using a gel based media and is therefore not hindered by the same problems associated with the 2-D gels. However, it requires a relatively high degree of technical expertise to obtain a sensible result and it can be expensive to run routinely. In an attempt to improve its suitability for biomarker investigations developments have included improving separations based on common groups associated with biomarkers, such as phosphorylated or glycosylated moieties<sup>[12]</sup>.

#### 1.5.7. Single nucleotide polymorphism (SNP) techniques

This is a relatively recent technique that uses a genomic marker or SNP that is associated with a particular phenotype, such as a diseased state, as the target biomarker. The nature of the DNA used in this process has meant that it is unsuitable as a diagnostic test and can only determine the subjects' predisposition for the disease<sup>[7]</sup>. For this reason SNP-based techniques are commonly used to determine the genotypic or phenotypic characteristics of certain cancer types. A disadvantage of this technique is the variation associated with the function of different regions of a gene. This can greatly affect the reliability of the SNP used and as a biomarker may lead to false positive results.

#### 1.5.8. Liquid chromatography/mass spectrometry (LC/MS)

The inability of mass spectrometry to decipher the complement of mixtures has limited its applications in the past to the analysis of relatively pure compounds. Its combination with a separation technique such as liquid chromatography has now made it both a sensitive and reliable analysis of biological samples. However, it is important to understand that it is vital to choose a suitable sample preparation or separation technique for a successful analysis of the sample by mass spectrometry.

### 1.6 Biological Matrix Selection, Preparation and Analysis Techniques

The procedures used for the preparation of the sample should be selected according to the sample type with the endeavour of obtaining optimum results at the analysis stage. In this investigation we have used the ultrafiltrate or 'dialysate' after exposure to the patients blood during haemodialysis. We deemed the following four sample

preparation techniques to be most suitable for the dialysate solution due to the wide range of analytes expected; size exclusion chromatography (SEC), affinity chromatography, precipitation, and ion exchange chromatography. The dialysate sample was initially chosen as it was expected to require less preparation than the more common biological matrices of blood and serum. The premise for this is the limiting molecular weight cut-off point of the dialysis membrane thought to minimise the presence of interferences, such as albumin, that usually affect the analysis stage. In addition, the vast majority of literature involving uremic toxins has used blood, serum or urine as the target biological matrix, rendering dialysate as relatively novel for this type of analysis.

Mass spectrometry has had a relatively sporadic past for the analysis of uremic toxins. A significant paper was released during 1998 and was a review of all of the articles prior to this period that have involved this type of analysis<sup>[14]</sup>. Gas chromatography/mass spectrometry (GC/MS) has been the most common mass spectrometry technique for low molecular weight solutes in the dialysate solution, such as organic acids and phenolic compounds. However, since GC/MS is not suitable for the non-volatile middle to high molecular weight solutes, analysis of these was achieved using fast-atom bombardment (FAB) mass spectrometry and in some cases liquid secondary ion mass spectrometry (LSIMS). These techniques enabled the analysis of molecules such as peptides and nucleosides that were previously considered difficult to monitor by mass spectrometry. The development of ionisation techniques, such as atmospheric pressure chemical ionisation (APCI), electrospray ionisation, and matrix-assisted laser desorption/ionisation has enabled the analysis of a wide range of molecular weight species at high sensitivity. The increasing range of biomolecules amenable by these techniques, and their high levels of sensitivity and



reproducibility provide a strong case for using this type of analysis for identifying novel uremic toxins. Mass spectrometry for the analysis of mixtures such as a biological matrix, requires some degree of sample preparation. A large proportion of bioanalysis is carried out with mass spectrometry interfaced with the separation technique, liquid chromatography. Liquid chromatography/mass spectrometry can enable the identification and characterisation of biomolecules at trace levels within large complex mixtures. Results can be obtained with a high degree of reproducibility and as part of a high throughput analytical protocol. These characteristics make LC/MS a vital component of bioanalysis laboratories of many pharmaceutical and research companies.

## 1.7 Separation Science

There are a vast number of techniques employed for the separation of biomolecules, each chosen to provide optimum isolation of the chosen analyte of interest. Method development of the following techniques was carried out for the investigation of haemodialysate to identify novel uremic solutes.

### 1.7.1. Solid phase extraction (SPE)

This preparative technique simply isolates a group of analytes by passing the sample through a solid phase material. It can be used for this purpose either by retaining interferences and collecting the analytes of interest, or by retaining the analytes with subsequent elution followed by the removal of any interferences. The latter method generally involves four main steps as shown in the diagram overleaf (figure 1.2).

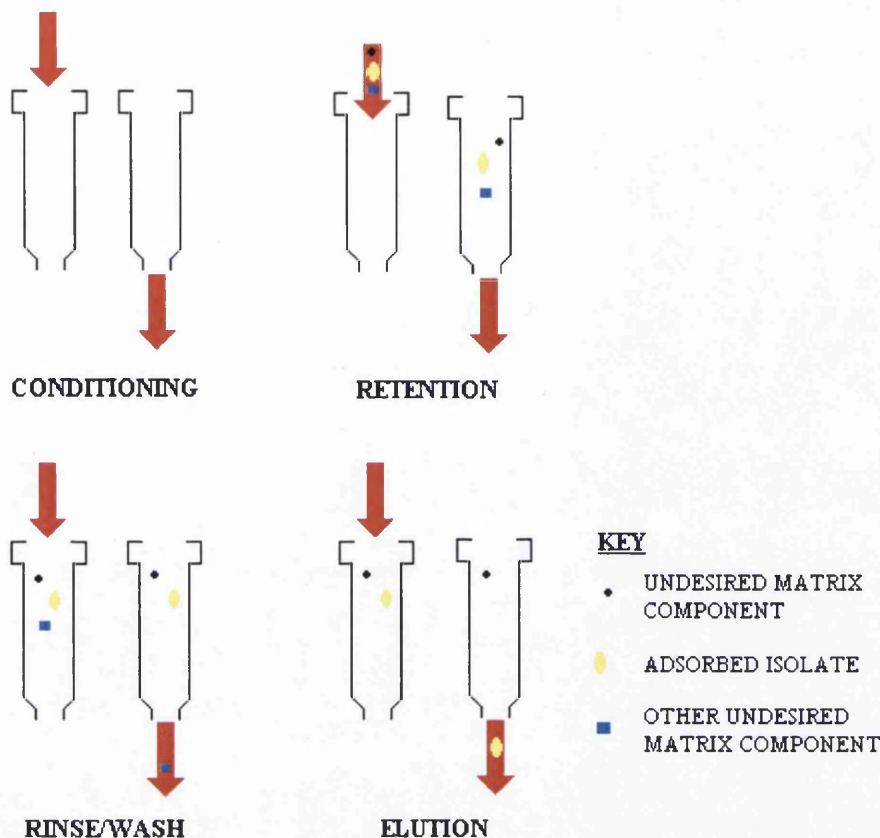


Figure 1.2: Illustration depicting four common steps of the SPE process.

Stage 1: Conditioning – ‘activates’ the sorbent bed to retain analytes and ensures reproducible retention of the analyte of interest (the isolate).

Stage 2: Retention – adsorption of isolate and some interferences but the remaining impurities freely pass through the sorbent.

Stage 3: Rinse/Wash – intended to remove any undesired bio-matrix components that are retained on the sorbent.

Stage 4: Elution – isolate is eluted by applying small volume of solvent that concentrates and purifies the sample; increases the detection limits and simplifies the analysis by removing the impurities.

The initial ‘conditioning’ stage of the SPE process is aimed at preparing or activating the solid phase packing for the retention of analytes of interest. This is followed by the application of the sample and retention of the analyte by adsorption to the column

packing. The third step of the extraction is commonly known as the 'rinse' stage and any retained interfering impurities are washed off the sorbent bed. The final step involves the elution of the analytes of interest with a small volume of solvent which effectively 'cleans-up and concentrates' the sample. The SPE process is therefore intended to simplify the analysis of the sample by simply removing any unwanted impurities that would normally interfere with this process and to provide lower detection limits for the analytes of interest<sup>[15]</sup>.

### 1.7.2. Size exclusion chromatography (SEC)

This separation technique is also known as desalting and is based on differences of hydrodynamic volume of analytes as they flow through a column containing a polystyrene resin stationary phase. The relationship of hydrodynamic volume and molecular weight of the polystyrene resin enables the determination of the molecular weight of an analyte. This calculated molecular weight will only be an approximation as this relationship does not provide a constant value for all polymer resins. The column resin is essentially a porous medium constructed from polymer beads, and can have a wide range of pore sizes which eventually determines the molecules that can be separated. This range is known as the fractionation or exclusion range of the resin and analytes that are too large to enter the pores will flow around the resin beads to elute first. The smaller low molecular weight analytes that are capable of entering the resin pores will have a longer path to travel within the column, and therefore, elute later.

If the sample is suspected to contain some form of protein it is suggested that the column is pre-equilibrated with a suitable buffer. Tris-HCl buffer at pH 7 is commonly used as it is suitable to collect the sample analytes without altering the

possible protein structures in the sample. Another important factor for a successful separation is to select an appropriate column size for the volume of the sample. For example, if the column is too large, it can result in the dilution of the sample and alter the analytical sensitivity for the target analytes. However, if the column is too small the low molecular weight contaminants will not be separated to a sufficient degree from the analytes of interest. A general rule of thumb for a successful separation is that the column selected should be capable of a volume that is 4-20 times greater than the sample volume<sup>[16]</sup>.

### 1.7.3. Liquid-liquid chromatography/extraction

This form of separation is based on the solubility of an analyte for a particular solvent and can involve repeated partitioning steps between a liquid stationary and mobile phase. Two immiscible solvents are combined in a container and the analyte of interest passes from the solvent of origin into a polarity compatible solvent. For example, an analyte with a large non-polar section contained in an aqueous biological matrix such as plasma, will partition into an immiscible non-polar solvent. The degree of separation is dictated by an expression known as the partition coefficient and is represented by the term  $\log P$ . This is essentially the log ratio of the concentration (C) of a solute in one phase to another phase at equilibrium conditions:

$$\log P = \frac{C_{m2}}{C_{m1}} \quad \text{Equation 1.1}$$

where,

$C_{m2}$  = Molar concentration of the solute in mobile phase 2

$C_{m1}$  = Molar concentration of the solute in mobile phase 1

The conventional conditions by which most partition coefficients are compared is the octanol:water system. This particular partition coefficient is used as a measure of the hydrophobicity or hydrophilicity of an analyte depending on the solubility in the octanol or water layers respectively<sup>[17]</sup>.

#### 1.7.4. Ion exchange chromatography (IEC)

Ion exchange chromatography involves passing a sample through a column containing a stationary phase with charged ionic groups that interact with oppositely charged functional groups of the sample analytes. Protein and peptide biomolecules have terminal functional groups capable of opposite charges and are known as zwitterion structures. This effect can result in an overall molecular net charge of zero, and is known as the isoelectric point (pI). When a protein or peptide is placed in a buffer solution of a pH greater than the pI value the carboxy terminus will deprotonate to give an overall negative charge. This will enable the protein to bind to any positively charged functional groups of the stationary phase within the column. Conversely, if the protein or peptide is in a buffer of pH lower than the pI value then the amino terminus will protonate providing it with an overall positive charge that can be retained by negatively charged groups of the column. Thus, varying the buffer pH following retention of the analyte of interest will cause its elution from the column<sup>[16]</sup>. This technique can also be used to desalt a sample which has proven to be a particular problem with hemodialysis concentrate.

#### 1.7.5. Reverse phase-high performance liquid chromatography (RP-HPLC)

This was initially developed from an existing separation technique known as normal phase chromatography. The column packing designed for normal phase conditions

consisted solely of polar silanol groups and capable of retaining polar compounds. However, this was unable to provide sufficient separation of mixtures containing relatively non-polar compounds, and hence, the polar packing material was modified to include non-polar groups such as C-8 or C-18 carbon chains. Thus, unlike normal phase conditions, reverse phase chromatography could separate mixtures of non-polar compounds using a non-polar stationary phase. Retention of analytes during reverse phase chromatography is achieved using a polar aqueous mobile phase with the gradual increase in its non-polar complement for analyte elution. This is one of two common techniques utilised in HPLC, known as gradient elution, where mobile phase composition is altered to steadily remove the analytes from the stationary phase. The second is the isocratic elution method and involves maintaining a constant mobile phase composition throughout the separation<sup>[18,19]</sup>.

There are several parameters used to assess the suitability of the chromatographic conditions and the overall performance of the system<sup>[17]</sup>. The main objective of a good chromatographic system is to achieve good separation with sharp well resolved peaks.

#### 1.7.5.1. Partition coefficient (K)

This parameter indicates the degree of analyte distribution in the stationary and mobile phases. It is defined as the ratio of the concentration of a solute in the stationary phase divided by its concentration in the mobile phase at equilibrium (see equation 1.2).

$$K = \frac{C_s}{C_m} \quad \text{Equation 1.2}$$

where,

$C_s$  = Molar concentration of the solute in the stationary phase

$C_m$  = Molar concentration of the solute in the mobile phase

This measure is related to the rate at which the analyte migrates through the column or its retention time ( $t_R$ ); the higher affinity the analyte has for the stationary phase (the higher its concentration in the phase) the more time it will spend on the column, resulting in a longer retention time. Using the expression above, a high value for the partition coefficient ( $K$ ) is indicative of a high concentration of analyte in the stationary phase and a longer retention time. To summarise, the partition coefficient is directly proportional to the retention time of the solute on the column.

#### 1.7.5.2. Resolution (R)

This may be defined in terms of chromatographic performance as a meaningful measure of separation. In its simplest guise the following expression can be used to quantify the resolution of components in a chromatogram, with the origin of the values shown by the illustration in figure 1.3.

$$R = \left( \frac{2(t_{R2} - t_{R1})}{W_1 + W_2} \right)$$

Equation 1.3

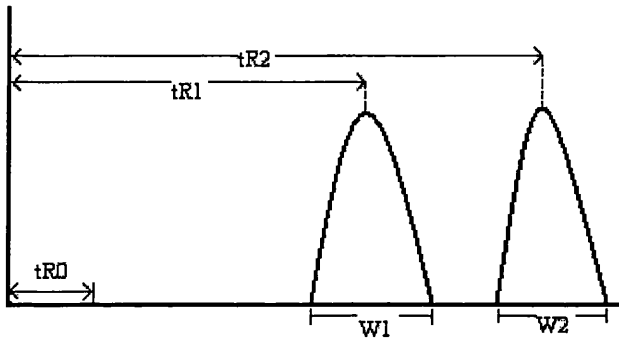


Figure 1.3: Hypothetical chromatogram showing two peaks with their associated retention times ( $t_{R1}$  and  $t_{R2}$ ) and peak widths ( $W_1$  and  $W_2$ ) respectively. The value  $t_{R0}$  represents the retention time of the mobile phase peak or solvent front. Resolution is a parameter dependent on the retention time separation of the components and peak width, and therefore considers the degree of peak overlap.

In addition to variations in solute polarity, the distance between two peaks ( $t_{R2} - t_{R1}$ ) is dependent on the selectivity or the separating power of the column, and the term ( $W_1 + W_2$ ) is affected by the column efficiency or the number of theoretical plates of the column ( $N$ ). Therefore the value obtained for peak resolution can provide information regarding the suitability of the column for the analysis. Essentially the principal aim in chromatography is to obtain the highest possible peak resolution within the shortest possible elapsed time. There are a number of conflicting parameters required for this and often a compromise between peak resolution and overall time of the chromatographic run is necessary. A more accurate expression for calculating resolution can be used to achieve this compromise and considers other chromatographic parameters each including the dimension of time (see equation 1.4).



$$R = \frac{1(\alpha - 1)(k')}{4(\alpha)(1 + k')} \times N^{\frac{1}{2}} \quad \text{Equation 1.4}$$

where,

$k'$  = Capacity or Retention Factor

$\alpha$  = Relative Retention or Selectivity or Separation Factor

$N$  = Column Efficiency

#### 1.7.5.2.1. Capacity factor ( $k'$ )

This is used to describe the rate of solute migration through the column and is related to the sorption or partition coefficient ( $K$ ) of a sample component with the compatibility of the stationary phase. It can be defined for a sample component as:

$$k' = \frac{t_R - t_0}{t_0} \quad \text{Equation 1.5}$$

where,

$t_R$  = retention time of the sample component

$t_0$  = retention time of an un-retained component

The capacity factor of the chromatographic system may be improved by using a more suitable mobile phase composition or stationary phase packing.

1.7.5.2.2. Relative retention ( $\alpha$ )

The relative retention for two sample components,  $R_1$  and  $R_2$ , on a column may be defined as the ratio of the affinity of each component for the stationary phase (the partition coefficients):

$$\alpha = \frac{K_{R_2}}{K_{R_1}} \quad \text{Equation 1.6}$$

where,

$R_2$  = least strongly held component

$R_1$  = most strongly held component

This relationship between relative retention and partition coefficient can provide an indirect link to calculating the previously described resolution parameter, the capacity factor ( $k'$ ).

1.7.5.2.3. Column efficiency (N)

This function of resolution is defined in terms of the number of theoretical plates or the number of suitable retentive sites for a component to reside in a column. It can be determined using two different expressions depending on the information available to the chromatographer (see equation 1.7).

$$N = 16 \times \left( \frac{t_R}{W_B} \right)^2 \quad \text{Equation 1.7}$$

where,

$t_R$  = retention time of the peak of the component R

$W_B$  = width at the base of peak R

or,

$$N = 5.545 \left( \frac{t_R}{W_{0.5h}} \right) \quad \text{Equation 1.8}$$

where,

$W_{0.5h}$  = peak width at half height of peak R

Therefore it can be observed that the column efficiency is related to peak broadening and the time spent on the column for a particular component. Column efficiency can also be expressed in terms of the column height equivalent to one theoretical plate or HETP:

$$\text{HETP} = \frac{\text{Column Length (L)}}{\text{Number theoretical plates (N)}} \quad \text{Equation 1.9}$$

Hence, it can be proposed that the efficiency of the chromatographic column increases, as the number of theoretical plates increase and the height equivalent to one plate decreases. Also, as resolution is proportional to the number of theoretical plates

to the power of a half ( $N^{1/2}$ ), peak resolution can be enhanced by either maximising the length or the number of theoretical plates of the column, and by minimising the height equivalent to one theoretical plate.

## **1.8. Mass Spectrometric Analysis**

### **1.8.1. Origins of mass spectrometry**

The pioneer of mass spectrometry, Sir J. J. Thomson first discovered the electron in 1897, and was followed by the invention of the mass spectrometer, then known as the parabola spectrograph<sup>[20]</sup>. In 1906 his contributions were recognised with the presentation of the Nobel Prize in Physics, and the next major advance in mass spectrometry was not until 13 years later in 1919. This was the development of a higher resolution mass spectrometer by Francis W. Aston for which he was awarded the Nobel Prize in Chemistry for isotope discovery<sup>[21]</sup>. The subsequent years entailed the introduction of the magnetic deflectron mass spectrometer<sup>[22]</sup> with direction focussing capabilities. This instrument also saw the first use of the electron impact ionisation source and is still commonly used in modern mass spectrometry. This type of mass spectrometer was continually developed and modified over the next twenty years and in the early 1940's Nier and co-workers tailored it for isotopic analysis<sup>[23]</sup>. Also, during this period the high-mass resolution double focussing mass spectrometer was invented by Dempster and first unveiled by Mattauch and Herzog which enabled mass assignments to be made with greater confidence. The late 1940's saw the invention of the time-of-flight (ToF) mass spectrometer developed by Cameron and Eggers<sup>[24]</sup> from the initial ideas proposed by William E. Stephens in 1946<sup>[25]</sup>. The subsequent 40 years involved vast improvements in the mass resolution of the ToF

analyser especially with the introduction of the reflectron lens, invented by Mamyrin<sup>[26, 27]</sup>. This degree of mass resolution was later surpassed by the invention of the Fourier-transform ion cyclotron resonance mass spectrometer (FT-ICR-MS) by Comisarow and Marshall in 1974<sup>[28]</sup>, and is still the leader in obtaining high mass resolution. Alongside the development of FT-ICR-MS was the invention of chemical ionisation source and led to the design of a similar source capable of operating at atmospheric pressure conditions, known as APCI, by Horning and co-workers<sup>[29]</sup>. Successive developments involved the interfacing of liquid chromatography (LC) systems to mass spectrometers by Arpino, Baldwin, and McLafferty<sup>[30]</sup> which was previously considered impractical due to the vacuum conditions of the mass spectrometer. This progress in mass spectrometry was accompanied by the introduction of other new ionisation sources such as fast-atom bombardment (FAB) in 1981 and matrix-assisted laser desorption/ionisation (MALDI) in 1988. Current bioanalysis frequently employs these new ionisation techniques and in particular the use of APCI, electrospray ionisation (ESI), and MALDI are becoming standard practise.

### 1.8.2. An introduction to mass spectrometry

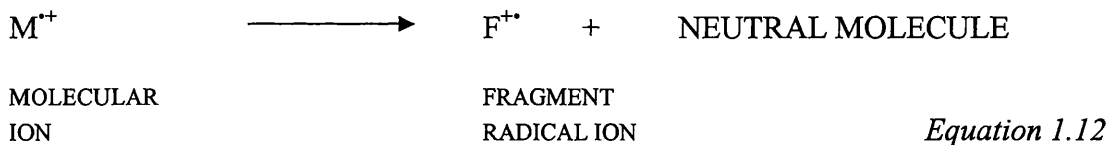
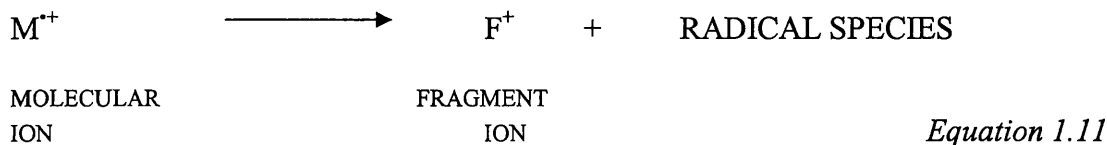
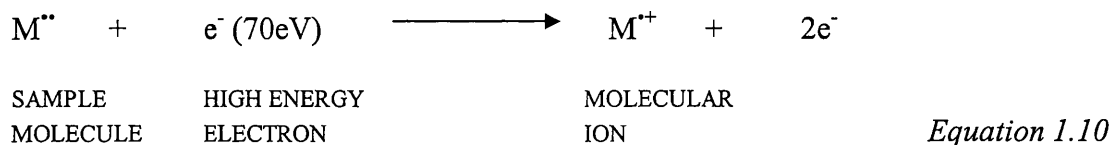
Mass spectrometry is a powerful analytical technique involving the separation of gas phase ionic species according to mass-to-charge and is considered to have three main functions:

1. molecular weight determination,
2. chemical structural characterisation, including isotope analysis and
3. qualitative and quantitative analysis of analyte(s) in a mixture.

The instruments capable of this technique are collectively known as mass spectrometers although, a wide variety of types are available, some of which are discussed later. A common feature of all these forms is that the sample analytes must first be ionised to enable their separation according to a property known as the mass-to-charge ratio ( $m/z$ ). A mass spectrometer can be divided into four main sections:

1. The sample inlet: where samples are introduced
2. The ion source: where charged ions are formed from the analytes
3. The mass analyser: separates ions according to their  $m/z$  values after passing through electric and/or magnetic fields
4. An ion collection system: collects the separated ions to identify their abundance by  $m/z$ .

Modern designs, such as, the quadrupole, quadrupole ion trap, Fourier transform-ion cyclotron resonance (FT-ICR) trap, Orbitrap and time-of-flight (ToF) instruments require specialised extraction and acceleration ion optics to transfer ions from the source to the mass analyser and shall be discussed later. Early mass spectrometer designs used an ionisation method known as electron impact (EI). Here the sample is vapourised and an electron is removed from the sample analyte, known as the secondary electron, by an energetic electron beam. This usually forms a radical cation referred to as the molecular ion  $M^{+\cdot}$  (equation 1.10), which has an excess internal energy that may be internally converted and cause the ion to either fragment further (equation 1.11) or to rearrange (equation 1.12).



These ions are then accelerated under vacuum conditions within the mass spectrometer to the mass analyser, with the aim of separating and measuring the mass (mass-to-charge) of the resulting ions.

The mass measurement of ions is achieved by monitoring different parameters depending on the type of mass analyser, such as time-of-flight or an appropriate magnetic field to maintain stable ion trajectory. Separated ions hit the detector producing a signal and can provide, not only, mass-to-charge ratio ( $m/z$ ) information of the ion from the analyser but its relative abundance within the sample. This information is converted into a chart known as a mass spectrum comparing ion intensity and  $m/z$  with the ion of highest abundance known as the base peak.

### 1.8.3. Sample ionisation and modern interfaces

The interface, in addition to ionisation, is required to efficiently transfer analytes from a solution phase to a gas. The connection of a LC system to a mass spectrometer

incorporates both the powerful separation capabilities of chromatography with the sensitivity for detection of mass spectrometry. Initially the major hindrance of LC/MS was the incompatibility of the then LC flow rates (0.5-2mL/min of normal or reversed phase systems) and the vacuum requirements of the mass spectrometer<sup>[31]</sup>. Since the 1970's many interfaces were developed to overcome this problem, each of which were suitable for particular mass ranges and polarity as shown in figure 1.4.

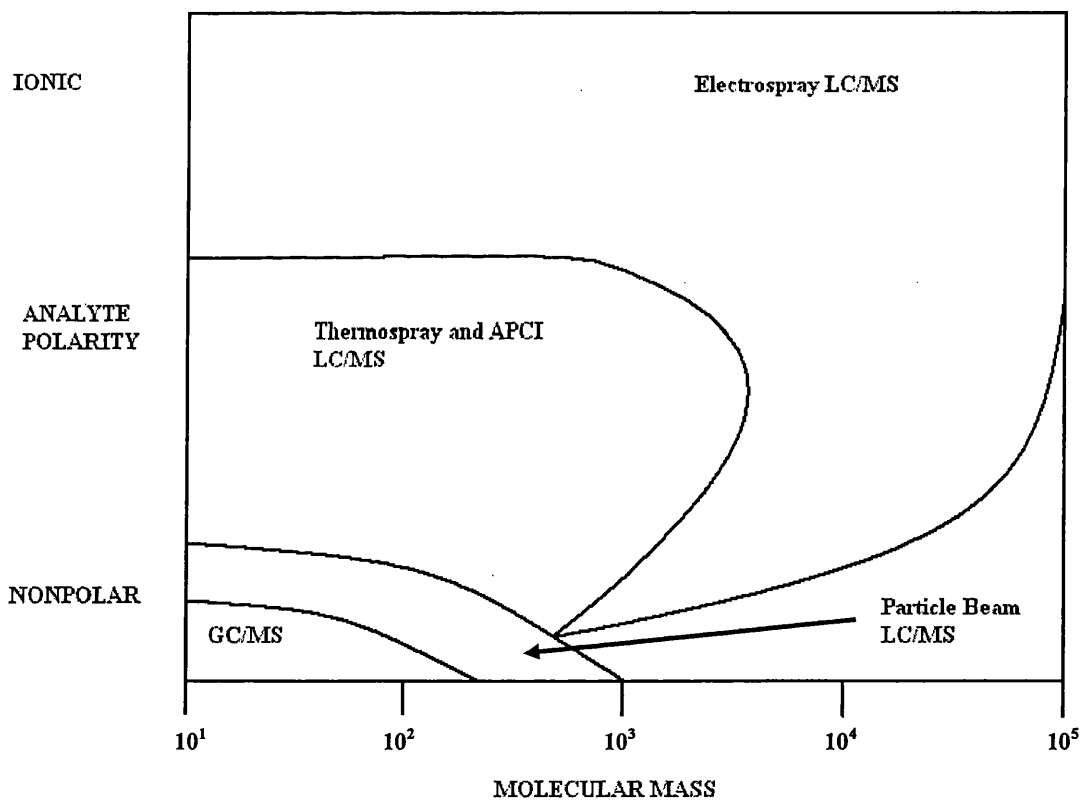


Figure 1.4: Suitability ranges of LC/MS interfaces regarding polarity and molecular mass of an analyte<sup>[31]</sup>.

The balance that the interface must achieve is to accept as much LC eluent as possible to attain maximum sensitivity and yet minimise the liquid load into the mass



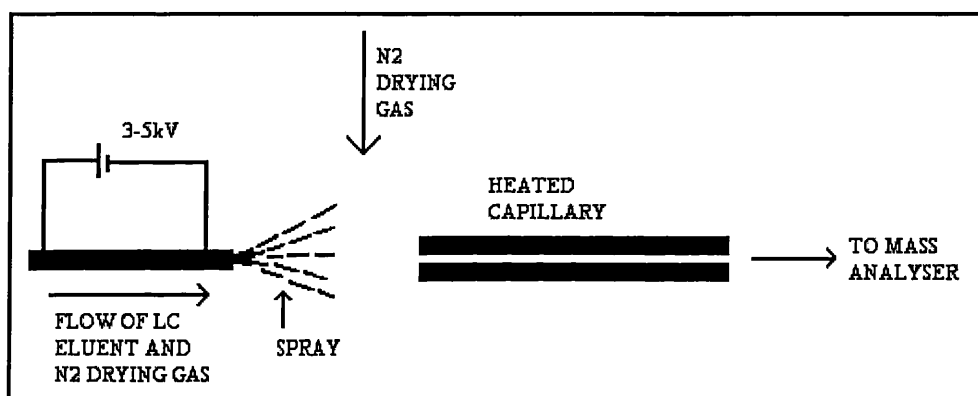
spectrometer. In addition to solvent flow other incompatibilities can arise particularly during bioanalytical applications, for example non-volatile buffers used for some high quality separations were unsuitable for the past ionisation sources. Several different methods to overcome these interfacing difficulties have been attempted ranging from;

- a) increasing the pumping capacity of the vacuum system of the mass spectrometer,
- b) minimising or elimination of solvent prior to entering the vacuum system,
- c) minimising solvent flow into the vacuum system by splitting the flow to waste, but at a loss of sensitivity,
- d) use of micro-LC systems capable of efficient separation at lower flow rates compatible with the mass spectrometer,
- e) additional pumps at ionisation source improving vacuum system and,
- f) development of an ionisation source capable of operating at atmospheric conditions.

Many interfaces utilising these approaches were developed although only a select number are utilised for the analysis of biosamples. Two of the most commonly used interfaces, electrospray ionisation (ESI) and atmospheric pressure chemical ionisation (APCI) function at atmospheric conditions thus avoiding many of the issues associated with LC/MS. Other ionisation sources used in the bioanalytical field such as matrix-assisted laser desorption/ionisation (MALDI), still remain separate from the LC system and therefore additional sample preparation is required.

### 1.8.3.1. Electrospray ionisation (ESI)

This is one of the main atmospheric pressure ionisation (API) techniques and was first used commercially in 1985 by Whitehouse and co-workers<sup>[32]</sup>. The sample is sprayed into the source through a very fine needle with a potential difference of approximately 3kV. Upon exiting the needle the spray forms a Taylor cone followed by the formation of droplets and are charged by an electric field of  $10^6 \text{ Vm}^{-1}$  that is applied to the tip of the capillary needle.



*Figure 1.5: Schematic diagram of an electrospray ionisation source indicating the flow of liquid from the LC system and the application of drying gas enabling the formation of ions.*

These charged droplets are then desolvated by a flow of hot nitrogen gas which results in charged or multiply charged analyte molecules<sup>[33]</sup>. There are two main theories describing the ionisation process; one model proposed by Dole and colleagues<sup>[34]</sup> is that the charged droplet undergoes a Coulombic explosion during desolvation and results in the formation of smaller droplets, eventually leading to desolvated analyte ions<sup>[35,36]</sup>.

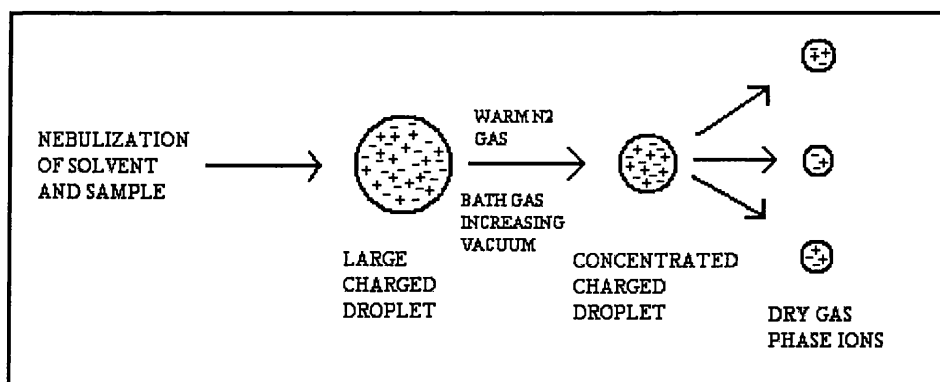


Figure 1.6: An illustration depicting the proposed Coulombic desolvation process of ESI. Charged sample droplets are evaporated causing an increase in the surface charge of the droplet promoting a Coulombic explosion to form smaller charged droplets.

The second model devised by Iribarne and Thomson<sup>[37, 38]</sup> is known as the 'ion-evaporation' model and is considered as the most likely as it provides the most adequate explanation for the ionisation process. The vast majority of ions present in the charged droplet are thought to be pre-formed and originate from the acid-base chemistry of the solution. Hence, pH manipulation to increase the abundance of charged basic analytes is of considerable importance in achieving good sensitivity when observing ions in positive mode.

ESI is considered a 'soft' ionisation technique as it involves minimal fragmentation of the sample analytes capable of providing molecular weight information. This method is suitable for the analysis of proteins due to the capability of multi-charging, in which one charge typically associates per 1000Da of protein. A protein of molecular mass 60,000Da will have a distribution of 40-80 positive charges, depending on the number of 'basic' residues and it will therefore have a  $m/z$  value of 1000-2000Da. This  $m/z$  value is easily attainable by most mass

spectrometers, although the actual  $m/z$  may be slightly different from the 'true' value due to the multiple charging. However, if a 'contaminant' is present it can suppress the ionisation of the target analyte, even if it is at a lower abundance, and consequently 'masks' the analyte signal from detection<sup>[39]</sup>. This should be considered when choosing ESI as an ionisation technique and the interpretation of the resulting mass spectrum. Ionisation suppression can pose more problems when running online LC/MS incorrectly as sample components from overloading the LC column may carry over into subsequent separations and suppress components from other samples.

#### 1.8.3.2. Matrix-assisted laser desorption/ionisation (MALDI)

MALDI essentially employs a matrix to transfer energy to the analyte to facilitate ionisation. The energy is supplied by a pulsed laser at a wavelength that is absorbed by the matrix. Chromophore containing matrices commonly use  $N_2$  and  $Nd^+/YAG$  lasers emitting at wavelengths of 337nm and 1064nm, respectively, the  $Nd^+/YAG$  is frequency tripled to a wavelength of 354nm. The laser ablates the matrix and ionisation occurs by a number of processes<sup>[40]</sup> with the resulting gas phase ions passing into the mass spectrometer. There is a wide selection of possible matrices, each suitable for ionising different types of synthetic and biological molecules. In addition to matrix choice other parameters such as solvent composition and the inclusion of additives, such as trifluoroacetic acid (TFA), facilitate and improve ionisation of the target molecule.

MALDI is usually coupled to a time-of-flight mass analyser and will be discussed later. MALDI mass spectra generally have poor resolution and mass accuracy with sensitivity primarily dependent on the analyte. They usually consist of radical molecular ions and protonated molecules of both sample components and

matrix. Additional peaks may be observed as a consequence of sample degradation during laser application, and adduct formation, from the presence of sodium ions in biosamples. These considerations must be taken into account during the interpretation of such mass spectra and have particular importance when applying a database search, such as those used in protein identification. The major upshot of MALDI-ToF is that it is easy to use and is capable of identifying very large molecular weights of up to approximately 500kDa. It is therefore used during the analysis of very large biomolecules such as proteins or DNA fragments.

This ionisation technique has not been commonly associated with LC, however, current developments have been directed towards interfacing these systems. For example, a study published in 2004<sup>[41]</sup> involved adapting the outlet via the LC system with a programmed heated needle that placed droplets of the sample mixture onto a MALDI target plate. This is still deficient in some requirements for an ideal interface and currently remains as a separate off-line technique to LC.

#### 1.8.4. Ion separation and analysis

Following ionisation, gas phase ions are guided into the mass analyser from the source. There are a number of types of mass analysers available each capable of determining the mass of an ion despite measuring different parameters. For example, both magnetic sector and quadrupole analysers rely on magnetic and electric fields and can undertake selected ion monitoring (SIM), collecting data from one or more target ions. SIM has the added advantage of increasing the sensitivity of the experiment by increasing the signal-to-noise ratio by as much as 100:1. However, mass analysers, such as, time-of-flight (ToF) simply obtain a mass measurement from the flight time of the ion and are used for experiments that can search a high mass

range for ions. Mass analysers capable of obtaining measurements with high peak resolution can obtain an ion accurate mass or the mass-to-charge of an ion to within four decimal places. Once obtained elemental composition information is suggested as only specific combinations of elements will match the obtained formulae. The suggested elemental formulae are evaluated according to four main parameters each important in selecting the most likely identification:

a. **Mass error** – this is the measure by which the accurate mass of the suggested elemental formula differs from the calculated accurate mass of the unknown ion. It is a value provided in mDa or ppm.

b. **Isotope pattern** – this should be checked against the suggested elemental formulae matching that of the unknown. Using Xcalibur software v.2.0 (Thermo Fisher Scientific) a theoretical isotope pattern can be generated and compared against that of the unknown to confirm if the elemental formula is suitable for the unknown.

c. **DBE** – the double bond rule. This is often determined by the following equation,

$$\text{DBE} = \text{Cx} - 1/2\text{Hy} + \text{Nz} + 1$$

where,

**x** = number of carbon atoms

**y** = number of hydrogen atoms

**z** = number of nitrogen atoms

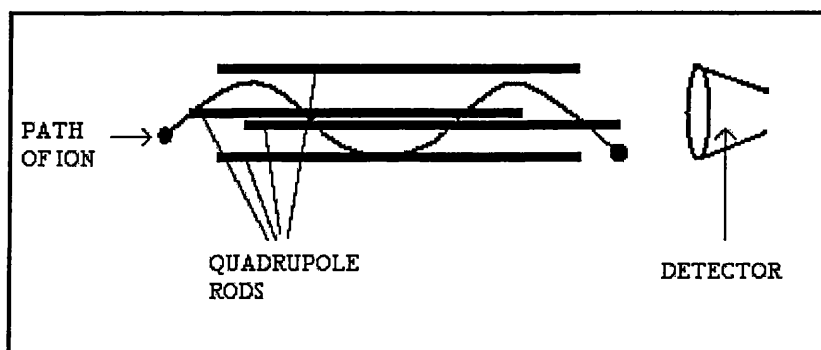
The calculated value obtained is either a whole number or a half, signifying an odd electron species ( $M^{+\bullet}$ ) and an even electron species ( $M^+$  or  $[M+H]^+$ ) respectively. Odd electron species are unlikely protonated parent molecules in electrospray ionization and so are excluded from the final selection of elemental formulae.

d. **Nitrogen rule** – this states that an even number of nitrogen atoms suggested for the elemental formulae should have an odd mass for the even electron ion  $[M+H]^+$ .

Analysers with poor mass separation or resolution often have other capabilities, such as  $MS^n$  experiments, and can be included with high resolution analysers as a hybrid instrument to benefit from both characteristics.

#### 1.8.4.1. Quadrupole mass filter

This, as the name suggests, consists of four metal rods and are arranged as shown in figure 1.7.



*Figure 1.7: Schematic of a quadrupole mass analyser showing the passage of an ion with a stable trajectory of specific  $m/z$  into the detector. Opposing pairs of poles have dc and rf fields applied.*

Opposing pairs of rods are connected electrically and a voltage consisting of both radiofrequency (rf) and direct-current (dc) field components are applied with the rf field present being  $180^\circ$  'out of phase' with the dc field. The potential between each rod ( $\phi$ ) can be calculated using the by the equations below:

$$\Phi_0 = +(U - V \cos \omega t) \text{ and } \Phi_0 = -(V \cos \omega t) \quad \text{Equation 1.13}$$

where,

$U$  = direct potential (dc) field

$V$  = amplitude of the rf field

$\omega$  = angular frequency =  $2\pi\nu$

$\nu$  = frequency of the rf field

At a specific voltage ions of the corresponding  $m/z$  follow a stable trajectory through the rods into the detector. Ion path stability is determined by two functions  $a$  and  $q$ , which are dependent on the dc and rf values respectively. These are calculated using the expressions 1.14 and 1.15.



$$a = \frac{8eU}{mr_0^2\omega^2} \quad \text{Equation 1.14}$$

$$q = \frac{4eV_0}{mr_0^2\omega^2} \quad \text{Equation 1.15}$$

where,

$r_0$  = radius of the field within the circular  
arrangement of the rods

$m$  = mass of ion

This relationship of ion stability is often depicted as a graph indicating the boundaries of the stable oscillation of an ion (see figure 1.8).

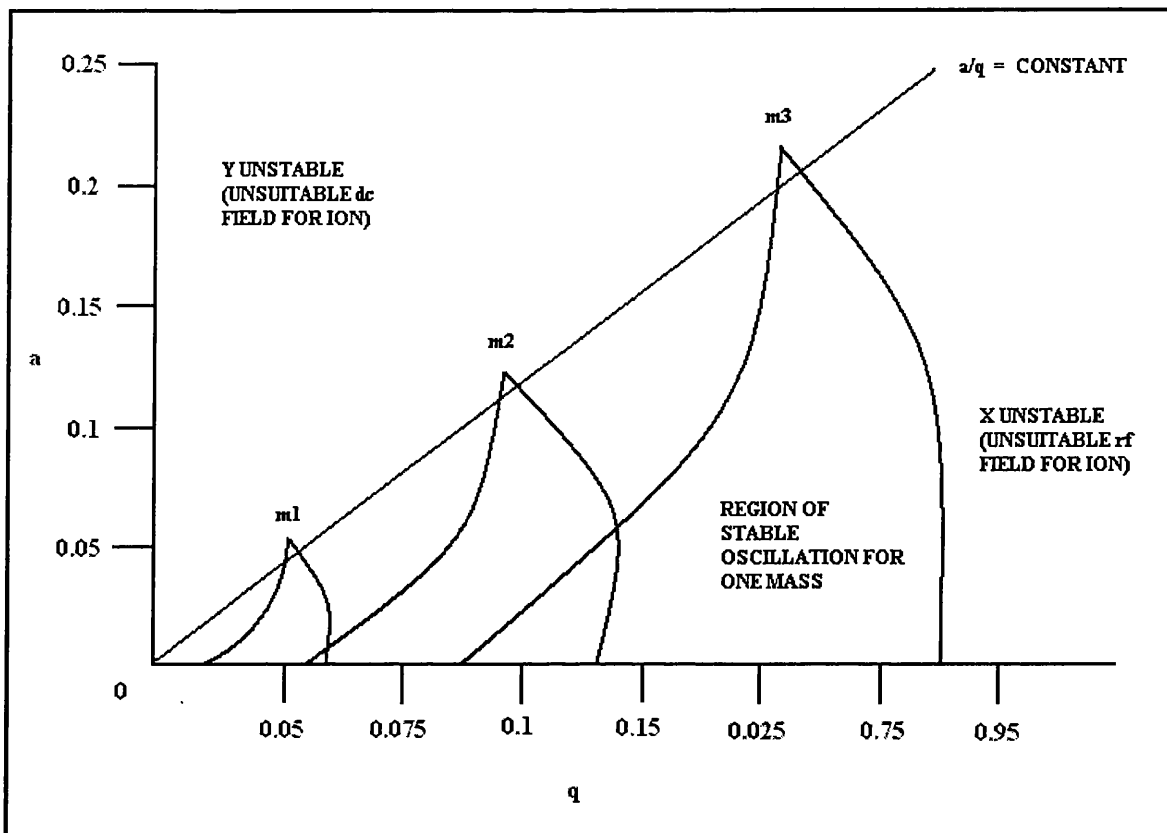


Figure 1.8: Stability diagram of ion trajectories within a linear quadrupole and the area under the curve depicting the dc ( $a$ ) and rf ( $q$ ) fields suitable for a stable oscillation of an ion through the quadrupole where  $m1 < m2 < m3$ .

A mass spectrum is produced by scanning  $U$  and  $V_0$ , so that  $a/q$  remains constant, and the mass of the ion transmitted is directly proportional to the values of  $a$  and  $q$ . The quadrupole is considered to be the ideal analyser for interfacing with liquid chromatography. This mass analyser is capable of fast mass scanning and uses a low accelerating voltage that increases its compatibility to high operating pressures, such as those encountered in LC/MS. In tandem mass spectrometry techniques (MS/MS), i.e. those involving the coupling of more than one mass analyser, the quadrupole mass analyser can be linked in succession to form a triple quadrupole analyser. This is the

most common apparatus used in MS/MS and consists of three sets of quadrupoles in a series;

- Q1 = A mass filter using both rf and dc fields for first stage of ion monitoring.
- Q2 = Collision cell containing inert gas and is enclosed to increase the localised gas pressure for fragmentation of ions transmitted by Q1. Uses rf only and focuses any product ions into Q3.
- Q3 = The second mass filter with both rf and dc fields for product ion monitoring.

The quadrupole may also be included in another type of tandem instrument known as the hybrid mass spectrometer. In this instrument the final quadrupole of a triple quadrupole system is replaced, for example, a ToF analyser. This Q-ToF mass spectrometer contains a quadrupole linked to a collision cell, and is linked to a ToF analyser capable of detecting ions that enter at a specific time.

#### 1.8.4.2. Quadrupole ion trap mass analyser

This mass analyser can be thought of as a folded quadrupole in which a circular centre ring electrode has two end-cap electrodes, one above and one below. Essentially the ions are dynamically stored within this three-dimensional quadrupole by the application of an rf field. Following their introduction into the trap they move in a stable, but complex trajectory stabilised by the presence of a buffer gas such as helium<sup>[42, 43]</sup>.

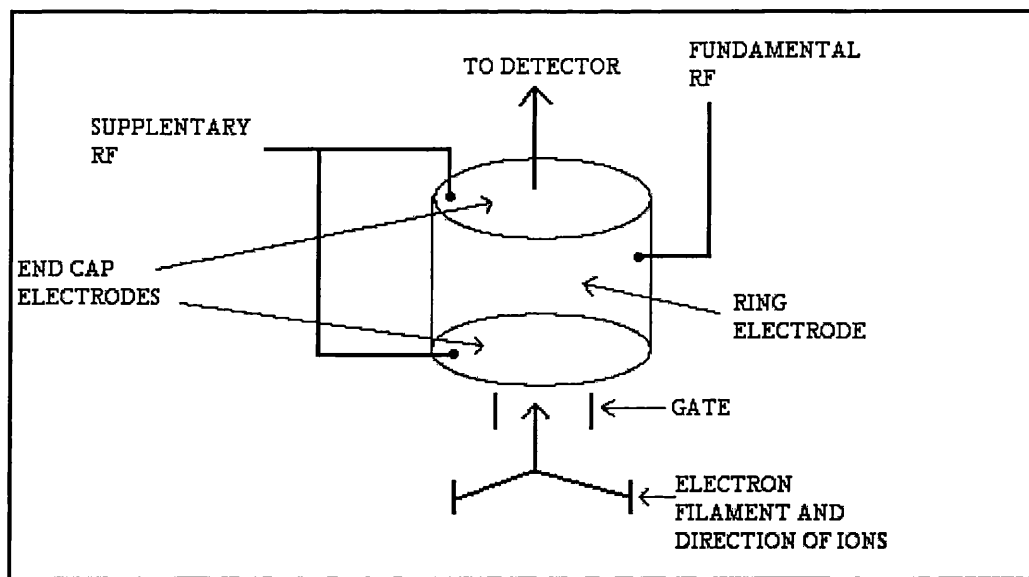


Figure 1.9: Schematic of a cylindrical quadrupole ion trap indicating the ion path via the electron filament of the ion source to the detector. The applied rf fields are used to control which ions are present in the trap. They are capable of retaining both all ions in the mass range of the analyser (a full mass scan) and specific ions (a single ion monitoring scan).

These trapped ions can be manipulated by varying the voltages applied to the trap to perform ion ejection, ion excitation, and mass-selective ejection. If this voltage manipulation is carried out in a systematic fashion a complete mass spectrum can be obtained and mass-tandem experiments (i.e. MS/MS, MS/MS/MS, and MS<sup>n</sup> etc) to obtain fragmentation information of the analyte. A major drawback that must be considered with this analyser is the reaction of ions with any neutral species present in the trap. This is known as 'self-chemical ionisation' and can affect the resulting mass spectrum<sup>[44]</sup>. A common technique to reduce this effect is to introduce the ions via an external source. This however, can lead to ion losses during transmission and trapping thus, resulting in a small reduction in the sensitivity of the method. Modern

ion traps are highly refined instruments and use techniques such as automatic gain control (AGC) which restrict ion density within the trap, preventing the 'self chemical ionisation' process.

#### 1.8.4.3. Time-of-flight (ToF) mass analyser

This is the simplest of all mass analysers and is based on the idea that all ions produced in the source have the same given kinetic energy. The velocity of each ion will therefore be inversely proportional to the square root of its mass and the time the ion takes to travel down a field-free flight path will be related to the  $m/z$  of the ion.

$$t = d \sqrt{\frac{m}{(2zeV)}} \quad \text{Equation 1.16}$$

where,

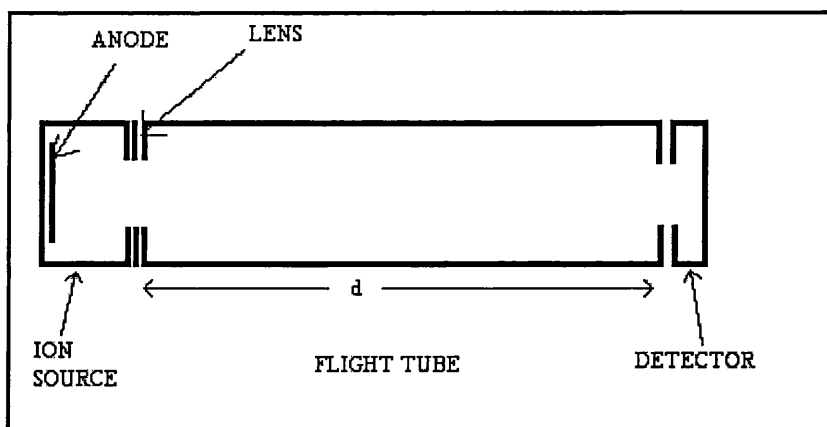
t = time of flight

d = distance travelled

m = mass of ion

V = accelerating voltage and

z = number of charges on the ion.



*Figure 1.10: Basic schematic of a time-of-flight mass analyser showing the three essential components of a mass spectrometer; the ion source, mass analyser (flight tube) and detector. The anode is used to propel the ions into the flight tube and the lens to focus their flight path.*

In order for mass-to-charge to be determined accurately it is essential that all ions from the source are transferred into the mass analyser at a known time i.e.  $t = 0$ . The first generation of ToF analysers involved passing the ions directly into the detector, which resulted in mass spectra with poor resolution. Poor peak resolution is more pronounced with high molecular weight species due to longer flight times and is overcome by the use of a reflectron lens as initially suggested by Mamyrin in 1966<sup>[26]</sup>. The resolution within ToF instruments is dependent on the ability to measure small differences in the time-of-flight of ions of similar mass-to-charge. Large ions of longer flight times will have an amplification of any disparity in ion energies and flight time resulting in peaks of poor resolution. The reflectron is an electrostatic lens which folds the ion beam and 'corrects' for any difference in the energy of ions of the same mass-to-charge. The ion energy dictates its path length within the lens, where ions of high energy will have a longer ion path than those of low energy. This device

causes ions of the same mass-to-charge to arrive at the detector with the same time-of-flight, minimising any small variations in flight time and resulting in greater peak resolution. Some modern high resolution ToF analysers use the reflectron to analyse at improved resolution large molecular weight species in conjunction with MALDI<sup>[45]</sup>, although this is often achieved with a loss in sensitivity.

#### 1.8.4.4. LTQ Orbitrap mass analyser

This has been developed as part of a hybrid mass spectrometer consisting of an atmospheric pressure ionisation source, leading into a linear ion trap where ions can be stored and detected. A facility of the Orbitrap design is the acquisition of reproducible accurate mass data. Following the linear ion trap ions can be axially ejected into a device known as the c-trap, where they are compressed into a small packet and injected into the Orbitrap mass analyser. In this mass analyser ions are electrostatically contained whilst rotating around a central electrode and causing them to oscillate in an axial dimension. This oscillation induces an image current in the two apices of the Orbitrap that is amplified and detected as a time carrying current (or voltage) signal. The mass-to-charge of an ion may be determined by the following formula:

$$\omega = \frac{k}{m/z} \quad \text{Equation 1.17}$$

$\omega$  = axial oscillation frequency

$k$  = instrumental constant

$m$  = mass of ion

$z$  = number of charges on ion

The time-domain signal of the ions within the Orbitrap can be very complex and require additional elucidation. This is achieved by applying a Fourier transformation calculation which converts time domain to frequency domain from which the masses are deduced using equation 1.17 at high accuracy and peak resolution. The trap design of this analyser with an automatic gain control function enables highly reproducible accurate mass measurements by ensuring a minimum quantity of ions required to generate the appropriate signal are present. In addition to elemental formula determination this instrument enables good separation and detection of low intensity isotopes and multiple stages of fragmentation to be carried out.

### **1.9 Project Brief and Hypothesis**

Past literature has described identifying toxins that contribute to the uremic condition of patients with renal insufficiency by using mass spectrometry<sup>[14]</sup>. However, these analyses involved biological matrices other than hemodialysate, which is essentially the ultrafiltrate of the dialysis procedure after exposure to the patients' blood. It was expected that this sample would require less preparation than the more common matrices of blood and serum. The molecular weight cut-off point of the dialysis membrane was thought to limit the presence of interferences that can affect analysis, such as albumin. In addition, dialysate is generally easy to obtain as it is available per patient in large volumes and has less stringent ethical considerations as it is considered as 'waste'. The aim of this project is to identify novel biomarker(s) that are uremic solutes suitable for assessing the efficacy of the haemodialysate procedure within dialysate solution by mass spectrometry. This is to be achieved by comparing the patients' dialysate throughout haemodialysis treatment, with the identification of novel analytes and an examination of their excretory behaviour.



## References

1. Satko, SG., Freedman, BI., and Moossavi, S., *Kidney International (Supplement)*, 2005. **67**: p. S46-S49.
2. Depner, TA., *Haemodialysis International*, 2005. **9**: p. 241-254.
3. Vanholder, RC., De Smet, R., and Lameire, NH., *Contributions to Nephrology*, 2001. **133**: p. 42-70.
4. Ashley, C., *Hospital Pharmacist*, 2004. **11**: p.54-61.
5. Vanholder, RC., and Ringoir, SM., *Kidney International*, 1992. **42(3)**: p. 540-558.
6. Hewitt, SM., Dear, J., and Star, RA., *Journal of the American Society of Nephrology*, 2004. **15**: p. 1677-1689.
7. Werner Zolg, J., and Langen, H., *Molecular and cellular proteomics*, 2004. **3(4)**: p. 345-354.
8. Hutchens, TW., and Yip, TT., *Rapid Communications in Mass Spectrometry*, 1993. **7(7)**: p. 576-580.
9. Gershwin, ME., Yip, TT., Vandewater, J., Hutchens, TW., *Hepatology*, 1995. **22(4)**: p.57-57.
10. Feilner, T., Kreutzberger, J., Niemann, B., Kramer, A., Possling, A., Seitz, H., Kersten, B., *Current Proteomics*, 2004. **1**: p. 283-295.
11. Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, MJ., Sauter, G., and Kallioniemi, OP., *Nature Medicine*, 1998. **4**: p. 844-847.
12. Lill, J., *Mass spectrometry reviews*, 2003. **22**: p.182-194.
13. Berndt, P., Hobohm, U., Langen, H., *Electrophoresis*, 1999. **20**: p. 3521-3526.
14. Niwa, T., *Mass Spectrometry Reviews*, 1998. **16**: p. 307-332.
15. Moors, M, Massart, DL., and McDowall, RD., *Pure Applied Chemistry*, 1994. **66(2)**: p. 277-304.
16. Aguilar, M-I., *Methods in Molecular Biology: HPLC of Peptides and Proteins - Methods and Protocols*, 2004. **251**: p. 23-35, 55-66, 89-92, 103-104.
17. Heftmann, E., *Chromatography: Fundamentals and Applications of Chromatography and Related differential Migration Methods*, 2004. Elsevier. **6**: p. 2-9, 47-53.
18. Snyder, LR., *Journal of Chromatographic Science*, 1972. **10(4)**: p. 200.

19. Snyder, LR. *Journal of Chromatographic Science*, 1972. 10(6): p. 369.
20. Thomson, JJ., *Rays of positive electricity and their application to chemical analysis*, 1913. Longmans, Green and Co. Ltd.
21. Aston, FW., *Mass spectra and isotopes*. Edward and Arnold and Co. 1942.
22. Dempster, AJ., *Physical Review*, 1918. 11: p.316-325.
23. Nier, AO., *Review of scientific instruments*, 1940. 11: p.212-216.
24. Cameron, AE., and Eggers, DF., *Review of scientific instruments*, 1948. 19(9): p. 605-607.
25. Stephens, WE., *Physical Review*, 1946. 69: p.691.
26. Tang, X., Beavis, R., Ens, W., Lafortune, F., Schueler, B., Standing, KG., *International Journal of Mass Spectrometry and Ion Processes*, 1988. 85: p. 43-67.
27. Karataev, VI., Mamyrin, BA., and Smikk, DV., *Soviet Physics-Technical Physics*, 1972. 16: p.1177.
28. Comisarow, MB., and Marshall, AG., *Chemical Physics Letters*, 1974. 25: p. 282-283.
29. Horning, EC., Carroll, DI., Dzidic, I., Haegele, KD., Horning, MD., Stillwell, RN., *Journal of Chromatographic Science*, 1974. 12: p. 725-729.
30. Arpino, PJ., Baldwin, MA., and McLafferty, FW., *Biological Mass Spectrometry*, 1974. 1(1): p. 80-82.
31. Kellner, R., Mermet, J-M., Otto, M., and Widmer, HM., *Analytical Chemistry*. Wiley VCH, 1998.
32. Whitehouse, CM., Dreyer, RN., Yamashita, M., and Fenn, JB., *Analytical Chemistry*, 1985. 57: p. 675-679.
33. deHoffmann, E. and Stroobant, V., *Mass Spectrometry: Principles and Applications*, 2003: p. 14-22, 34-35, 100-102, 214-238.
34. Dole, M., Hines, RL., Mack, LL., Mobley, RC., Ferguson, LD., Alice, MB., *Journal of Chemical Physics*, 1968. 49: p. 2240-2249.
35. Gaskell, S., *Journal of Mass Spectrometry*, 1997. 32: p. 677-688.
36. Kebarle, P., *Journal of Mass Spectrometry*, 2000. 35: p. 804-817.
37. Iribarne, JV., and Thomson, BA., *Journal of Chemical Physics*, 1976. 64(6): p.2287-2294.
38. Thomson, BA., Iribarne, JV., and Dziedzic, PJ., *Analytical Chemistry*, 1982. 54: p. 2219-2224.

39. King, R., Bonfiglio, R., Fernandez-Metzler, C., Miller-Stein, C., and Olah, T., *Journal of American Society for Mass Spectrometry*, 2000. **11**: p. 942-950.
40. Zhang, J. and Zenobi, R., *Journal of Mass Spectrometry*, 2004. **39**(7): p. 808-816.
41. Zhang, B., McDonald, C., and Li, L., *Analytical Chemistry*, 2004. **76**(4): p. 992-1001.
42. Todd, JFJ., *Mass Spectrometry Reviews*, 1991. **10**(1): p. 3-52.
43. Stafford, JG., *Journal of American Society for Mass Spectrometry*, 2002. **13**(6): p. 589-596.
44. March, RE., *Journal of Mass Spectrometry*, 1997. **32**: p. 351-369.
45. Chernushevich, IV., Loboda, AV., and Thomson, BA., *Journal of Mass Spectrometry*, 2001. **36**: p. 849-865.

## CHAPTER 2:

### Materials and Instrumentation

#### 2.1. Chemicals

Solvents, standard reference materials and gases used are shown in table 2.1.

Chemical	Grade	Supplier
Methanol	HPLC	Thermo Fisher Scientific
Deionised water (DI)	Mill-Q water	Milli-Q purification system (Millipore, USA)
Formic acid		BDH
Ethanol	HPLC	Thermo Fisher Scientific
Trifluoroacetic acid		Thermo Fisher Scientific
Phosphate buffered saline		Thermo Fisher Scientific
'5X' AntigenPlus retrieval buffer		Novagen
Xylene		Thermo Fisher Scientific
Nitrogen (OFN)	Oxygen free	BOC
Helium	99.99%	BOC
Air	99.99%	BOC
Adenosine monophosphate		Sigma
$\alpha$ -cyano-4-hydroxycinnamic acid		Sigma
Angiotensin I		Sigma
Creatinine		Sigma
Hypoxanthine		Sigma
Indole-3-acetic acid		Sigma
$\beta$ -guanidinopropionic acid		Sigma
N- $\alpha$ -acetylarginine		Sigma
N-acetyltryptophan		Sigma
Uric acid		Sigma
<i>Myo</i> -inositol		Sigma
Xanthosine		Sigma
Hydroquinone		Sigma
2,5-dihydroxybenzoic acid		Sigma
Sinapinic acid		Sigma

*Table 2.1: Table containing materials and associated suppliers used within this project.*

## 2.2. Instrumentation

### 2.2.1 Chromatographic equipment

#### 2.2.1.1 UV screening of crude pooled dialysate samples

Agilent (Santa Clara, CA, USA) 8453UV/Vis Spectrophotometer including; diode array detector capable of fast scanning over an absorbance range of 190-1100nm.

#### 2.2.1.2 Size exclusion

Applied Biosystems Ltd (Foster City, CA, USA) *Vision* multi-dimensional LC system including; UV, conductivity and pH detectors and pump was used with Biobasic SEC-60 (300 x 7.8mm i.d., 5 $\mu$ m, 60Å) size exclusion column (Thermo Fisher Scientific, Waltham, MA, USA).

#### 2.2.1.3 Reverse phase HPLC

Dionex/LC Packings (Dionex, Sunnyvale, CA, USA) Ultimate HPLC system including; Famos autosampler unit and Ultimate gradient pumping system was used throughout this research. The column used was a C18 PepMap<sup>TM</sup> reverse phase HPLC column (25cm x 1000 $\mu$ m i.d., 5 $\mu$ m) Dionex and 5 $\mu$ L injection loop. This system was run under micro-flow rate conditions using the MIC-1000 cartridge and 75 $\mu$ m connective tubing.

### 2.2.2 Chromatographic conditions

#### 2.2.2.1 Size exclusion

This LC system used a Biobasic SEC-60 (300 x 7.8mm i.d., 5 $\mu$ m, 60Å) size exclusion column (Thermo Fisher Scientific) at room temperature. A mobile phase consisting of deionised water only was run isocratically at a flow rate of 0.5mL/min over a

period of 1 hour. 20 $\mu$ L of sample was applied to the column using a full-loop injection method.

#### 2.2.2.2 Reverse phase HPLC

This involved a C18 PepMap<sup>TM</sup> reverse phase HPLC column (25cm x 1000 $\mu$ m i.d., 5 $\mu$ m) Dionex at room temperature. A gradient elution system of mobile phase A: 0.1% formic acid in water, and B: 100% methanol was used at a flow rate of 30 $\mu$ L/min. The gradient elution profile is shown below:

Time (min)	%A	%B
0	95	5
5	95	5
10	86	14
15	77	23
20	68	32
25	59	41
30	50	50
35	41	59
40	32	68
45	23	77
50	14	86
55	5	95
60	0	100
83	0	100
88	25	75
93	75	25
98	95	5
118	95	5

*Table 2.2: Table containing the mobile phase composition for the reverse phase chromatographic system.*

A sample volume of 5 $\mu$ L was injected onto the column using an injection programme to ensure accurate sample volumes per injection.

#### 2.2.2.2.1. Injection programme

This involved drawing up 9 $\mu$ L of sample at low syringe speed into the 5 $\mu$ L loop and at a height of 3mm from the base of the autosampler vial. This is left for 5 seconds to distribute homogeneously within the loop and then injected onto the column. The sampler needle is subsequently washed with 10 times the needle volume (500 $\mu$ L) of 100% Milli-Q deionised water (Millipore, Billerica, MA, USA).

### 2.2.3 Mass spectrometers

#### 2.2.3.1 Ion trap

LCQ ion trap (Thermo Fisher Scientific) equipped with an ESI source. Operated in positive mode over a mass-to-charge range of 50-2000 Th unless otherwise stated and the conditions were used as described in the analytical methodology sections. The software used was Xcalibur (v.1.3).

#### 2.2.3.2 LTQ Orbitrap

LTQ Orbitrap (Thermo Fisher Scientific) hybrid mass spectrometer equipped with an ESI source was used in positive mode over a mass-to-charge range of 5-500 Th unless otherwise stated in the analytical methodology. This more contemporary mass spectrometer used Xcalibur (v. 2.0), capable of generating elemental formulae from the calculated accurate masses. Accurate mass and elemental formula assignments for masses below 400 Da were made from the search criteria in table 2.3. A parent ion elemental formula was chosen according to the lowest error from the measured accurate mass and its agreement with the elemental formula obtained for the fragment ions. Fragment ions of the unknowns were then searched again using narrowed element parameters according to the parent ion elemental formula to obtain the lowest

error for the measured accurate mass. All accurate masses were obtained within an error of 5ppm unless stated otherwise in the relevant chapters and the neutral exact mass values are calculated in error by the mass of an electron.

<b>Element</b>	<b>Number Included in Search</b>
C	30
H	60
N	15
O	15
S	4
P	4
Na	2

*Table 2.3: Elemental composition information used in assigning unknown ions.*

#### 2.2.3.3. Voyager DE-STR

All MALDI-ToF analyses were carried out on a Voyager DE-STR (Applied Biosystems, Foster City, CA, USA) mass spectrometer. Four instrument files for both positive and negative ionisation modes were created for mass ranges of 50-2kDa, 1-7kDa, 2k-20kDa, and 5-100kDa. Each mass range was calibrated using the same instrument files and appropriate Sequazyme calibration mix (Applied Biosystems). The low mass gate function was activated to minimize the amount of data collected and set at a mass that defined the lowest limit of the mass spectra reported.



2.2.4 Mass spectrometric conditions2.2.4.1. Ion trap methods2.2.4.1.1. Full MS scan

<b>Instrument Parameter</b>	<b>Reading</b>
MS Run Time (minutes)	118.00
Scan Events	1.00
Scan Event Details	Full MS Scan ( $m/z$ 50-2000)
Capillary Temperature ( $^{\circ}\text{C}$ )	200.00
AGC (Automatic Gain Control)	On
Sheath Gas Flow	40.00
Auxillary/Sweep Gas Flow	10.00
Source Voltage (kV)	3.50
Source Current ( $\mu\text{A}$ )	80.00
Capillary Voltage (V)	23.00
Tube Lens Offset (V)	30.00
Multipole RF Amplifier (Vp-p)	400.00
Multipole 1 Offset (V)	-7.75
Multipole 2 Offset (V)	-9.50
Inter-Multipole Lens Voltage (V)	-16.00
Entrance Lens (V)	-40.00
Trap DC Offset Voltage (V)	-10.00
Full Micro Scans	3.00
Full Max Ion Time (ms)	50.00

*Table 2.4: Table illustrating the instrument conditions used with the LCQ DECA ion trap mass spectrometer operated during full mass scan positive mode.*

### 2.2.4.1.2. Fragmentation (MS/MS) methods

Instrument Parameter	Reading
MS Run Time (minutes)	118.00
Scan Events	4.00
Scan Event Details	<ol style="list-style-type: none"> <li>1. MS/MS Scan of chosen ion</li> <li>2. DDA MS<sup>n</sup> Scan (most intense ion chosen from MS/MS scan)</li> <li>3. MS/MS Scan of next chosen ion</li> <li>4. DDA MS<sup>n</sup> Scan (most intense ion from MS/MS scan)</li> </ol>
MS/MS Settings:	
Isolation Width	1.00
Normalized Collision Energy	35.00
Activation Q	0.25
Activation Time	30.00
Data Dependent Analysis (DDA) Settings:	
Mass Exclusion List	None
Isolation Width	2.00
Normalized Collision Energy	35.0
Activation Q	0.25
Activation Time	30.00
Minimum Signal Required	100000
Minimum MS <sup>n</sup> Signal Required	5000
Dynamic Exclusion Repeat Count	2.00
Dynamic Exclusion Repeat Duration	1.00
Dynamic Exclusion List Size	50
Dynamic Exclusion Mass Width Low	0.50
Dynamic Exclusion Mass Width High	0.50
Capillary Temperature (°C)	200.00
AGC (Automatic Gain Control)	On
Sheath Gas Flow	40.00
Auxillary/Sweep Gas Flow	10.00
Source Voltage (kV)	3.50
Source Current (μA)	80.00
Capillary Voltage (V)	23.00
Tube Lens Offset (V)	30.00
Multipole RF Amplifier (Vp-p)	400.00
Multipole 1 Offset (V)	-7.75
Multipole 2 Offset (V)	-9.50
Inter-Multipole Lens Voltage (V)	-16.00
Entrance Lens (V)	-40.00
Trap DC Offset Voltage (V)	-10.00
Full Micro Scans	3.00
Full Max Ion Time (ms)	50.00

Table 2.5: Table displaying the instrument parameters of the LCQ DECA ion trap operated during MS/MS fragmentation experiments.

### 2.2.4.1.3. Data dependent fragmentation (MS<sup>n</sup>) methods

Instrument Parameter	Reading
MS Run Time (minutes)	118.00
Scan Events	3.00
Scan Event Details	1. Full MS Scan ( <i>m/z</i> 50-2000) 2. DDA MS/MS Scan (most intense ion from full scan) 3. DDA MS <sup>n</sup> Scan (most intense ion from MS/MS scan)
Data Dependent Analysis (DDA) Settings:	
Mass Exclusion List	None
Default Isolation Width	1.00
Normalized Collision Energy	35.0
Activation Q	0.25
Activation Time	30.00
Minimum Signal Required	100000
Minimum MS <sup>n</sup> Signal Required	5000
Dynamic Exclusion Repeat Count	2.00
Dynamic Exclusion Repeat Duration	1.00
Dynamic Exclusion List Size	50
Dynamic Exclusion Mass Width Low	1.50
Dynamic Exclusion Mass Width High	1.50
Capillary Temperature (°C)	200.00
AGC (Automatic Gain Control)	On
Sheath Gas Flow	40.00
Auxillary/Sweep Gas Flow	10.00
Source Voltage (kV)	3.50
Source Current (μA)	80.00
Capillary Voltage (V)	23.00
Tube Lens Offset (V)	30.00
Multipole RF Amplifier (Vp-p)	400.00
Multipole 1 Offset (V)	-7.75
Multipole 2 Offset (V)	-9.50
Inter-Multipole Lens Voltage (V)	-16.00
Entrance Lens (V)	-40.00
Trap DC Offset Voltage (V)	-10.00
Full Micro Scans	3.00
Full Max Ion Time (ms)	50.00

*Table 2.6: Table containing the instrument parameters of the LCQ DECA ion trap operated during data dependent fragmentation experiments.*

## 2.2.4.2. Orbitrap methods

2.2.4.2.1. Full mass and data dependent scan: uremic analytes 1, 2, 3 at  $m/z$  214, 241,

275.

Instrument Parameter	Reading
MS Run Time (minutes)	85.00
Segments	2.00
Segment 1	0-10min
Scan Events	2.00
Scan Event Details	1. Full MS Scan ( $m/z$ 150-300) 2. MS/MS DDA parent ion list of 214.1298 and 241. If parent ion not present fragment nth most intense ion @ CE25%, Q = 0.25, IsoW = 3.0
Segment 2	30-50min
Scan Events	2
Scan Event Details	1. Full MS Scan ( $m/z$ 150-300) 2. MS/MS 275.2 @ CE45%, Q = 0.25, IsoW = 3.0
Capillary Temperature (°C)	275.00
AGC (Automatic Gain Control)	On
Sheath Gas Flow	10.00
Auxillary/Sweep Gas Flow	5.00
Source Voltage (kV)	5.00
Source Current ( $\mu$ A)	100.00
Capillary Voltage (V)	42.00

*Table 2.7: Table containing instrument parameters for the LTQ Orbitrap mass spectrometer operated in positive mode. These were used when acquiring both full mass scan data and data dependent fragmentation information for the novel uremic analytes present in size exclusion 2.*

2.2.4.2.2. Full mass and data dependent scan: uremic analytes 4, 5, 6 at  $m/z$  270, 381(359, 180), 335.

Instrument Parameter	Reading
MS Run Time (minutes)	85.00
Segments	3.00
Segment 1	0-10min
Scan Events	3.00
Scan Event Details	1. Full MS Scan ( $m/z$ 150-400) 2. MS/MS 270.1563 ( $m/z$ 55-300) @ CE35%, Q = 0.25, IsoW = 3.0 3. MS2 270.1563 → 253.1296 ( $m/z$ 65-300) @ CE35%, Q=0.25, IsoW=3.0 MS3 @ CE30%, Q=0.25, IsoW=3.0
Segment 2	20-35min
Scan Events	2
Scan Event Details	1. Full MS Scan ( $m/z$ 150-400) 2. MS/MS 335 ( $m/z$ 100-400) @ CE28%, Q=0.25, IsoW=3.0
Segment 3	35-45min
Scan Events	2
Scan Event Details	1. Full MS Scan ( $m/z$ 150-400) 2. MS/MS 380.8 ( $m/z$ 100-400) @ CE28%, Q=0.25, IsoW=3.0 3. MS/MS 358.8 ( $m/z$ 100-400) @ CE28%, Q=0.25, IsoW=3.0 4. MS/MS 179.9 ( $m/z$ 100-400) @ CE28%, Q=0.25, IsoW=3.0
Capillary Temperature (°C)	275.00
AGC (Automatic Gain Control)	On
Sheath Gas Flow	10.00
Auxillary/Sweep Gas Flow	5.00
Source Voltage (kV)	5.00
Source Current (μA)	100.00
Capillary Voltage (V)	42.00

*Table 2.8: Table containing instrument parameters for the LTQ Orbitrap mass spectrometer operated in positive mode. These were used when acquiring both full mass scan data and data dependent fragmentation information for the novel uremic analytes present in size exclusion 4.*

2.2.4.2.3. Mass targeted fragmentation scan: uremic analytes 1, 2, 3, at  $m/z$  214, 241,275.

Instrument Parameter	Reading
MS Run Time (minutes)	85.00
Segments	2.00
Segment 1	0-10min
Scan Events	4.00
Scan Event Details	1. Full MS Scan ( $m/z$ 150-300) 2. MS/MS 214.1298 ( $m/z$ 55-300) @ CE25%, Q = 0.25, IsoW = 3.0 3. MS3 214.1298 → 197.1033 ( $m/z$ 55-250) @ CE25%, Q = 0.25, IsoW = 3.0 4. MS/MS 240.95 ( $m/z$ 55-300) @ CE25%, Q = 0.25, IsoW = 3.0
Segment 2	30-50min
Scan Events	2
Scan Event Details	1. Full MS Scan ( $m/z$ 150-300) 2. MS/MS 275.2 CE45%, Q = 0.25, IsoW = 3.0
Capillary Temperature (°C)	275.00
AGC	On
Sheath Gas Flow	10.00
Auxillary/Sweep Gas Flow	5.00
Source Voltage (kV)	5.00
Source Current ( $\mu$ A)	100.00
Capillary Voltage (V)	42.00

*Table 2.9: Table containing instrument parameters for the LTQ Orbitrap mass spectrometer operated in positive mode. These were used when acquiring targeted mass-to-charge fragmentation information for the novel uremic analytes present in size exclusion 2.*

2.2.4.2.4. Mass targeted fragmentation scan: uremic analytes 4, 5, 6 at  $m/z$  270, 381

(359, 180), 335.

Instrument Parameter	Reading
MS Run Time (minutes)	85.00
Segments	3.00
Segment 1	0-10min
Scan Events	5.00
Scan Event Details	1. Full MS Scan ( $m/z$ 150-400) 2. MS/MS 270.1563 ( $m/z$ 55-300) @ CE35%, Q = 0.25, IsoW = 3.0 3. MS2 270.1563 → 253.1296 ( $m/z$ 65-300) @ CE35%, Q=0.25, IsoW=3.0 MS3 @ CE30%, Q=0.25, IsoW=3.0 4. MS2 270.1563 → 210.1236 ( $m/z$ 55-300) @ CE35%, Q=0.25, IsoW=3.0 MS3 @ CE30%, Q=0.25, IsoW=3.0 5. MS2 270.1563 → 253.1296 → 225.1345 ( $m/z$ 55-300) @ CE35%, Q=0.25, IsoW=3.0 MS3 @ CE30%, Q=0.25, IsoW=3.0 MS4 @ CE30%, Q=0.25, IsoW=3.0
Segment 2	20-35min
Scan Events	3
Scan Event Details	1. Full MS Scan ( $m/z$ 150-400) 2. MS/MS 335 ( $m/z$ 100-400) @ CE28%, Q=0.25, IsoW=3.0
Segment 3	35-45min
Scan Events	3
Scan Event Details	1. Full MS Scan ( $m/z$ 150-400) 2. MS/MS 380 ( $m/z$ 100-400) @ CE28%, Q=0.25, IsoW=3.0 3. MS/MS 358.8 ( $m/z$ 100-400) @ CE28%, Q=0.25, IsoW=3.0 4. MS/MS 179.9 ( $m/z$ 100-400) @ CE28%, Q=0.25, IsoW=3.0
Capillary Temperature (°C)	275.00
AGC (Automatic Gain Control)	On
Sheath Gas Flow	10.00
Auxillary/Sweep Gas Flow	5.00
Source Voltage (kV)	5.00
Source Current (μA)	100.00
Capillary Voltage (V)	42.00

*Table 2.10: Table containing instrument parameters for the LTQ Orbitrap mass spectrometer operated in positive mode. These were used when acquiring targeted mass-to-charge fragmentation information for the novel uremic analytes present in size exclusion 4.*

2.2.4.3. MALDI-ToF methods2.2.4.3.1. Mass range 50-2000Da

<b>Instrument Settings</b>	<b>Positive Mode</b>	<b>Negative Mode</b>
<b>Mode of operation</b>	Reflector	Reflector
<b>Extraction mode</b>	Delayed	Delayed
<b>Accelerating voltage (V)</b>	20000	20000
<b>Grid voltage</b>	68%	68%
<b>Mirror voltage ratio</b>	1.12	1.12
<b>Extraction delay time</b>	200nsec	200nsec
<b>No. laser shots/spectrum</b>	50	50
<b>Total no. laser shots</b>	150	150
<b>Laser repetition rate</b>	20Hz	20Hz
<b>Vertical scale 0</b>	500mV	500mV
<b>Vertical offset</b>	0.35%	0.35%

*Table 2.11: Instrument parameters specific for measuring ions within the mass range of 50-2000Da.*

2.2.4.3.2. Mass range 1k-7kDa

<b>Instrument Settings</b>	<b>Positive Mode</b>	<b>Negative Mode</b>
<b>Mode of operation</b>	Reflector	Reflector
<b>Extraction mode</b>	Delayed	Delayed
<b>Accelerating voltage (V)</b>	20000	20000
<b>Grid voltage</b>	68%	68%
<b>Mirror voltage ratio</b>	1.12	1.12
<b>Extraction delay time</b>	200nsec	200nsec
<b>No. laser shots/spectrum</b>	50	50
<b>Total no. laser shots</b>	150	150
<b>Laser repetition rate</b>	20Hz	20Hz
<b>Vertical scale 0</b>	500mV	500mV
<b>Vertical offset</b>	0.35%	0.35%

*Table 2.12: Instrument parameters specific for measuring ions within the mass range of 1000-7000Da.*



2.2.4.3.3. Mass range 2k-20kDa

<b>Instrument Settings</b>	<b>Positive Mode</b>	<b>Negative Mode</b>
<b>Mode of operation</b>	Linear	Linear
<b>Extraction mode</b>	Delayed	Delayed
<b>Accelerating voltage (V)</b>	25000	25000
<b>Grid voltage</b>	90%	90%
<b>Mirror voltage ratio</b>	-	-
<b>Extraction delay time</b>	400nsec	400nsec
<b>No. laser shots/spectrum</b>	60	60
<b>Total no. laser shots</b>	180	180
<b>Laser repetition rate</b>	20Hz	20Hz
<b>Vertical scale 0</b>	50mV	50mV
<b>Vertical offset</b>	0.50%	0.50%

*Table 2.13: Instrument parameters specific for measuring ions within the mass range of 2000-20,000Da.*

2.2.4.3.4. Mass range 5k-100kDa

<b>Instrument Settings</b>	<b>Positive Mode</b>	<b>Negative Mode</b>
<b>Mode of operation</b>	Linear	Linear
<b>Extraction mode</b>	Delayed	Delayed
<b>Accelerating voltage (V)</b>	25000	25000
<b>Grid voltage</b>	90%	90%
<b>Mirror voltage ratio</b>	-	-
<b>Extraction delay time</b>	750nsec	750nsec
<b>No. laser shots/spectrum</b>	100	100
<b>Total no. laser shots</b>	300	300
<b>Laser repetition rate</b>	20Hz	20Hz
<b>Vertical scale 0</b>	50mV	50mV
<b>Vertical offset</b>	0.50%	0.50%

*Table 2.14: Instrument parameters specific for measuring ions within the mass range of 5000-100,000Da.*

### 2.2.5 Laboratory apparatus

- Scientific Industries (Bohemia, NY, USA) *Vortex Genie 2* mixer
- HTL (Warsaw, Poland) *LabMate* 1000, 100, 10 $\mu$ L air displacement pipettes
- Elkay UK (Basingstoke, UK) 1000 $\mu$ L air displacement pipette tips
- Bio Medical Laboratory Supplies Ltd (Birmingham, UK) 100 $\mu$ L air displacement pipette tips
- StarLab GMBH (Ahrensburg, Germany) 10 $\mu$ L air displacement pipette tips
- Hamilton (Reno, NV, USA) 100 $\mu$ L spiking syringe
- Hamilton (Reno, NV, USA) 250 $\mu$ L spiking syringe
- Edwards (Crawley, UK) Lyophilizer
- Terumo Medical Corporation (Somerset, NJ, USA) 1mL plastic syringe
- Sartorius AG (Goettingen, Germany) Minisart High-flow single use syringe  
0.2 $\mu$ m non-pyrogenic hydrophilic micro filters
- Eppendorf (Hamburg, Germany) 1.5mL vortex tubes
- Eppendorf (Hamburg, Germany) 0.5mL vortex tubes
- Thermo Fisher Scientific (Waltham, MA, USA) 20mL glass scintillation vials
- SunSri (Brockville, Ontario, Canada) 250 $\mu$ L polypropylene autosampler vial and snap caps
- Hoover (Glenwillow, OH, USA) -20 $^{\circ}$ C chest freezer

## 2.3. Standard Solutions

### 2.3.1. SEC test solutions

#### 2.3.1.1. $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA)

A 10mg/mL solution was made in 50:50 acetonitrile:0.1% trifluoroacetic acid in dionised water.

### 2.3.1.2. Adenosine monophosphate (AMP)

A 1mg/mL solution was made in 50:50 methanol:0.1% formic acid in deionised water.

### 2.3.1.3. Angiotensin I

A 1mg/mL solution was made in 50:50 methanol:0.1% formic acid in deionised water.

### 2.3.2. Known uremic toxin standard solutions

All stock standard solutions were made at 1mg/mL concentration in 0.1% formic acid in deionised water to enable maximum stability through freezing at -20°C. Dilutions were made of each stock solution in 50:50 methanol:0.1% formic acid in deionised water and infused at a flow rate of 8µL/minute into the LCQ DECA mass spectrometer and their respective MS/MS fragmentation mass spectra obtained.

<b>Standard Stock Solution</b>	<b>Diluted Concentration (pmole/µL)</b>
Creatinine	10.0
Hypoxanthine	10.0
Indole-3-acetic acid	8.0
β-guanidinopropionic acid	10.0
N-α-acetylarginine	5.0
N-acetyltryptophan	5.0
Uric acid	5.0
Myo-inositol	8.0
Xanthosine	5.0
Hydroquinone	10.0

*Table 2.15: Table containing standard solutions used to obtain reference MS/MS mass spectra and the relevant concentrations.*

## **CHAPTER 3:**

# **Validation and Stability Investigations of Novel Methodology for Screening Haemodialysate Using LC/MS and LC/MS-MS**

### **3.1. Introduction**

Spent dialysate fluid from 8 renal patients was taken at 5 minutes post-connection (sample 2) and 5 minutes pre-disconnection (sample 3) to the dialysate unit to characterise diseased and healthy states respectively. The collections of 5 minutes post-connection for different patients were pooled into one volume and named sample 2. The same procedure was applied to the 5 minutes pre-disconnection fluid and named sample 3. These were intended to represent the total complement of uremic toxins associated with renal failure before and after treatment, and were lyophilised once frozen. For separation 1/10<sup>th</sup> of the solid samples were reconstituted in the minimum volume of deionised water and filtered using a Millipore (Billerica, MA, USA) 5µm mini-filter. Sample quantities and their subsequent volumes are shown in table 3.1.

<b>Sample Number</b>	<b>Weight of Fraction of Total Sample (g)</b>	<b>Volume of Reconstituted Sample (mL)</b>
2	0.4910	3.0
3	0.4606	4.0

*Table 3.1: Table containing information regarding sample composition for the pooled samples obtained 5 minutes post-connection (sample 2) and 5 minutes pre-disconnection (sample 3) from the dialysis unit.*

The filtered samples were then aliquoted into 50 $\mu$ L volumes to limit sample used and the number of freeze-thaw cycles.

### 3.2. Method Validation 1: Validation of Analytical Protocols for Crude Pooled Samples

#### 3.2.1. UV analysis of crude sample

In order to obtain an accurate UV representation of the crude samples their maximum and minimum absorbencies were determined with a UV spectrophotometer (Agilent, Santa Clara, CA, USA). The sample chromatograms had a maximum absorbance at a wavelength of 242nm. This wavelength was used for subsequent UV analyses and is characteristic of several different chromophoric functional groups from a diazoacetic ester, *cis*-crotonic acid, an unsaturated ketone or a ketone monosubstituted benzene ring<sup>[1]</sup>.

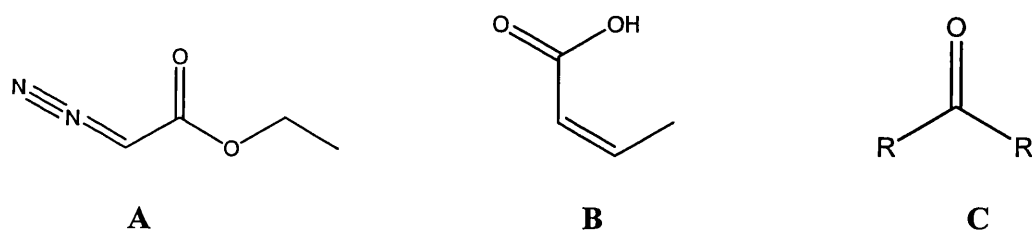


Figure 3.1: Chemical structure of the chromphoric groups that absorb at the UV wavelength of 242 nm; A = diazoacetic ester, B = *cis*-crotonic acid, C = unsaturated ketone, where R contains unsaturated functional groups.

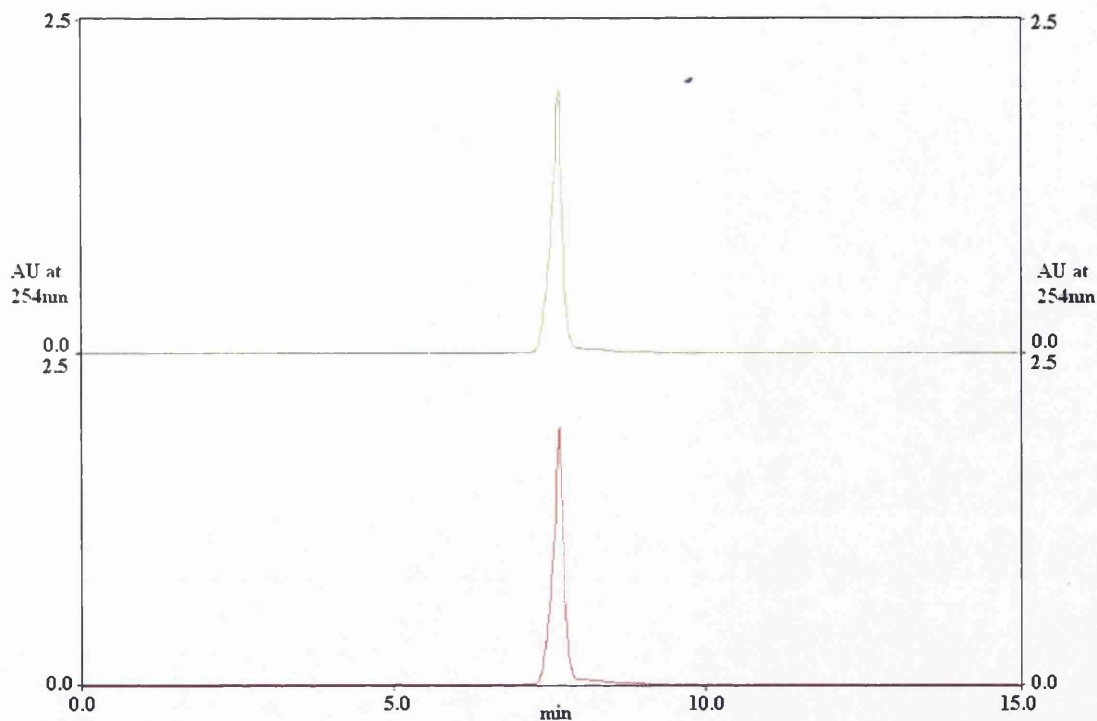
### 3.2.2. Size exclusion chromatography

#### 3.2.2.1. UV absorbance

Approximate retention times for specific molecular masses were determined using several standard solutions. A volume of 20 $\mu$ L of  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA), adenosine monophosphate (AMP), and angiotensin I standards were injected for two successive runs onto the size exclusion column as described in Chapter 2 with their UV absorbance's measured at 254 and 242nm. The observed approximate retention times can be used to predict molecular mass ranges of any subsequent fractions collected of the haemodialysate sample.

<b>Standard Solution</b>	<b>Molecular Weight (Da)</b>	<b>Retention Time (min)</b>
CHCA	189.17	30.5
AMP	347.22	8.5
Angiotensin I	1296.48	5.0

*Table 3.2: Standard reference materials (Sigma Aldrich, MO, USA) used to predict retention time bands within the size exclusion column.*



*Figure 3.2: Overlaid UV chromatograms of the size exclusion standard AMP of molecular weight 347.22Da. The UV peak of absorbance 1.9units at 254nm, for this molecular weight is at a retention time of 8.5 minutes.*

Initially the standards were run at the column manufacturers recommended flow rate of 1mL/min and under these conditions the low molecular weight standards were poorly separated. The reduction of eluent flow rate to 0.5mL/min greatly improved the separation and was used for fractionating the dialysate samples.

Of the crude pooled sample aliquots a volume of 20 $\mu$ L was injected onto a Biobasic SEC-60 (300 x 7.8mm i.d., 5 $\mu$ m, 60 $\text{\AA}$ ) size exclusion column (Thermo Fisher Scientific, Waltham, MA, USA) as part of the Applied Biosystems Ltd (Foster City, CA, USA) *Vision* multi-dimensional LC system. The separation of sample components using the column showed two main UV active sections in both chromatograms (see figure 3.3).

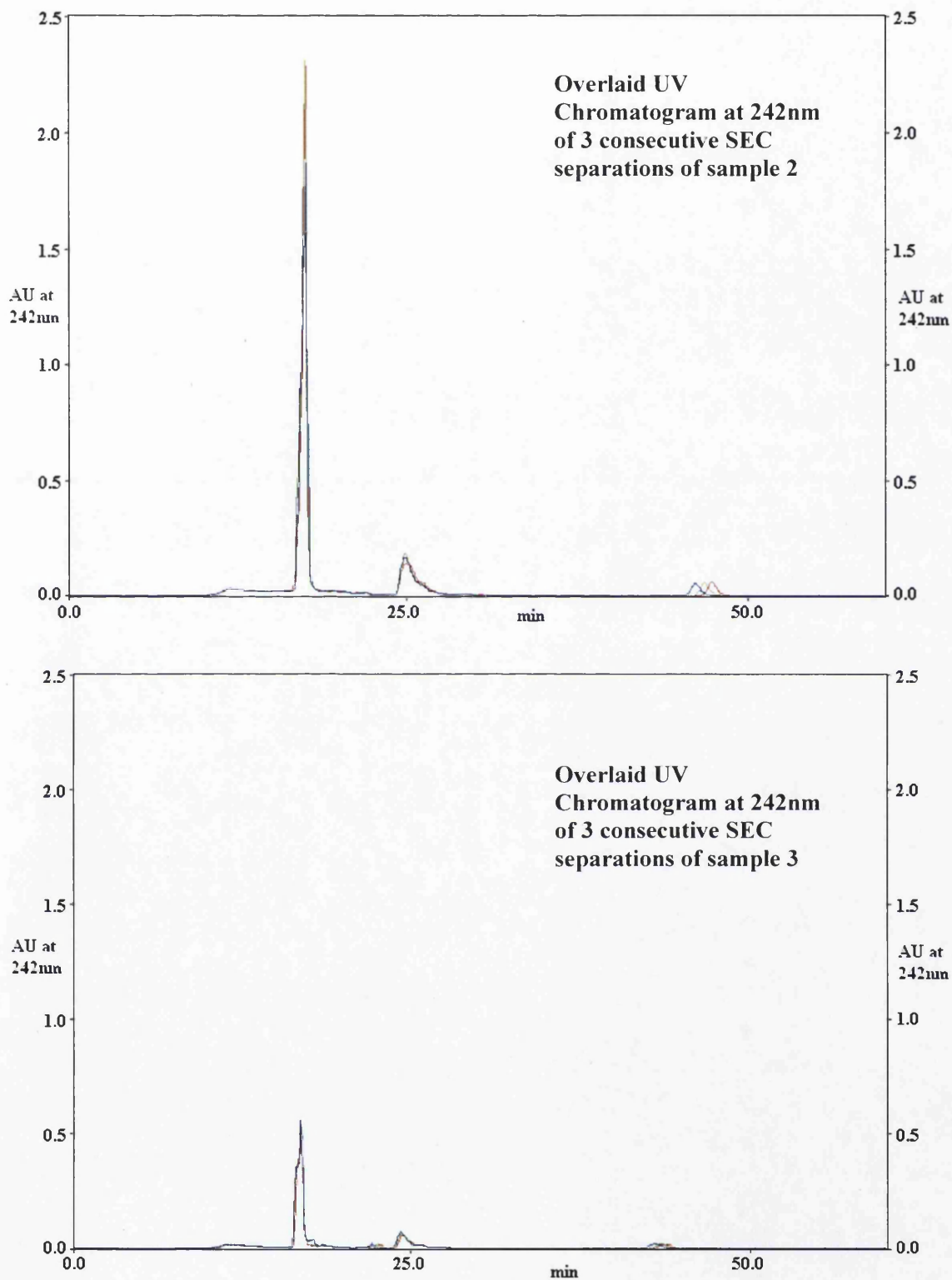


Figure 3.3: Overlaid UV chromatograms of three consecutive SEC separations of sample 2 and 3 at 242nm. Both chromatograms show comparative areas of UV activity supporting the use of the same fractionation procedure.



Fractions of different volumes were collected with the emphasis on obtaining a greater number in the UV active areas and followed the pattern shown in figure 3.4.

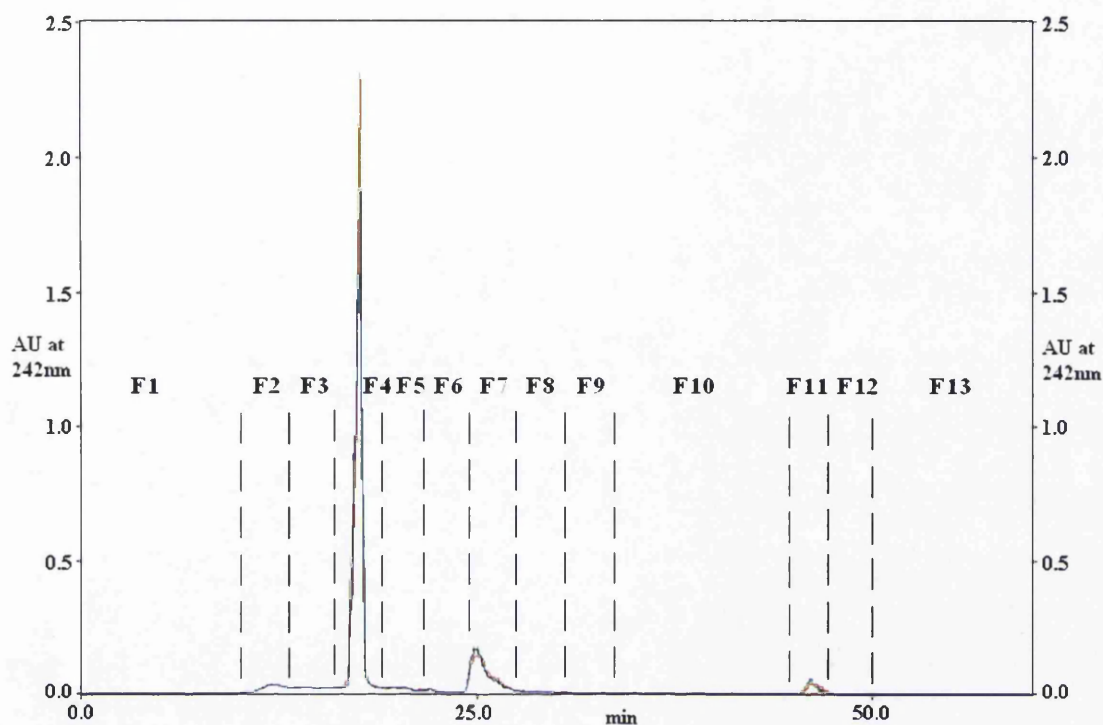


Figure 3.4 : UV chromatogram of sample 2 (S2 – 5 minutes post-connection) indicating the areas of UV activity; F1 = 0-10minutes, F2 = 10-13minutes, F3 = 13-16minutes, F4 = 16-19minutes, F5 = 19-22minutes, F6 = 22-25minutes, F7 = 25-28minutes, F8 = 28-31minutes, F9 = 31-34minutes, F10 = 34-45minutes, F11 = 45-48minutes, F12 = 48-50minutes, and F13 = 50-60minutes.

The fractions were collected over specified time periods and the volume of each is shown in table 3.3.

Fraction	Retention Time Range (min)	Fraction Volume (mL)
F1	0-10	5.0
F2	10-13	1.5
F3	13-16	1.5
F4	16-19	1.5
F5	19-22	1.5
F6	22-25	1.5
F7	25-28	1.5
F8	28-31	1.5
F9	31-34	1.5
F10	34-45	5.5
F11	45-48	1.5
F12	48-50	1.0
F13	50-60	5.0

*Table 3.3: Fraction sizes collected with regards to time elapsed with the corresponding volume during the size exclusion chromatographic run.*

This fraction collection regime was also applied to sample 3 (5 minutes pre-disconnection to the dialysis unit) since it showed a similar separation regarding UV absorbance to sample 2. The fractionation of each sample showed good inter- and intra-reproducibility indicated quite clearly by the following overlaid chromatograms and tabulated data (see figure 3.5 and tables 3.4-3.6). Intra-reproducibility of maximum absorbance retention time for each fraction is represented by a confidence interval value shown in the furthest right hand column of tables 3.4 and 3.5. This statistical test has been carried out with  $P = 0.05$ , and hence we can be 95% confident that the values in the table for each run will not deviate by more than  $\pm 0.634$  minutes from the mean standard deviation for sample 2 separations. This statistical test has also been carried out for sample 3 and show that each run will not deviate more than  $\pm 0.794$  minutes from the mean standard deviation. These results indicate that the

retention times of the fraction maximum absorbance will not differ significantly between separations carried out the same day for both samples.

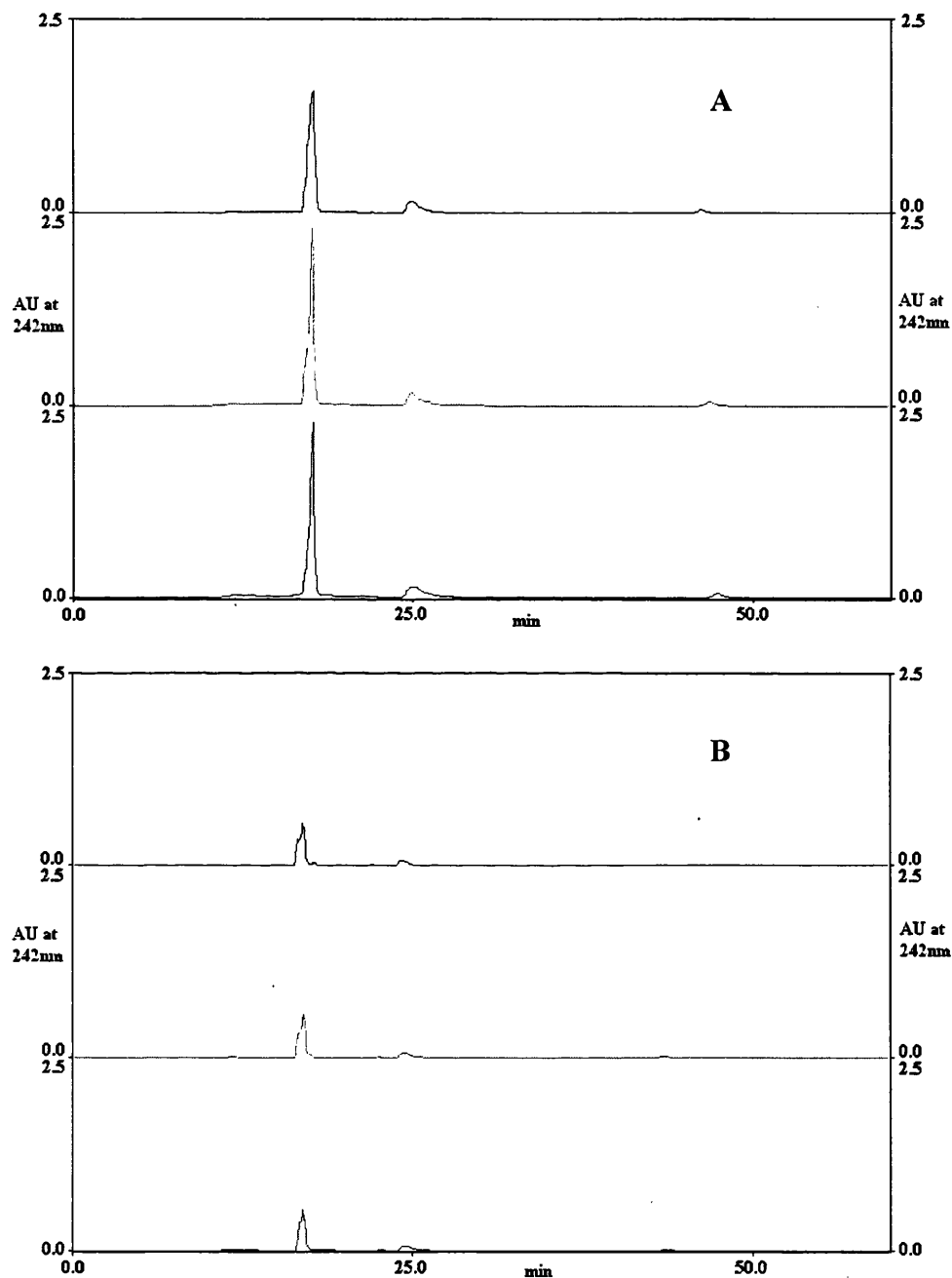


Figure 3.5: A – UV chromatogram at 242nm of multiple runs of sample 2 (5minutes post-connection) using the size exclusion LC system, indicating good reproducibility of the separation with the consistent UV active regions of the chromatogram.

B – UV chromatogram of multiple runs of sample 3 (5minutes pre-disconnection) using size exclusion, illustrating good reproducibility

and significant difference in the observed absorbance when compared to sample 2.

SEC Fraction	Retention Time Reproducibility (in minutes) of Maximum UV Absorbance Associated with a Size Exclusion Fraction for Sample 2 Day 1									Probability Result is Within 5% of Mean
	Intra run 1	Intra run 2	Intra run 3	Intra run 4	Intra run 5	Mean	Std Dev.	Mean Std Dev	Variance	
F1	9.9	9.9	9.9	9.9	9.9	9.9	0.000	0.000	0.000	0.000
F2	12.0	12.0	12.0	12.0	12.1	12.0	0.040	0.032	0.002	0.035
F3	13.0	13.0	13.0	13.0	13.0	13.0	0.000	0.000	0.000	0.000
F4	17.6	17.6	17.6	17.5	17.5	17.6	0.049	0.048	0.003	0.043
F5	19.0	19.7	19.0	19.0	19.0	19.1	0.280	0.224	0.098	0.245
F6	24.9	24.9	24.9	24.9	24.9	24.9	0.000	0.000	0.000	0.000
F7	25.0	25.0	25.0	25.0	25.0	25.0	0.000	0.000	0.000	0.000
F8	28.0	28.4	28.3	28.1	28.0	28.2	0.162	0.152	0.033	0.142
F9	31.0	31.0	31.0	31.0	31.0	31.0	0.000	0.000	0.000	0.000
F10	34.0	34.0	34.0	34.0	34.0	34.0	0.000	0.000	0.000	0.000
F11	47.3	46.8	46.2	45.7	45.3	46.3	0.723	0.632	0.653	0.634
F12	48.0	48.0	48.0	48.0	48.0	48.0	0.000	0.000	0.000	0.000
F13	50.0	50.2	50.1	50.0	50.1	50.1	0.075	0.064	0.007	0.066

SEC Fraction	Retention Time Reproducibility (in minutes) of Maximum UV Absorbance Associated with a Size Exclusion Fraction for Sample 2 Day 2									Probability Result is Within 5% of Mean
	Intra run 1	Intra run 2	Intra run 3	Intra run 4	Intra run 5	Mean	Std Dev.	Mean Std Dev	Variance	
F1	9.9	9.9	9.9	9.9	9.9	9.9	0.000	0.000	0.000	0.000
F2	11.9	11.9	11.8	12.0	11.7	11.9	0.102	0.088	0.013	0.089
F3	13.0	13.0	13.0	13.0	13.0	13.0	0.000	0.000	0.000	0.000
F4	17.4	17.5	17.6	17.4	17.4	17.5	0.080	0.072	0.008	0.070
F5	19.1	19.1	19.2	19.0	19.0	19.1	0.075	0.064	0.007	0.066
F6	24.3	24.4	24.6	24.9	24.7	24.6	0.214	0.184	0.057	0.187
F7	25.5	25.0	25.0	25.0	25.0	25.1	0.200	0.160	0.050	0.175
F8	28.2	28.0	28.1	28.0	28.0	28.1	0.080	0.072	0.008	0.070
F9	31.0	31.0	31.0	31.0	32.1	31.2	0.440	0.352	0.242	0.386
F10	34.9	34.0	34.0	34.0	34.5	34.3	0.366	0.336	0.167	0.320
F11	47.8	47.4	47.0	46.6	45.8	46.9	0.688	0.576	0.592	0.603
F12	48.0	48.0	48.0	48.0	48.0	48.0	0.000	0.000	0.000	0.000
F13	50.0	50.0	50.0	50.0	50.1	50.0	0.040	0.032	0.002	0.035

Table 3.4: Intra-reproducibility data of retention time of the maximum UV absorbance associated with each size exclusion fraction obtained for sample 2 (5 minutes post-connection) during two alternate days. Intra-reproducibility,

*represented by a confidence interval value, has been carried out with  $P = 0.05$ , and hence we can be 95% confident that the values in the table for each run will not deviate more than  $\pm 0.634$  minutes from the mean standard deviation.*

The inter-reproducibility of maximum UV absorbance of the relevant size exclusion fraction has been determined by an F-test calculation (see table 3.6). This indicates that there is no significant difference between the retention times of separations carried out for both samples on the same and alternate days.

SEC Fraction	Retention Time Reproducibility (in minutes) of Maximum UV Absorbance Associated with a Size Exclusion Fraction for Sample 3 Day 1									Probability Result is Within 5% of Mean
	Intra run 1	Intra run 2	Intra run 3	Intra run 4	Intra run 5	Mean	Std Dev.	Mean Std Dev.	Variance	
F1	9.9	9.9	9.9	9.9	9.9	9.9	0.000	0.000	0.000	0.000
F2	11.7	11.7	11.7	11.7	11.6	11.7	0.040	0.032	0.002	0.035
F3	13.0	13.0	13.0	13.0	13.0	13.0	0.000	0.000	0.000	0.000
F4	17.0	17.1	17.1	16.9	16.9	17.0	0.089	0.080	0.010	0.078
F5	19.6	19.5	19.4	19.2	19.0	19.3	0.215	0.192	0.058	0.189
F6	24.6	24.5	24.3	24.2	24.2	24.4	0.162	0.152	0.033	0.142
F7	25.0	25.0	25.0	25.0	25.0	25.0	0.000	0.000	0.000	0.000
F8	28.0	28.0	28.0	28.0	28.0	28.0	0.000	0.000	0.000	0.000
F9	31.7	31.6	31.5	31.6	31.3	32.8	0.532	0.112	0.023	0.466
F10	44.6	44.0	43.4	42.9	42.6	43.5	0.727	0.640	0.660	0.637
F11	45.0	45.0	45.0	45.0	45.0	45.0	0.000	0.000	0.000	0.000
F12	48.0	48.5	48.0	48.0	48.6	48.2	0.271	0.264	0.092	0.238
F13	52.1	51.4	50.5	50.2	50.0	50.8	0.791	0.728	0.783	0.694

SEC Fraction	Retention Time Reproducibility (in minutes) of Maximum UV Absorbance Associated with a Size Exclusion Fraction for Sample 3 Day 2									Probability Result is Within 5% of Mean
	Intra run 1	Intra run 2	Intra run 3	Intra run 4	Intra run 5	Mean	Std Dev.	Mean Std Dev	Variance	
F1	9.9	9.9	9.9	9.9	9.9	9.9	0.000	0.000	0.000	0.000
F2	11.7	11.7	11.7	11.7	11.7	11.7	0.000	0.000	0.000	0.000
F3	13.0	13.0	13.0	13.0	13.0	13.0	0.000	0.000	0.000	0.000
F4	17.0	17.0	16.9	17.0	17.0	17.0	0.040	0.032	0.002	0.035
F5	19.1	19.0	19.0	19.0	19.0	19.0	0.040	0.032	0.002	0.035
F6	24.4	24.4	24.3	24.2	24.3	24.3	0.075	0.064	0.007	0.066
F7	25.0	25.0	25.0	25.0	25.0	25.0	0.000	0.000	0.000	0.000
F8	28.0	28.0	28.0	28.0	28.0	28.0	0.000	0.000	0.000	0.000
F9	32.1	33.4	33.2	33.1	31.0	32.8	0.906	0.808	1.013	0.794
F10	43.7	43.5	43.2	43.0	42.7	43.2	0.354	0.304	0.157	0.311
F11	45.0	45.0	45.0	45.0	45.0	45.0	0.000	0.000	0.000	0.000
F12	48.0	48.1	48.0	48.0	48.0	48.0	0.040	0.032	0.002	0.035
F13	51.5	51.1	50.8	50.4	50.1	50.8	0.496	0.424	0.307	0.434

*Table 3.5: Intra-reproducibility data of the retention time of the maximum UV absorbance associated with each size exclusion fraction obtained for sample 3 (5 minutes pre-disconnection) during two alternate days. Intra-reproducibility is represented by a confidence interval value shown in the furthest right hand column. This statistical test has been carried out with  $P = 0.05$ , and hence we can be 95% confident that the values in the table for each run will not deviate more than  $\pm 0.794$  minutes from the mean standard deviation.*

Inter Reproducibility of Maximum UV Absorbance Associated with a SEC Fraction				Inter Reproducibility of Maximum UV Absorbance Associated with a SEC Fraction			
SEC Fraction	Variance of S2 Fraction Day 1	Variance of S2 Fraction Day 2	Inter Reproducibility by F-Test	SEC Fraction	Variance of S3 Fraction Day 1	Variance of S3 Fraction Day 2	Inter Reproducibility by F-Test
F1	0.000	0.000	0.8291	F1	0.000	0.000	0.8169
F2	0.002	0.013		F2	0.002	0.000	
F3	0.000	0.000		F3	0.000	0.000	
F4	0.003	0.008		F4	0.010	0.002	
F5	0.098	0.007		F5	0.058	0.002	
F6	0.000	0.057		F6	0.033	0.007	
F7	0.000	0.050		F7	0.000	0.000	
F8	0.033	0.008		F8	0.000	0.000	
F9	0.000	0.242		F9	0.023	1.013	
F10	0.000	0.167		F10	0.660	0.157	
F11	0.653	0.592		F11	0.000	0.000	
F12	0.000	0.000		F12	0.092	0.002	
F13	0.007	0.002		F13	0.783	0.307	

Table 3.6: Reproducibility data of the size exclusion separation carried out over different days and is represented by an F-test calculation.

### 3.2.2.2. pH and conductivity data

In addition to obtaining a UV profile of the separation, both pH and conductivity data was acquired for samples 2 and 3. The pH profile of both samples also showed good inter and intra-reproducibility indicating a drop in pH to  $4.21(\pm 0.295)$ . This data implies that the eluted component at this time point is acidic in nature and corresponds to the largest peak in the chromatogram present in the 16-19 minute size exclusion fraction (F4). This effect is also present in the chromatograms obtained for sample 3 and can provide further evidence in supporting the reproducibility of the separation method. In addition, it may also indicate that the same or similar acidic component is present in both samples, although, as indicated by the UV chromatogram, at different concentrations.

Reproducibility of LC Data: Sample 2			Reproducibility of LC Data: Sample 3		
Peak	Variation in pH		Peak	Variation in pH	
	Inter Mean	Intra Mean		Inter Mean	Intra Mean
F1	0.87	0.30	F1	0.25	0.30
F2	0.92	0.29	F2	0.30	0.21
F3	0.35	0.34	F3	0.07	0.18
F4	0.42	0.35	F4	0.10	0.24
F5	0.60	0.30	F5	0.14	0.19
F6	0.67	0.23	F6	0.16	0.16
F7	0.33	0.26	F7	0.02	0.14
F8	0.32	0.25	F8	0.05	0.16
F9	0.48	0.23	F9	0.11	0.16
F10	0.58	0.20	F10	0.16	0.16
F11	0.59	0.18	F11	0.16	0.14
F12	0.60	0.18	F12	0.16	0.14
F13	0.65	0.19	F13	0.17	0.14

*Table 3.7: Inter- and intra-reproducibility data of pH measurements for each size exclusion fraction throughout the chromatographic run for both sample 2 (5 minutes post-connection) and sample 3 (5 minutes pre-disconnection).*



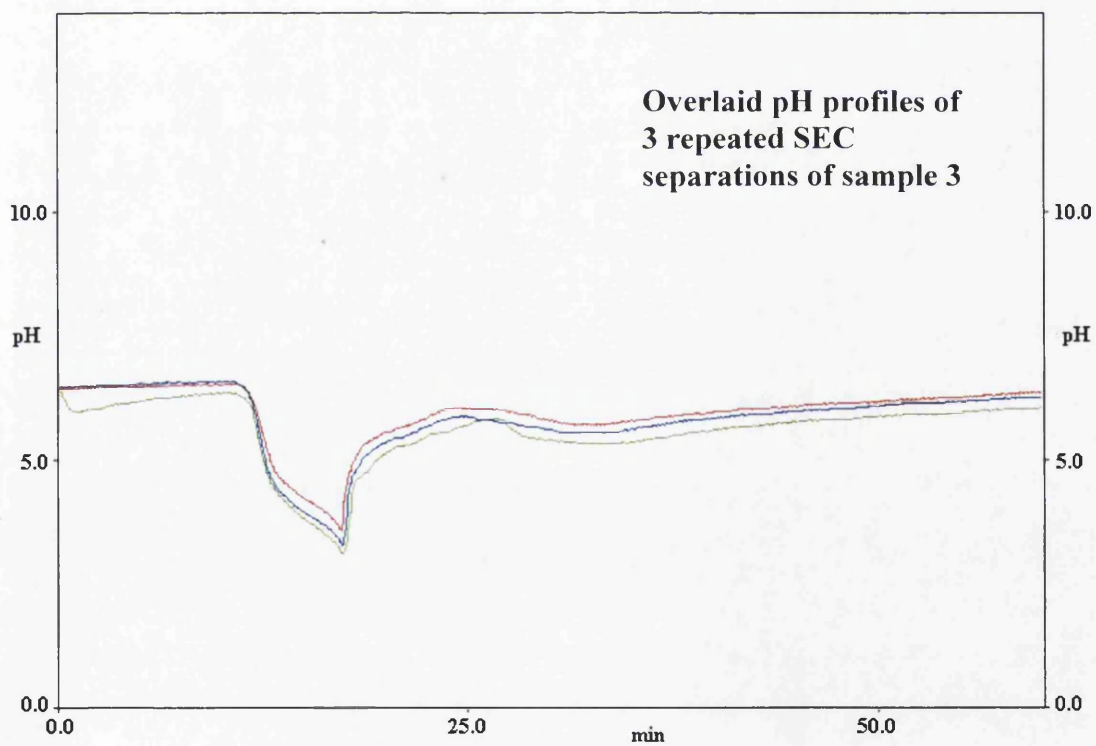
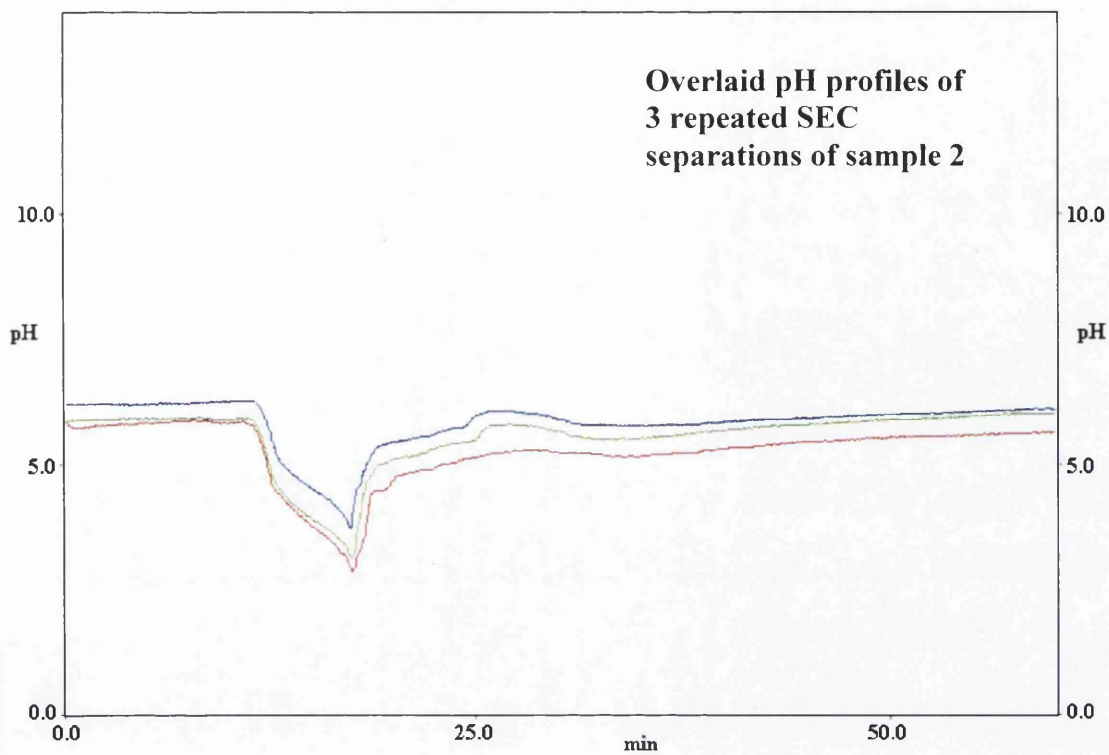


Figure 3.6: Overlaid pH profiles of three repeated SEC separations for sample 2 and 3 illustrating the drop in pH within the 16-19minute fraction.

Overall the pH results displayed good reproducibility as shown by the data present in table 3.7 and the profiles for both samples. The conductivity data also correlates with this change in pH, with a singular peak present in the 16-19 minute fraction. This again shows good reproducibility between separations and displays a variation in conductivity between the two samples. The results illustrate that sample 2 has a greater degree of conductivity at 16-19 minutes than the corresponding fraction of sample 3. A possible explanation for this is that there is an increased level of protons ( $H^+$ ) present in solution as represented by the decrease in pH.

Reproducibility of LC Data: Sample 2			Reproducibility of LC Data: Sample 3		
Peak	Conductivity		Peak	Conductivity	
	Inter Mean	Intra Mean		Inter Mean	Intra Mean
F1	0.00	0.00	F1	0.00	0.00
F2	0.00	0.00	F2	0.00	0.00
F3	0.09	0.02	F3	0.02	0.02
F4	0.08	0.18	F4	0.05	0.06
F5	0.00	0.00	F5	0.00	0.00
F6	0.00	0.00	F6	0.00	0.00
F7	0.00	0.00	F7	0.00	0.00
F8	0.00	0.00	F8	0.00	0.00
F9	0.00	0.00	F9	0.00	0.00
F10	0.00	0.00	F10	0.00	0.00
F11	0.00	0.00	F11	0.00	0.00
F12	0.00	0.00	F12	0.00	0.00
F13	0.00	0.00	F13	0.00	0.00

*Table 3.8: Inter- and intra-reproducibility data of conductivity measurements for each size exclusion fraction throughout the chromatographic run for both sample 2 (5 minutes post-connection) and sample 3 (5 minutes pre-disconnection).*

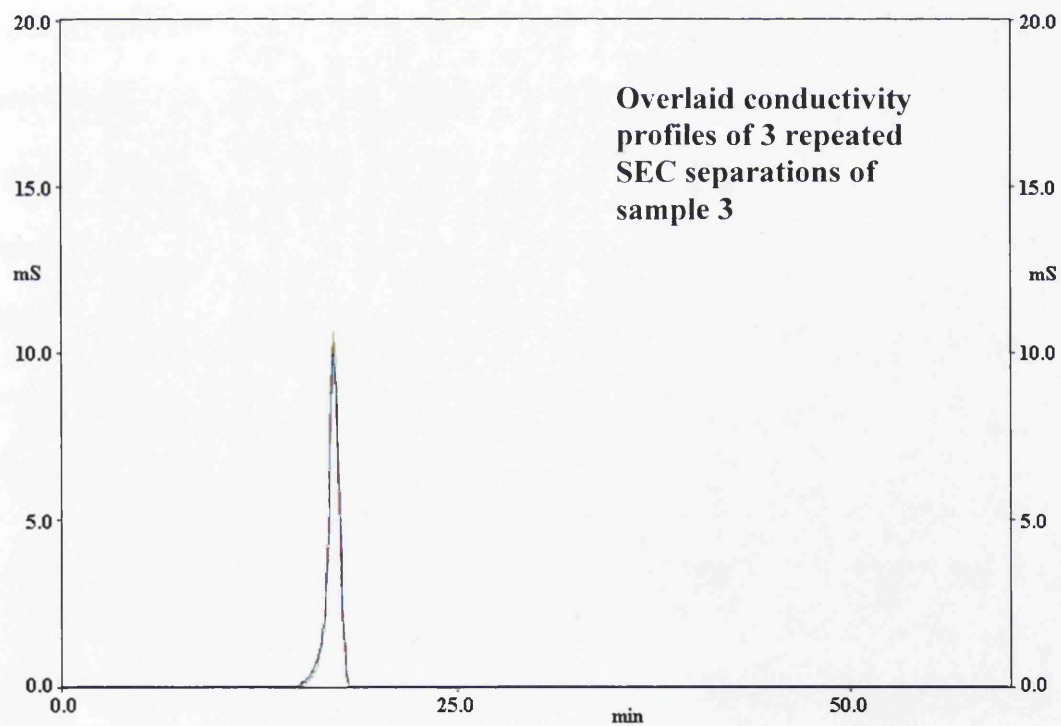
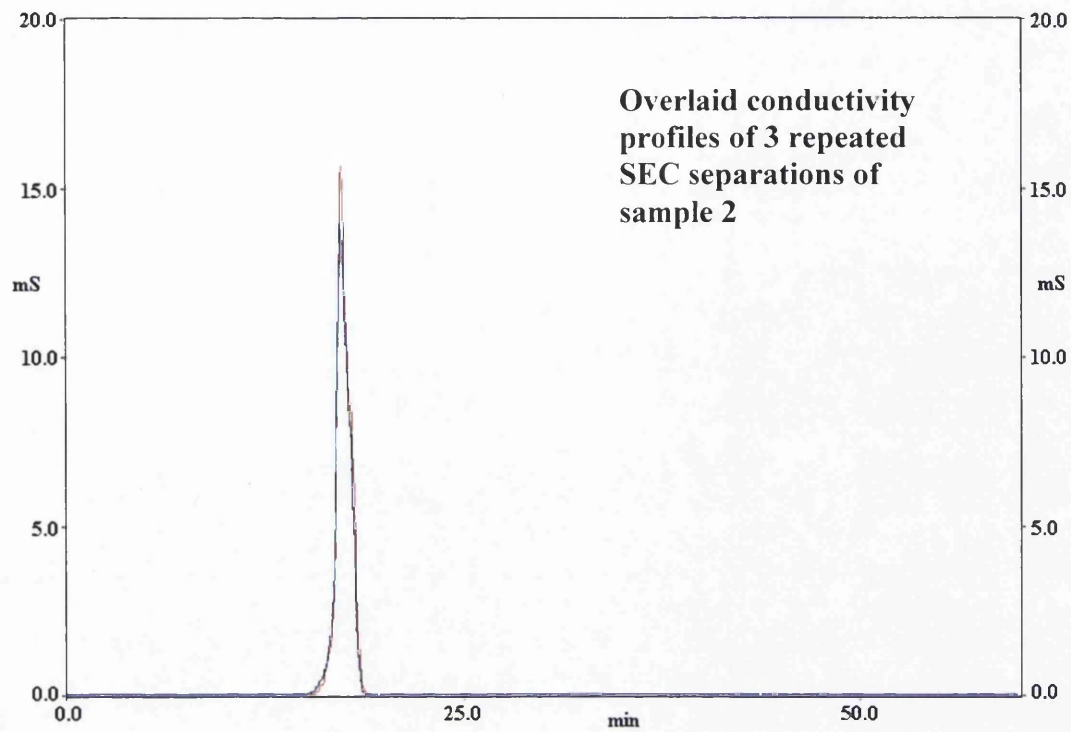


Figure 3.7: Overlaid conductivity profiles of three repeated SEC separations for samples 2 and 3 showing a peak within the 16-19 minute fraction at a maximum value of 15.7mS and 10.6mS respectively.

The size exclusion separation is essentially a fractionation procedure aiming to minimize the complexity of samples 2 and 3, making them more manageable. This degree of separation is still insufficient for the samples to be analysed by mass spectrometry, and hence an additional preparative step is required.

### 3.2.3. Reverse phase chromatography

The details of this separation is described in Chapter 2 section 2.2.2.2, and was setup as an online procedure, with the column effluent passing directly into the ionisation source of the mass spectrometer. Prior to injection, the size exclusion fractions were lyophilised to concentrate any analytes present. Each fraction was reconstituted in 150 $\mu$ L of deionised water, vortexed for 20 seconds and then transferred into a 200 $\mu$ L autosampler vial. Initial reverse phase separations of the SEC fractions appeared to have average reproducibility and only a few observed differences of the UV and mass spectrometric data of sample 2 and 3 were apparent. However, upon re-injection of a fraction after a few days at room temperature a large proportion of the UV activity of the chromatogram was absent (see figure 3.7). This prompted stability tests to be carried out in order to obtain a valid candidate ion or biomarker with confidence by mass spectrometry.

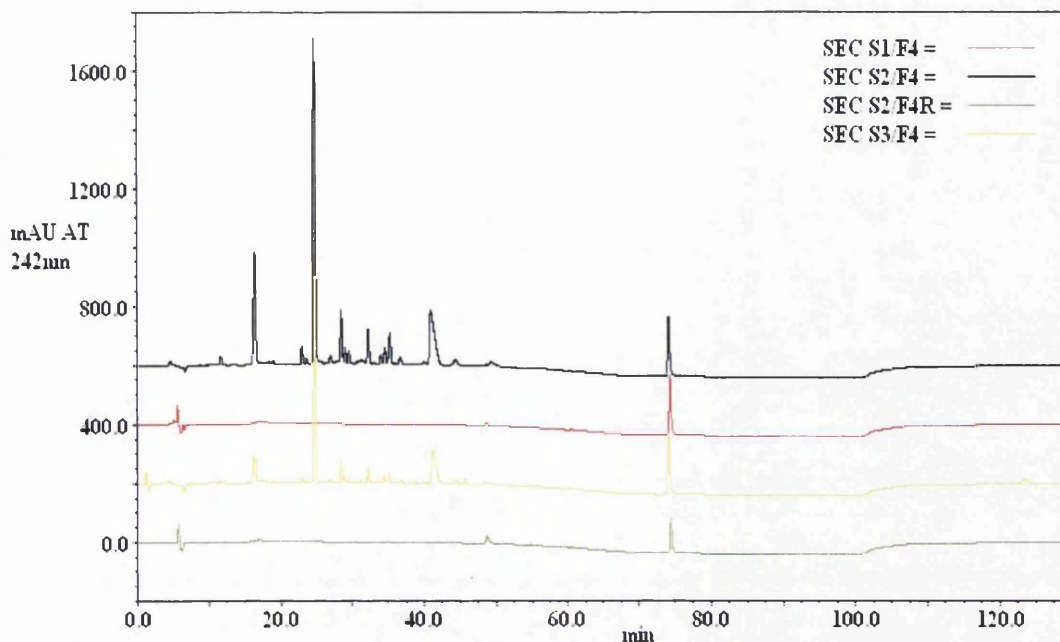


Figure 3.8: Illustration of overlaid UV chromatograms of fraction 4 of the size exclusion separation of samples 1 (blank dialysate solution prior to contact with the patient – S1/F4), 2 (S2/F4), and 3 (S3/F4). Fraction 4 of sample 2 was re-injected after 2 days (S2/F4R) and shows the absence of a cluster of peaks within the 10-50 minute region of the chromatogram of the original injection.

### 3.2.3.1. Stability investigations of each size exclusion fraction

In order for the online reverse phase method to be reproducible and valid for overnight analyses sample thermal lability was determined over a 16 hour period. This consisted of repeat injections of each size exclusion fraction at 0, 8, and 16 hour time points at 10°C into the LCQ DECA ion trap mass spectrometer (Thermo Fisher Scientific). Stability mass spectrometric profiles were obtained for each fraction and intensity values for observed peaks at the relevant time points recorded. From comparing the total ion chromatograms of the fractions there were observed differences in the results over the 16 hour period.

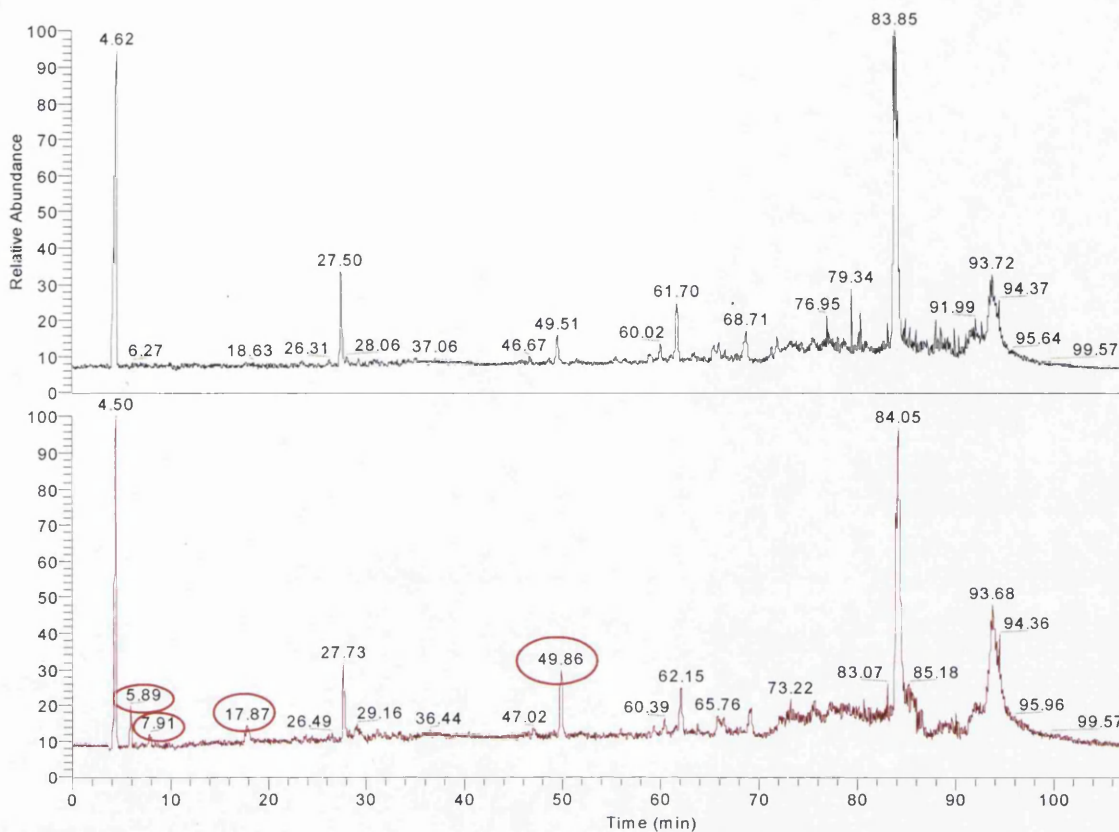


Figure 3.9: Total ion chromatograms of fraction 10 of the size exclusion separation of sample 2 for 0 hour (top) and 16 hour (bottom). Circled are some of the differences revealed by the stability investigations indicating the appearance of new analytes (at retention time of 49.86 minutes) or the accumulation of existing analytes (at retention times of 5.89, 7.91 and 17.87 minutes) over 16 hours.

A thorough analysis of the numerical data obtained for these investigations showed that for both sample 2 and sample 3 there was a large number of analytes present that had very poor stability during the 16 hour period. This indicates that some of the analytes present are thermally labile when existing as a mixture in the dialysate solution. Furthermore, these sample components may also be labile when exposed to UV light and thus any subsequent sample collection will involve appropriate UV protective containers.

In order to have a reliable final biomarker, those chosen from the comparative investigations of sample 2 and 3 should display considerable stability over the analysis time.

### 3.2.3.2. Reproducibility study of the reverse phase chromatography

Retention time is of considerable importance in confirming the presence of a sample component and reproducible chromatography is essential in obtaining valid analyte identification. In order for chromatographic runs to be compared the inter- and intra-reproducibility of retention time must be determined to identify any variations in analyte separation. Lyophilised size exclusion fractions were reconstituted in 150 $\mu$ L of deionised water prior to injection using the chromatographic conditions as described in Chapter 2, section 2.2.2.2. Retention time reproducibility was monitored using both UV and mass spectrometric detection as an on-line separation method. Each size exclusion fraction showed good intra-reproducibility with little variation observed between each run, deviating no more than  $\pm 0.124$  minutes from the mean retention time (see table 3.9). However, the subsequent injections to test inter-reproducibility did show some variations when compared to the data obtained the previous day but very little when compared to that on the same day. A likely explanation is the poor stability observed for some of the size exclusion fraction components. This was confirmed with separations carried out at a reduced temperature of 10 $^{\circ}$ C. At these optimum conditions each size exclusion fraction was separated with a high degree of reproducibility for alternate days determined by an F-test statistical calculation (see table 3.10). Thus, this again highlights the importance of choosing a final biomarker that has good degree of stability over a sufficient period of time, i.e. the duration of the analysis.

**Intra-reproducibility of the Relative Retention Times within Typical Reverse Phase Chromatographic (RP-HPLC) Runs**

<b>SEC Fraction</b>	<b>Standard Deviation of RRT</b>	<b>Variance of RRT</b>	<b>Probability Result is within 5% of Mean</b>
F1 (0-10 min)	0.061	0.0037	0.069
F2 (10-13 min)	0.055	0.0030	0.062
F3 (13-16 min)	0.110	0.0121	0.124
F4 (16-19 min)	0.080	0.0064	0.091
F5 (19-22 min)	0.020	0.0004	0.023
F6 (22-25 min)	0.041	0.0017	0.046
F7 (25-28 min)	0.150	0.0225	0.170
F8 (28-31 min)	0.130	0.0169	0.147
F9 (31-34 min)	0.097	0.0094	0.110
F10 (34-45 min)	0.078	0.0061	0.088
F11 (45-48 min)	0.046	0.0021	0.052
F12 (48-50 min)	0.080	0.0064	0.091
F13 (50-60 min)	0.067	0.0045	0.076

*Table 3.9: Intra-reproducibility data for the reverse-phase chromatographic runs at 10°C for the peak at m/z 803 which is present in all size exclusion fractions of sample 2. Intra-reproducibility is represented by a confidence interval value shown in the furthest right hand column. This statistical test has been carried out with P = 0.05, and hence we can be 95% confident that the values in the table for each run will not deviate more than ± 0.124 minutes from the mean standard deviation.*



<b>Inter-reproducibility of the Relative Retention Times within Typical Reverse Phase Chromatographic (RP-HPLC) Runs</b>			
<b>SEC Fraction</b>	<b>Standard Deviation of RRT Day 1</b>	<b>Standard Deviation of RRT Day 2</b>	<b>Inter-reproducibility by F-test</b>
<b>F1 (0-10 min)</b>	0.087	0.091	0.939
<b>F2 (10-13 min)</b>	0.100	0.105	
<b>F3 (13-16 min)</b>	0.157	0.162	
<b>F4 (16-19 min)</b>	0.146	0.140	
<b>F5 (19-22 min)</b>	0.080	0.070	
<b>F6 (22-25 min)</b>	0.094	0.091	
<b>F7 (25-28 min)</b>	0.194	0.193	
<b>F8 (28-31 min)</b>	0.182	0.182	
<b>F9 (31-34 min)</b>	0.140	0.134	
<b>F10 (34-45 min)</b>	0.127	0.121	
<b>F11 (45-48 min)</b>	0.099	0.095	
<b>F12 (48-50 min)</b>	0.133	0.125	
<b>F13 (50-60 min)</b>	0.099	0.098	

*Table 3.10: Reproducibility data of reverse phase chromatographic separation carried out over different days. This inter-reproducibility is represented by an F-test calculation.*

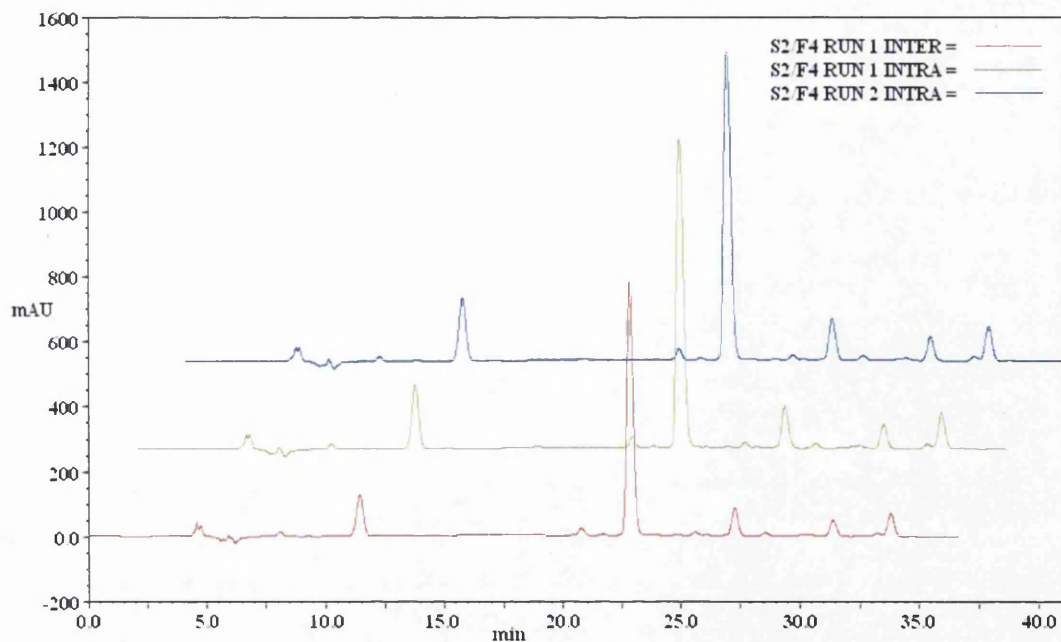


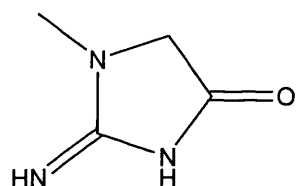
Figure 3.10: UV chromatogram of repeated reverse phase separations of fraction 4 from the size exclusion separation of sample 2. Good inter and intra reproducibility was observed although slight variations in the chromatograms may be attributed to the poor stability of some of the sample components.

### 3.3. Method Validation 2: Identification of Known Uremic Toxins

Extraction and detection of known uremic toxins can provide additional evidence that this methodology for dialysate preparation has the potential to identify new uremic analytes. Mass spectrometric detection was carried out using both the LCQ DECA ion trap and the LTQ Orbitrap (Thermo Fisher Scientific, CA, USA) and highlighted the presence of a range of ions suspected to be known uremic toxins. The assignments of identities were confirmed by considering chromatographic elution conditions, elemental formula and fragmentation data for each suspected ion. All ion accurate masses were obtained within an error of 5ppm unless stated otherwise in the relevant chapters.

### 3.3.1. Creatinine

Creatinine is one of the conventional markers used for the assessment of dialysis adequacy. The serum level can vary between patients as it is dependent on several factors other than accumulation, such as muscle mass and protein intake. Creatinine clearance is not affected to this degree and is used with other solutes including urea as a predictive parameter of dialysis adequacy. This uremic solute is usually measured using an enzymatic test as the conventional Jaffe reaction can be affected by glucose, ketoacids and particular classes of drugs, such as antibiotics<sup>[2]</sup>.



Molecular Formula –  $C_4H_7N_3O$

Protonated Molecule  $[M+H]^+$  Formula –  $C_4H_8N_3O$

*Figure 3.11: Chemical structure and related molecular information of the uremic toxin creatinine.*

In solution creatinine is present as an ion of  $m/z$  114. A peak of corresponding mass was observed in size exclusion fractions 6 and 7, consistent with the expected greater retention time associated with smaller molecular mass. Again, both the standard reference material and the unknown were fragmented by MS/MS and the mass spectra compared (see figure 3.12). The fragmentation of the standard shows a loss of 28Th associated with the transition  $m/z$  114  $\rightarrow$  86, due to the ejection of carbon monoxide from the protonated molecule (see figure 3.12). Creatinine is also capable of losing a highly stable structure of  $CH_2N_2$  consistent with the transition of  $m/z$  114  $\rightarrow$  72 and both losses are present not only in the MS/MS mass spectrum of the

standard but also in the unknown. It is therefore clear from the similarities of the ions present and their relative intensities that the unknown at  $m/z$  114 is creatinine.

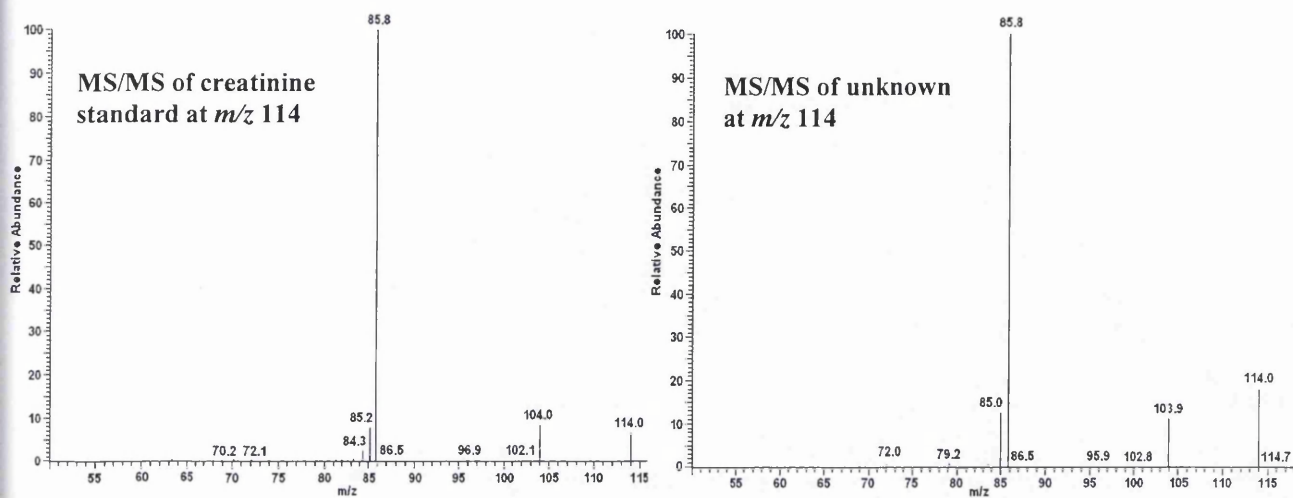
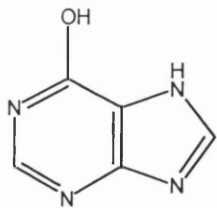


Figure 3.12: Fragmentation patterns of creatinine reference standard and an unknown of corresponding mass and retention time within the haemodialysate solution.

### 3.3.2. Hypoxanthine

This purine when retained in renal patients is known to affect several important physiological mechanisms such as vasoconstriction, platelet-induced vasorelaxation<sup>[3]</sup> and the repair of the endothelial barrier within the kidney<sup>[4]</sup>. This uremic toxin has only previously been detected in serum and blood samples with detection by UV spectrophotometry<sup>[5, 6]</sup> which is prone to interferences from other chromophores with similar reverse phase retention times. Hence, the ability to detect and monitor the levels of this uremic toxin within haemodialysate by mass spectrometry with greater specificity has the potential to limit such malfunctions of these physiological systems.



Molecular Formula - C<sub>5</sub>H<sub>4</sub>N<sub>4</sub>O  
 Molecular Mass - 136.0380

Protonated Molecule [M+H]<sup>+</sup> Formula - C<sub>5</sub>H<sub>5</sub>N<sub>4</sub>O  
 Protonated Molecule [M+H]<sup>+</sup> Mass - 137.0458  
 Protonated Molecule Measured Mass Error of 0.1ppm

Figure 3.13: Chemical structure of hypoxanthine with its expected molecular formula and mass for the neutral and positively charged moiety.

An unknown ion of corresponding mass-to-charge ( $m/z$  137) was observed within the haemodialysate solution at an appropriate mobile phase composition for hypoxanthine. A commercial standard of hypoxanthine was fragmented by collision induced dissociation (CID) and compared to the unknown fragmentation pattern (see figure 3.14). The mass spectrum of the latter showed all associated ion transitions of hypoxanthine from the initial loss of water ( $m/z$  137  $\rightarrow$  119) to those involving ring opening (see figure 3.14).

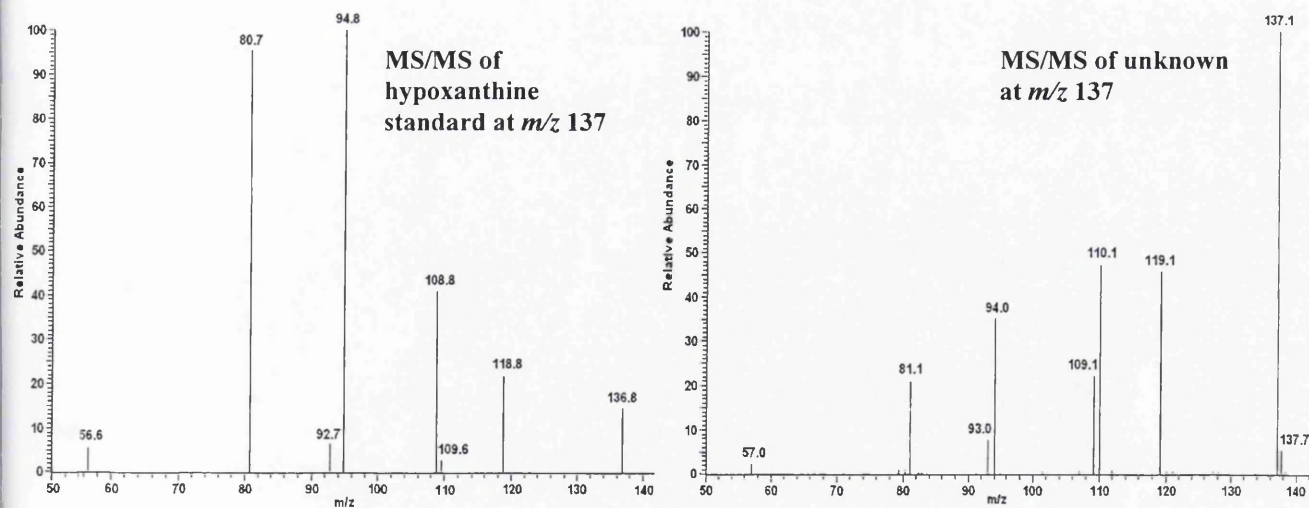


Figure 3.14: MS/MS fragmentation spectra of the unknown ion at  $m/z$  137 and the standard reference material of a known uremic toxin hypoxanthine. Assignment of identity is made on the similarities of the ions present in both mass spectra.

The accurate mass and fragmentation MS/MS studies of both the standard reference material and the unknown enabled a partial fragmentation reaction scheme to be created as shown figure 3.15. The elemental formula of the fragments and the neutral losses are consistent with hypoxanthine and with the remaining fragmentation data indicates that this is the identity of the unknown ion.

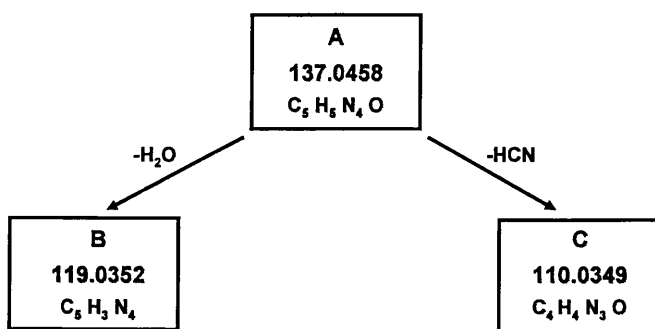
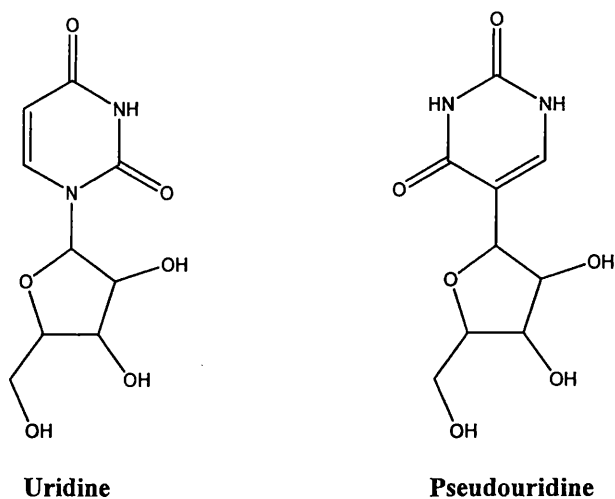


Figure 3.15: Fragmentation reaction scheme generated from the accurate masses of ions generated in the MS/MS mass spectrum of the unknown ion at  $m/z$  137.

### 3.3.3. Pseudouridine

This nucleoside is an isomer of uridine and both are common components of RNA nucleic acid. When charged they are both observed at  $m/z$  245 and differ only by the position of a nitrogen atom within the 6-membered ring changing the bond to the sugar moiety from a carbon-nitrogen bond to a carbon-carbon bond as shown in figure 3.16. This structural variation results in very different mass spectral fragmentation patterns and can enable the identification of the isomer without other analytical techniques such as NMR (see figure 3.18). Past studies on modified nucleosides have included their potential use as cancer markers<sup>[7, 8]</sup> and as analogues for developing antiviral treatments<sup>[9, 10]</sup>. Pseudouridine that originates from tRNA is normally

excreted in urine as an intact molecule and has been reported to accumulate in the serum of renal patients<sup>[11, 12, 13, 14, 15, 16]</sup>.



*Figure 3.16: Chemical structures of the nucleosides uridine and pseudouridine. These are isomers and contain a structural difference involving an amino group within the 6-membered ring resulting in change from a carbon-carbon bond to a carbon-nitrogen bond to the sugar moiety.*

An unknown parent ion was observed at  $m/z$  245 within the dialysate solution. The reverse phase chromatographic retention time, corresponding to approximately 15% organic solvent, indicated that this ion could be either of these nucleosides from the list of published uremic toxins. This was supported further by the calculated accurate mass which generated a correct elemental formula for the ions of these isomers of  $C_9H_{13}N_2O_6$  at an error of 1.3ppm. An accurate mass was also obtained for each of the fragment ions present in the unknown MS/MS mass spectrum and a product ion reaction scheme generated (see figure 3.16). This shows several possible

losses of water consistent with both isomers and is more likely with pseudouridine as shown in the MS/MS mass spectrum.

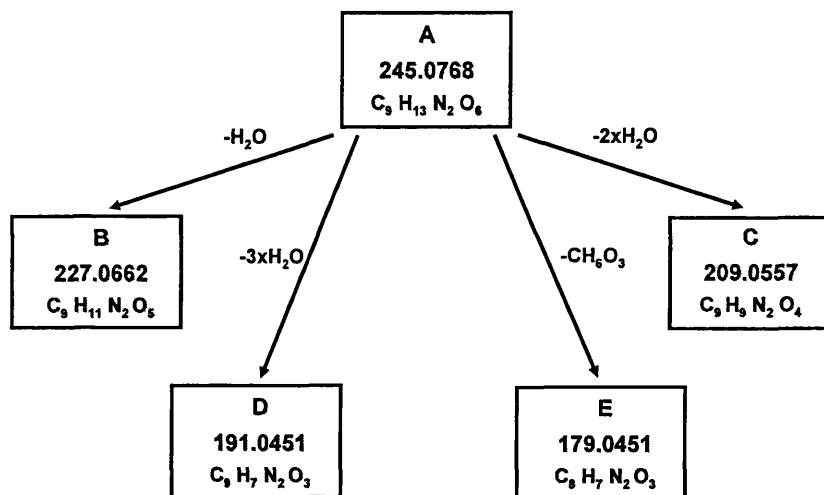


Figure 3.17: Product ion reaction scheme generated from the calculated accurate mass of ions present in the MS/MS mass spectrum of the unknown ion at m/z 245.

A comparison of the standard<sup>[17]</sup> MS/MS fragmentation pattern with the accurate mass fragmentation reaction scheme and MS/MS mass spectrum of the unknown reveals that it this ion is protonated pseudouridine.



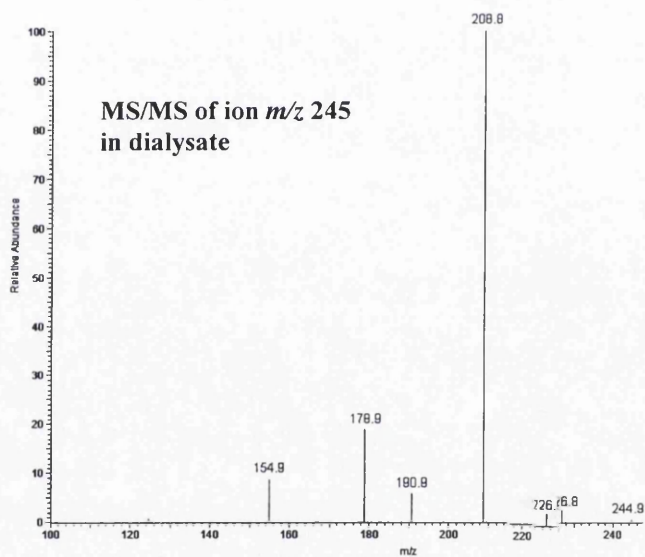
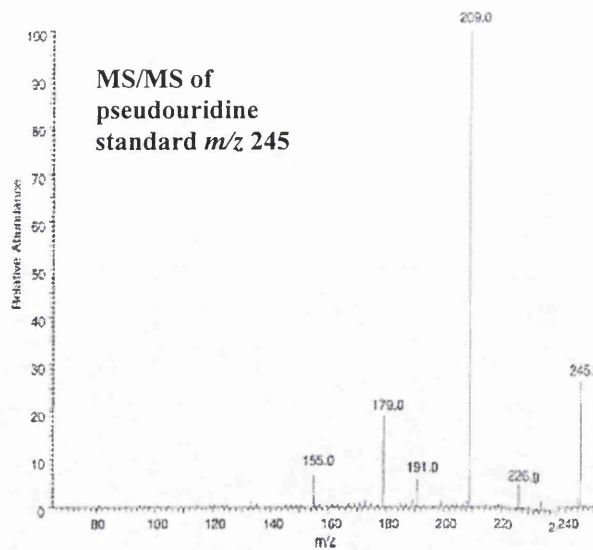
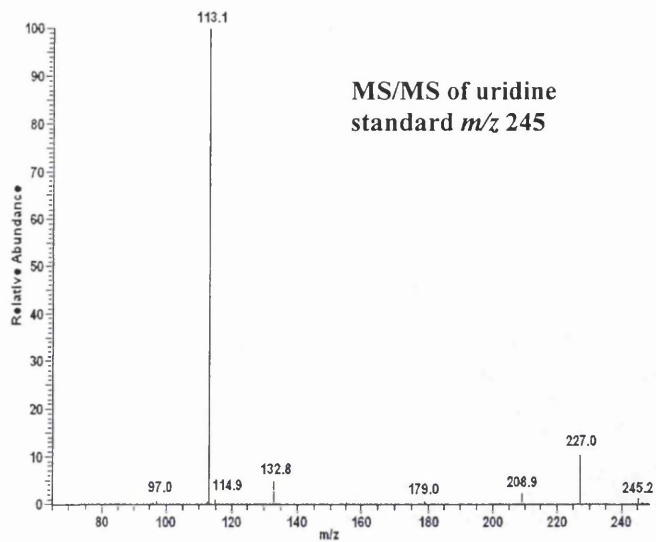


Figure 3.18: MS/MS fragmentation pattern of uridine and pseudouridine<sup>[17]</sup> standard reference materials and an unknown parent ion of corresponding mass-to-charge present within the haemodialysate sample. All MS/MS mass spectra were obtained on an LCQ DECA ion trap mass spectrometer.

### 3.3.4. Advanced glycation end products

These are formed as an eventual product of the Maillard reaction; a series of complex non-enzymatic reactions between a reducing sugar such as glucose and the amino groups of biomolecules in the form of proteins, peptides and amino acids<sup>[18]</sup>. These AGE compounds are known to exist at elevated levels in uremic patients and can be found as much as 10 times the levels associated with a healthy person<sup>[19, 20]</sup>. This is thought to be due to a combination of both an increase in AGE synthesis by oxidative or carbonyl stress and a decrease in AGE excretion by the kidney after digestion to form AGE-peptides<sup>[19]</sup>. Accumulation of modified peptides can result in further production of AGE-related toxins and biochemical interactions with molecules such as lipoproteins and collagen<sup>[21]</sup>. In addition to AGE-peptides, smaller moieties resulting from reactions with amino acids can be formed, such as N<sup>ε</sup>-(carboxymethyl)lysine (CML), N<sup>ε</sup>-(carboxyethyl)lysine (CEL) and pentosidine.

An ion of corresponding mass-to-charge to a common AGE, N<sup>ε</sup>-(carboxyethyl)lysine or CEL, was discovered within the haemodialysate. The calculated accurate mass suggested that this ion had an elemental formula (C<sub>9</sub>H<sub>19</sub>N<sub>2</sub>O<sub>4</sub>), obtained with an error of 2.4ppm and a relevant fragmentation pattern was generated for additional information (see figure 3.19). This included a partial product ion reaction scheme of ions with their calculated accurate mass and associated elemental formula as shown in figure 3.20. All data matched published work indicating that this unknown at *m/z* 219 is the protonated molecule, [M+H]<sup>+</sup>, of the AGE N<sup>ε</sup>-(carboxyethyl)lysine<sup>[22, 23, 24]</sup>.

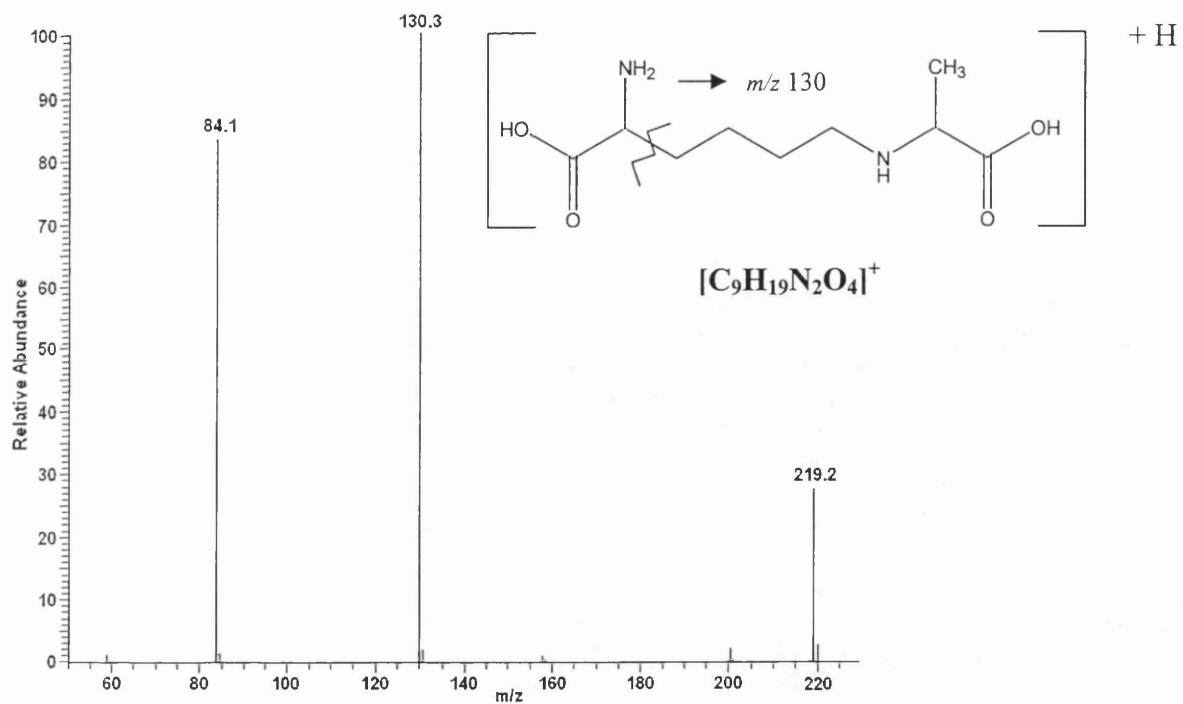


Figure 3.19: MS/MS fragmentation pattern of an unknown at  $m/z$  219 within the haemodialysate solution.

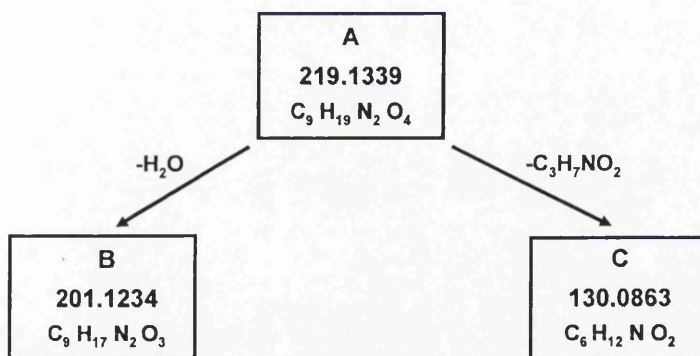
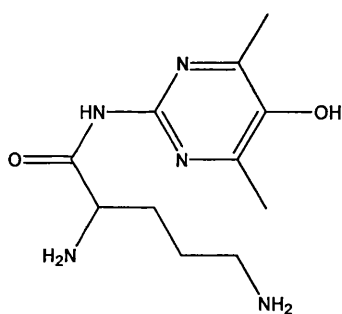


Figure 3.20: Fragmentation reaction scheme of the unknown ion at  $m/z$  219 present in the haemodialysate and suspected to be the uremic toxin CEL.

In addition to this AGE moiety, an ion present at  $m/z$  254 generated a calculated accurate mass and subsequent elemental formula at an error of -1.7ppm that corresponded to the molecule Argpyrimidine (2,5-Diamino-pentanoic acid (5-hydroxy-4,6-dimethyl-pyrimidin-2-yl)-amide). This was selected for further structural elucidation by fragmentation and generated the MS/MS mass spectrum shown in figure 3.22. Unfortunately all past research regarding detection and quantitation of this AGE compound by mass spectrometry has involved synthesis in-house with multiple reaction monitoring (MRM) analyses<sup>[22]</sup>. Thus, due to time constraints with this project structural elucidation and assignments for this ion were made from the proposed fragment matches from predicted degradative pathways and published information available.



Molecular Formula -  $C_{11}H_{19}N_5O_2$

Molecular Mass - 253.1533

Protonated Molecule  $[M+H]^+$  Formula -  $C_{11}H_{20}N_5O_2$

Protonated Molecule  $[M+H]^+$  Mass - 254.1612

Protonated Molecule Measured Mass Error of 1.7ppm

*Figure 3.21: Chemical structure of the AGE compound Argpyrimidine or 2,5-Diamino-pentanoic acid (5-hydroxy-4,6-dimethyl-pyrimidin-2-yl)-amide, and the molecular mass and protonated molecule mass information.*

The fragmentation pattern of the ion suspected to be Argpyrimidine contained the expected  $m/z$  140 (see figure 3.22) as stated by MRM experiments in published literature<sup>[22]</sup>, however there are several other ions also present.

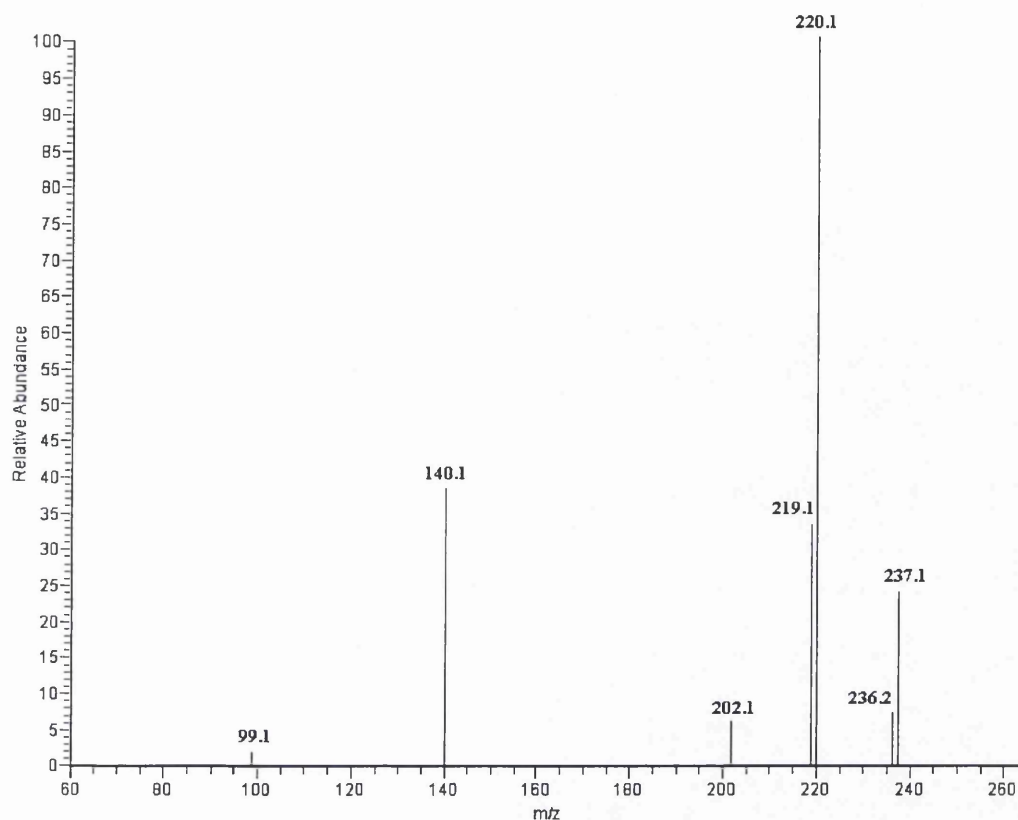


Figure 3.22: MS/MS fragmentation mass spectrum of the unknown ion at  $m/z$  254 within the haemodialysate suspected to be the AGE compound Argpyrimidine.

From the chemical structure and MS/MS mass spectrum obtained at high accuracy we can predict that the ions at  $m/z$  237 and 220 are as a result of successive losses of the primary amino groups in the form of ammonia. The transition of  $m/z$  254 to  $m/z$  236 of 18Th is commonly associated with hydroxyl groups where they are lost as a water molecule, and the next fragment at  $m/z$  219 may form due to the ejection of a primary amino group again as ammonia. The MS/MS mass spectrum also shows another decrease of 18Th from  $m/z$  220 to  $m/z$  202, and could be due to a loss of water following the subtraction of two ammonia molecules. This information enables a product ion reaction scheme to be generated as shown in figure 3.23.

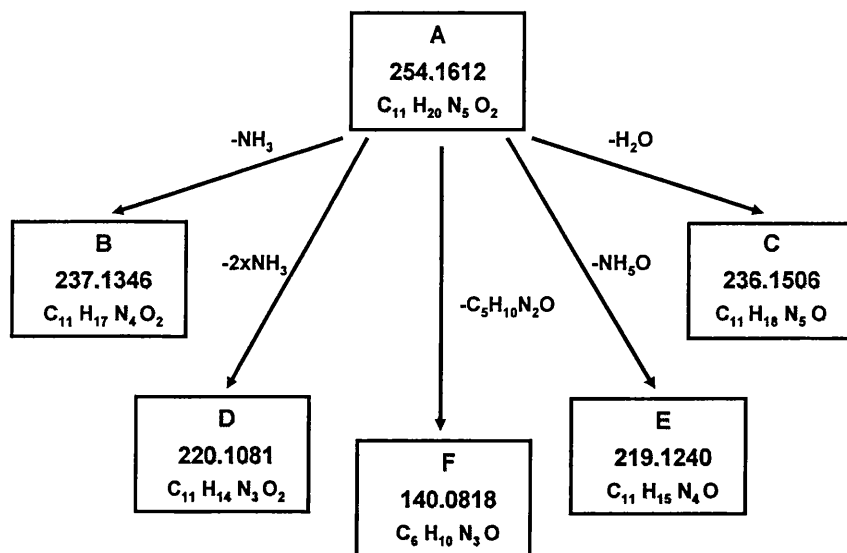


Figure 3.23: Product ion fragmentation reaction scheme of the unknown ion at  $m/z$  254 suspected to be the uremic toxin and AGE compound Argpyrimidine.

A comparison of elemental and fragmentation information with the suggested chemical structure and published literature indicate that this unknown ion at  $m/z$  254 is Argpyrimidine. However, in order to confirm this suspicion the results shown here would require comparison against a standard reference material.

### 3.3.5. Dimethylglycine

This is a by-product of the betaine-homocysteine methyltransferase (BHMT)-mediated biochemical pathway involved in synthesis and reuse of the amino acid methionine<sup>[25, 26]</sup>. This physiological mechanism is essential for the correct formation of phospholipids and eventual cell membrane construction. The starting material is a homologue of the amino acid cysteine, homocysteine. This uremic toxin is frequently used as a risk factor for atherosclerosis<sup>[25]</sup>, a clinical condition common in dialysis patients. Past research indicates that both dimethylglycine and homocysteine

accumulate in renal patients. During chronic renal failure (CRF) dimethylglycine is used as an independent predictor of homocysteine levels and an indirect marker for the development of atherosclerosis<sup>[26]</sup>.

An ion at  $m/z$  104 was observed within the haemodialysate at low organic solvent composition (~10% methanol) consistent with the elution of dimethylglycine<sup>[27]</sup>. The accurate mass of this ion obtained at an error of 3.7ppm suggests it has an elemental formula of  $C_4H_{10}NO_2$  and is equivalent to that of ionized dimethylglycine ( $[M+H]^+$ ).

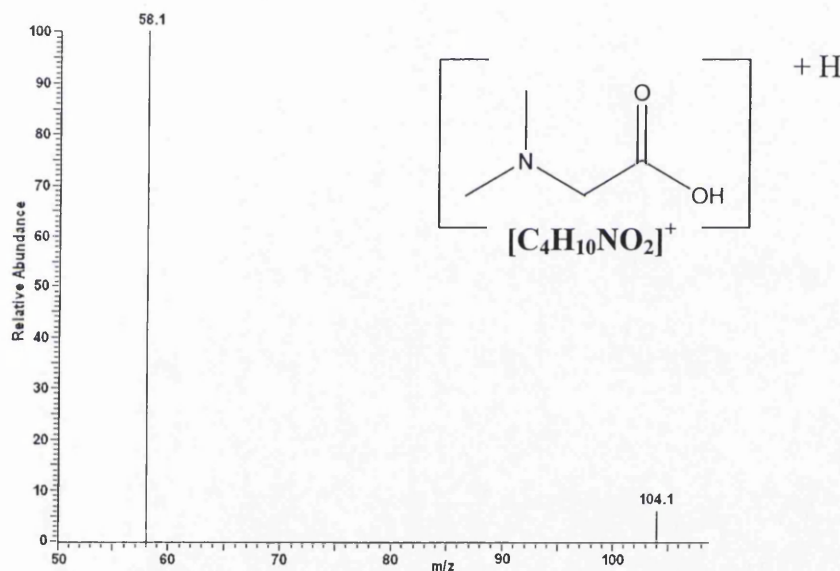


Figure 3.24: Fragmentation mass spectrum of unknown at  $m/z$  104 present in the haemodialysate sample. Both elemental formula and MS/MS data correspond to dimethylglycine standard reference material and related work<sup>[27]</sup>.

When fragmented, the unknown formed an almost identical pattern to that of the dimethylglycine standard<sup>[27]</sup> in respect of the ions present and their relative abundances (see figure 3.24). Characteristically this unknown ion had just one product ion at  $m/z$  58, corresponding to the loss of carbon monoxide and water or

formic acid from the protonated molecule of dimethylglycine and is formed due to its high stability.

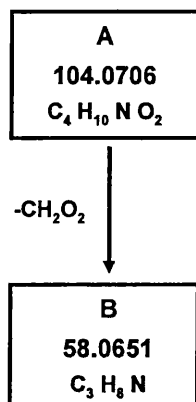


Figure 3.25: Product ion reaction scheme of an unknown ion at  $m/z$  104 detected within the haemodialysate solution and suspected to be the uremic solute dimethylglycine.

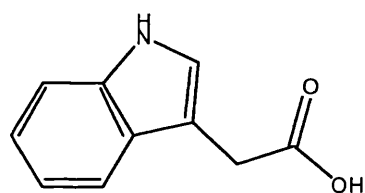
This ion within the unknown MS/MS mass spectrum has an accurate mass with an elemental formula indicative of this dimethylglycine fragment (see figure 3.25) and when included with the other information described previously confirms the presence of this uremic toxin in haemodialysate.

### 3.3.6. Indole-3-acetic acid

This uremic toxin is classed as ‘protein-bound’ exerting a larger effective molecular weight than expected and results in a limited haemodialytic removal. Its propensity to bind to proteins can affect the toxicity of administered drugs as it can compete for protein receptors or binding sites. Studies have also shown that this indole disrupts ion transport within the kidney tubules and interferes with the reabsorption of necessary solutes<sup>[28]</sup>. This work was carried out mainly with UV detection following



reverse phase HPLC separation and states that this solute will elute from the column with 40% organic mobile phase<sup>[6]</sup>. The reverse phase separation of prepared haemodialysate shows a peak present at  $m/z$  176 corresponding to elution conditions of 40-45% organic mobile phase and was suspected of being indole-3-acetic acid. Further evidence for this was obtained by calculating the accurate mass of the unknown ion at an error of -1.4ppm and generated an exact elemental formula.



Molecular Formula -  $C_{10}H_9NO_2$

Molecular Mass - 175.0628

Protonated Molecule  $[M+H]^+$  Formula -  $C_{10}H_{10}NO_2$

Protonated Molecule  $[M+H]^+$  Mass - 176.0706

Protonated Molecule Measured Mass Error of -1.4ppm

*Figure 3.26: Chemical structure of unbound indole-3-acetic acid and its molecular formula and mass for both the neutral and positively charged species.*

Fragmentation of indole-3-acetic acid generates a distinctive MS/MS mass spectrum. One product ion is formed, clearly observed at  $m/z$  130, and is as a result of the loss of the carboxylic acid group. The unknown ion at  $m/z$  176 fragments to an almost identical pattern to indole-3-acetic acid with regards to the ions present and their relative intensities.

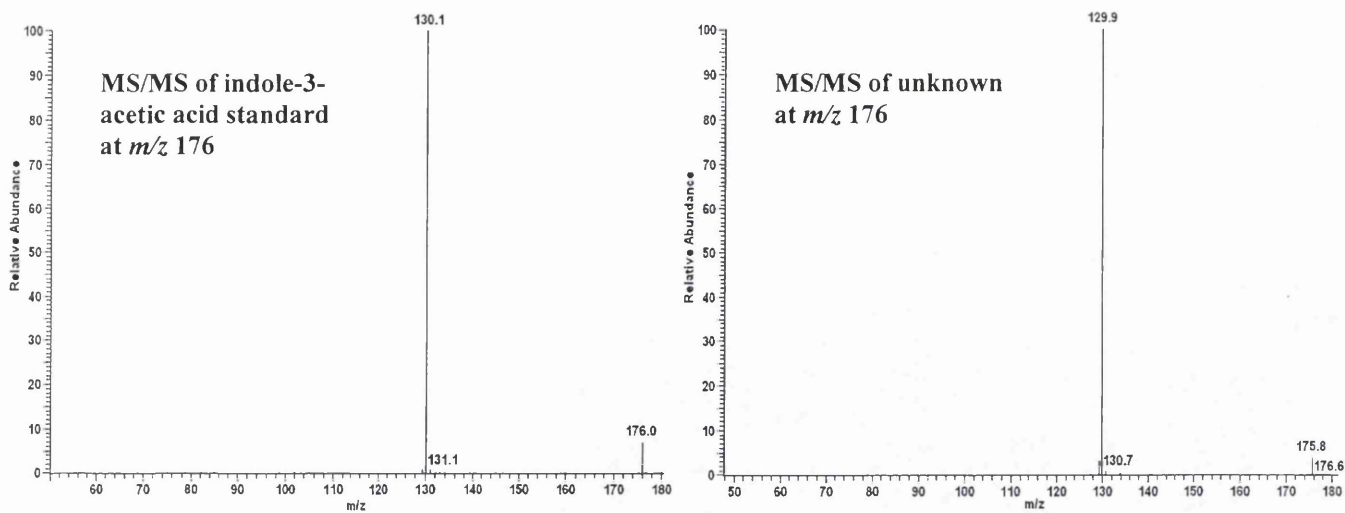


Figure 3.27: Fragmentation patterns of indole-3-acetic acid at  $m/z$  176 and an ion within the haemodialysate solution of corresponding mass.

This suggestion is supported further by the calculated accurate mass of the fragment ion,  $m/z$  130, producing an elemental formula indicative of the loss associated with the protonated indole-3-acetic acid (see figure 3.28).

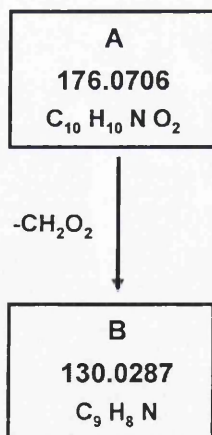
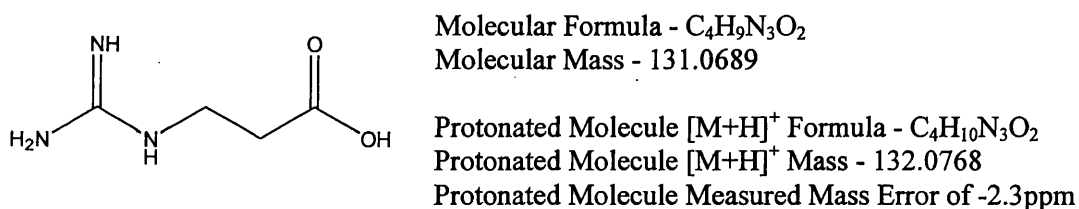


Figure 3.28: Fragmentation pattern of the unknown ion at  $m/z$  176 suspected to be the uremic solute indole-3-acetic acid.

3.3.7.  $\beta$ -guanidinopropionic acid

This is an analogue of another uremic toxin, creatine, and has an inhibitory affect on creatine uptake into the cell. The reduction of intracellular creatine disrupts the enzyme, creatine kinase, and can lead to an eventual interference in bone formation and repair<sup>[29]</sup>. It is also thought to have an additional role in the ‘metabolic syndrome’ and research has shown that it can cause an elevated level of protein modification<sup>[30]</sup>. Previous work comparing the removal of this analyte using different dialysis techniques has shown that despite having similar excretion kinetics to the current biomarkers, urea and creatinine, levels remain higher following treatment than in healthy patients.



*Figure 3.29: Chemical structure of the uremic toxin  $\beta$ -guanidinopropionic acid and associated molecular information.*

In solution this analyte is observed at  $m/z$  132 and forms ions of  $m/z$  115, 114, 90, 86, 72 and 60 when fragmented using collision induced dissociation as shown in figure 3.31. The first two fragments correspond to the loss of ammonia and water directly from the protonated molecule and the peak at  $m/z$  86 is generated following the loss of carbon monoxide from the ion at  $m/z$  114.

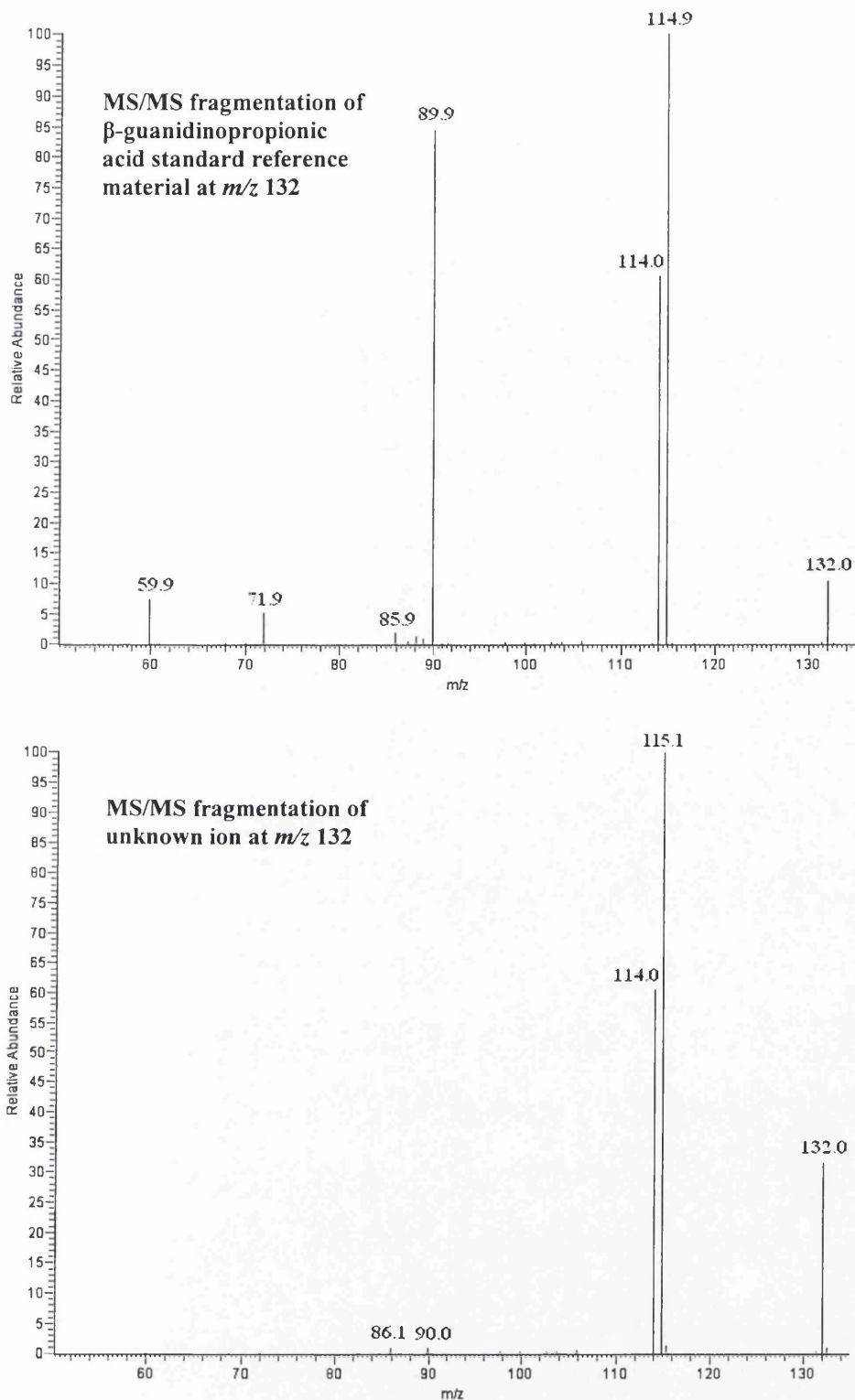


Figure 3.30: Fragmentation MS/MS spectra of the uremic toxin  $\beta$ -guanidinopropionic acid with a parent ion at  $m/z$  132 and an unknown ion present in the haemodialysate solution of corresponding mass-to-charge. The standard of  $\beta$ -guanidinopropionic acid shows characteristic fragments at  $m/z$  115, 114, and 90 which are also present in the unknown mass spectrum.

It is these three ions that are present in the fragmentation MS/MS mass spectrum of an unknown at  $m/z$  132 in the haemodialysate solution and has an accurate mass suggesting an elemental formula of  $C_4H_{10}N_3O_2$  at an error of -2.3ppm. The accurate mass of the product ions in the MS/MS mass spectrum are also indicative of those associated with  $\beta$ -guanidinopropionic acid (see figure 3.31). Therefore, this information in combination with the similarities of the fragmentation patterns suggest that the unknown ion at  $m/z$  132 is the uremic toxin  $\beta$ -guanidinopropionic acid.

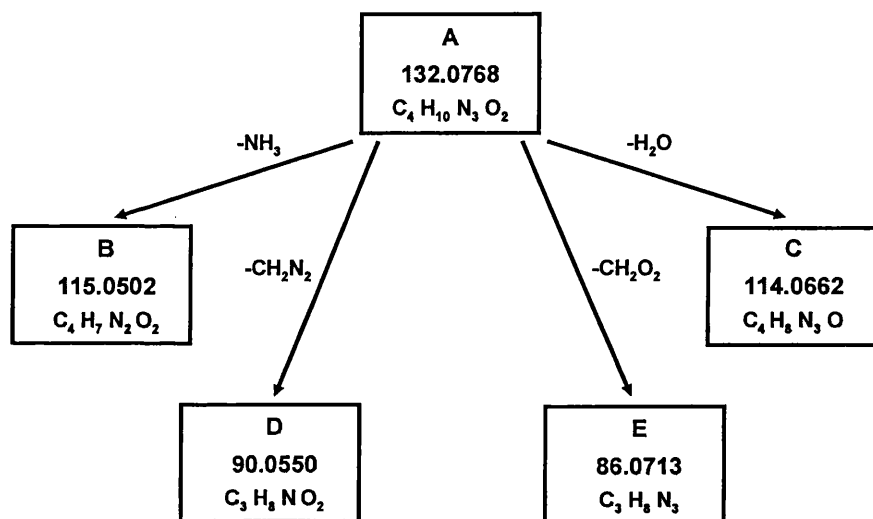


Figure 3.31: Product ion reaction scheme of the unknown ion at  $m/z$  132 proposed to be the uremic toxin  $\beta$ -guanidinopropionic acid.

### 3.3.8. N- $\alpha$ -acetylarginine

N- $\alpha$ -acetylarginine is one of several guanidino compounds known to accumulate within renal patients to toxic levels<sup>[31, 32, 33, 34]</sup>. They are all essentially metabolites of proteins and peptides and contain the guanidino functional group. The vast majority of published research regarding the physiological role of these compounds has involved two of the most established guanidines, creatine and arginine, and work

carried out with N- $\alpha$ -acetylarginine has primarily involved toxicity studies. This was to assess the extent it contributes to symptoms associated with renal failure<sup>[35, 36]</sup>.

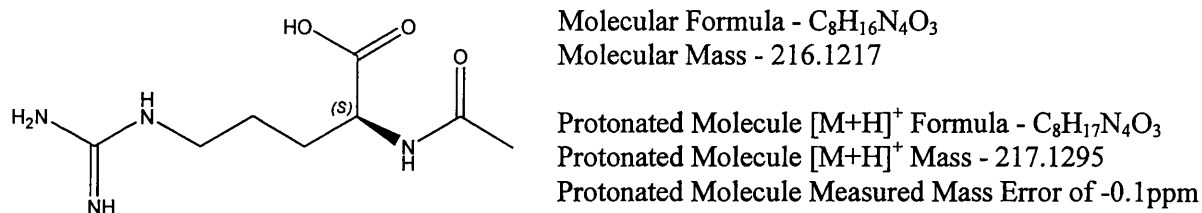


Figure 3.32: Chemical structure of the uremic toxin N- $\alpha$ -acetylarginine and its associated mass information.

From the chemical structure it is likely to encounter a loss of either ammonia or water due to the presence of a primary amino and hydroxyl group, respectively during CID fragmentation. This is illustrated by the MS/MS mass spectrum of the commercial N- $\alpha$ -acetylarginine standard by the first two fragment peaks at  $m/z$  200 and 199, associated with the subtraction of 17 and 18Th from the parent ion. The remaining product ions are formed as a result of hydrogen rearrangements priming the structure for further fragmentation involving the heteroatoms present. Analysis of the haemodialysate sample revealed an ion of  $m/z$  217 that had an accurate mass and a proposed elemental formula of C<sub>8</sub>H<sub>17</sub>N<sub>4</sub>O<sub>3</sub> to within -0.1ppm. This, and the reverse phase retention time corresponded to the known uremic toxin of N- $\alpha$ -acetylarginine, which was fragmented for a comparison with the unknown (see figure 3.33). Visually it appears as though the fragmentation patterns were obtained from the same compound due to the similarities of the masses of ions present. This is also supported by the accurate masses obtained for the ions at  $m/z$  200, 199, 158 and 157 which correspond to the losses ammonia, water, a guanidine-type fragment of CH<sub>5</sub>N<sub>3</sub>, and

the terminal carboxylic acid group from N- $\alpha$ -acetylarginine, respectively (see figure 3.34).

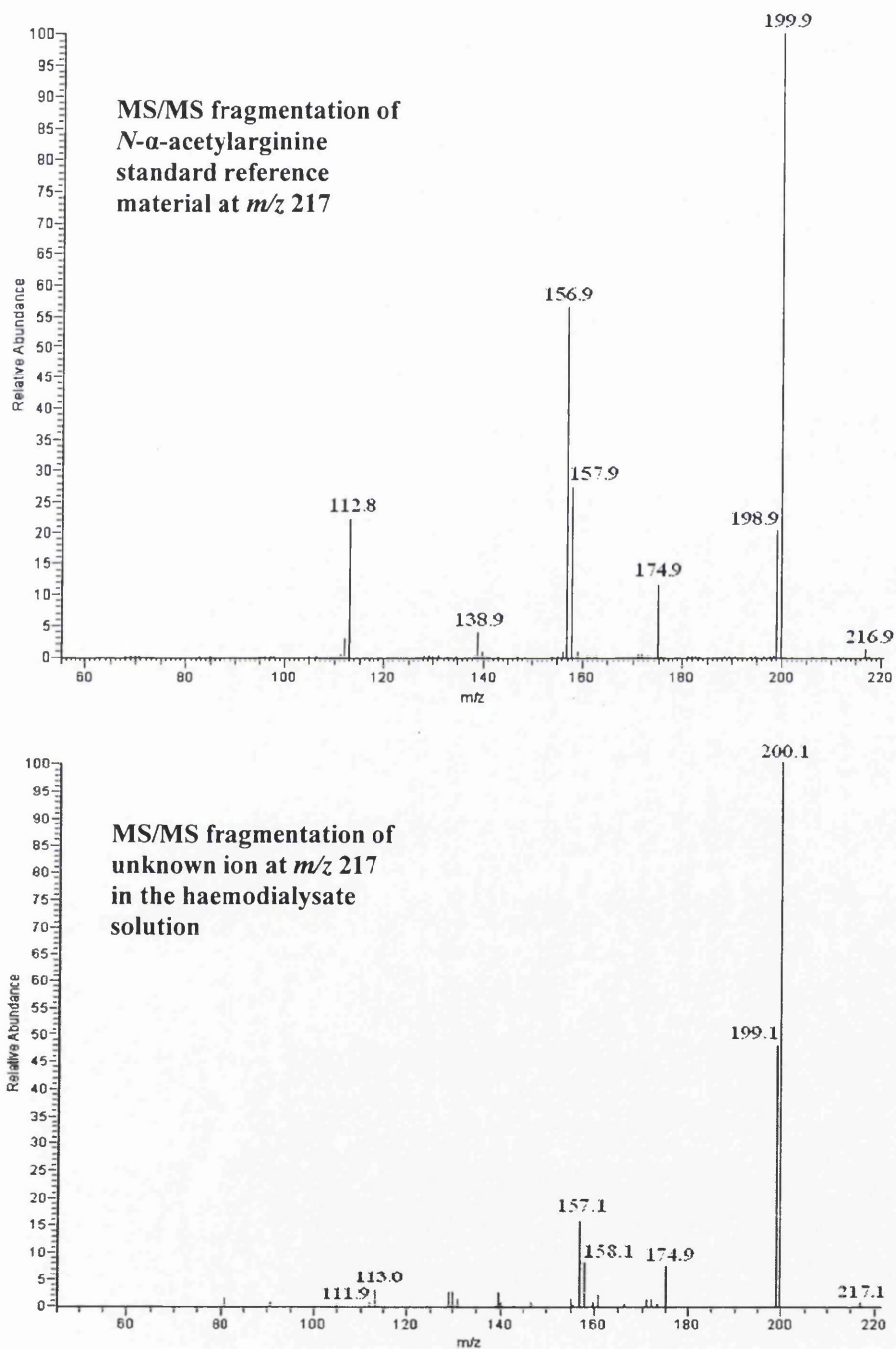


Figure 3.33: MS/MS fragmentation mass spectra of the commercially available standard N- $\alpha$ -acetylarginine with a parent ion of  $m/z$  217 and an ion of corresponding mass-to-charge within the haemodialysate concentrate.

Hence, with the other mass spectrometric and chromatographic information obtained, it appears that N- $\alpha$ -acetylarginine is present in the haemodialysate solution at  $m/z$  217.

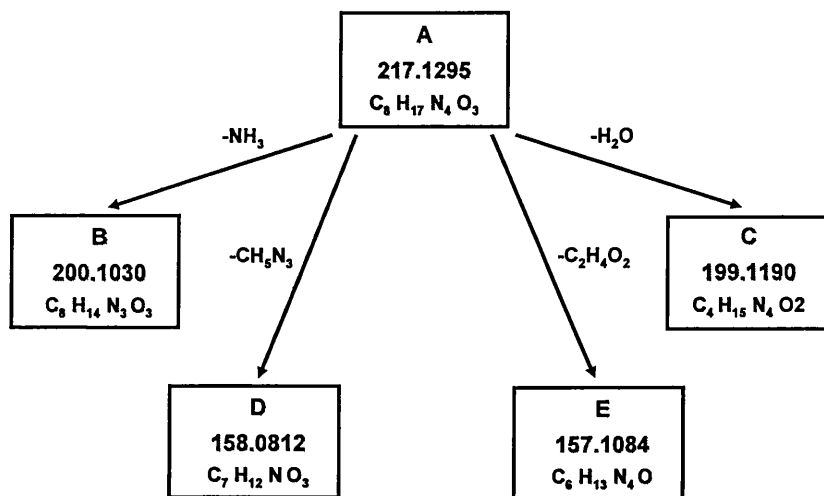


Figure 3.34: Product ion reaction scheme of the unknown ion at  $m/z$  217.

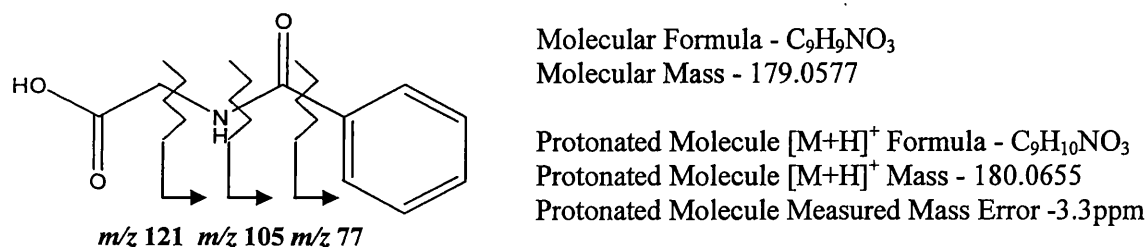
### 3.3.9. Hippuric acid (benzoylglycine)

This is an aromatic acid formed as a result of glycine metabolism involving the conjugation of this amino acid with a benzene ring. Research has indicated that this has the potential to be used as a biomarker for a dysfunction in the glucuronidation stage of metabolism. It highlighted that levels of hippuric acid can more than double from metabolism malfunction as a result of over stimulation of the nuclear receptor peroxisome proliferator-activated receptor alpha (PPAR $\alpha$ ). Therefore the potential of this analyte as a biomarker has been investigated as part of a range of candidates in monitoring disturbances of intermediary metabolism<sup>[37]</sup>.

This analyte has also been found to accumulate within renal patients receiving dialysis treatment and is hence classed as a uremic toxin. Previous studies involving



this patient cohort have involved monitoring this toxin within blood, plasma and urine but not haemodialysate. This analyte has proved detectable by mass spectrometry and has been quantified in urine samples by monitoring a particular reaction path of its fragmentation by tandem mass spectrometry or MS/MS.



*Figure 3.35: Chemical structure of hippuric acid showing the main fragmentation sites of the uremic toxin.*

It is known that hippuric acid will fragment by collision induced dissociation into three specific regions, illustrated by figure 3.35 and generate ions of  $m/z$  121, 105, and 77. An ion of comparative mass and fragmentation was discovered at an appropriate retention time for hippuric acid within the haemodialysate solution (see figure 3.36).

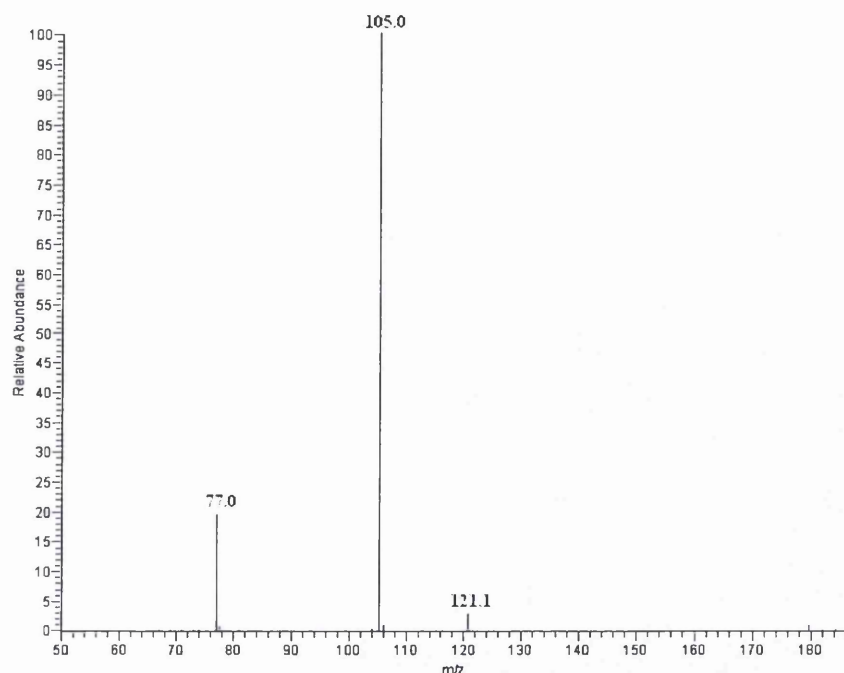


Figure 3.36: Fragmentation mass spectrum of the unknown ion at  $m/z$  180.0649 obtained using the LTQ Orbitrap mass spectrometer.

This had a calculated accurate mass at an error of -3.3ppm and subsequent elemental formula that supported this suspicion. Additional evidence was provided by the accurate mass of some of the fragment ions associated with the unknown (see figure 3.37), providing comparative elemental formulae for the hippuric acid product ions at  $m/z$  105 and 77. This information indicates that hippuric acid can be detected and identified within haemodialysate by this mass spectrometry methodology.

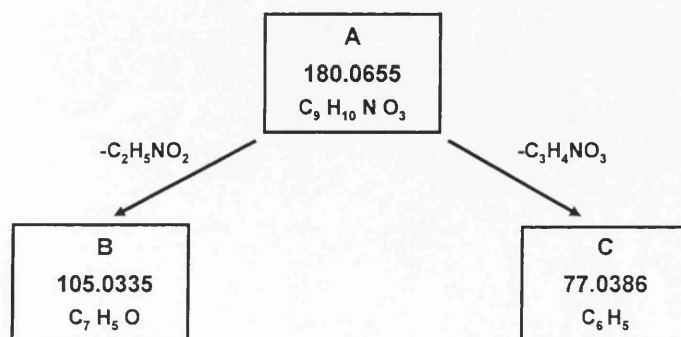
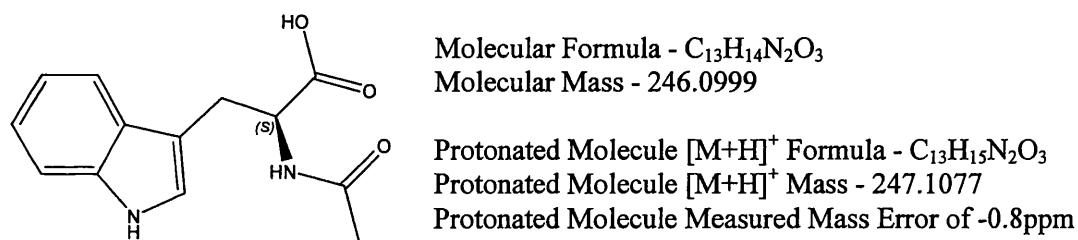


Figure 3.37: Product ion accurate mass reaction scheme of the unknown at  $m/z$  180.

3.3.10. N-acetyltryptophan

This is formed as a consequence of tryptophan metabolism and is one of a number of amino acid metabolites that are used to monitor disease progression with patients suffering from chronic renal failure. Past research with this patient population has shown that it exists as both a free molecule and as a protein-bound moiety but will accumulate primarily in the free form during renal impairment<sup>[38]</sup>. In addition to this ailment, abnormal levels and malfunctions of the biochemical pathways of tryptophan and its analogues have been implicated in several forms of mental illness<sup>[39]</sup>.

When ionised by protonation this analyte is observed at  $m/z$  247 and fragments in a specific pattern forming ions at  $m/z$  229, 211, 205, 201, 188, 187, 159, and 130 (see figure 3.39). The first two transitions involve successive losses of water, initially as the hydroxyl group of the carboxylic acid and then through a rearrangement of the opposing ketone group. The remaining ions will be generated by fragmenting the aliphatic section of the molecule with losses involving the carboxylic acid group as a whole, the ketone section and the amine group.



*Figure 3.38: Chemical structure of the uremic toxin N-acetyltryptophan and its molecular information.*

An ion was detected within the haemodialysate with a calculated accurate mass of 247.1075 obtained at an error of -0.8ppm, indicative of the elemental formula

$C_{13}H_{15}N_2O_3$ . This with comparative reverse phase elution conditions, suggest that this unknown ion could be *N*-acetyltryptophan. Fragmentation studies generated a similar mass spectrum regarding higher mass ions, although some were of lower intensities than the standard and could possibly be due to an inherent effect within the haemodialysate solution (see figure 3.39).

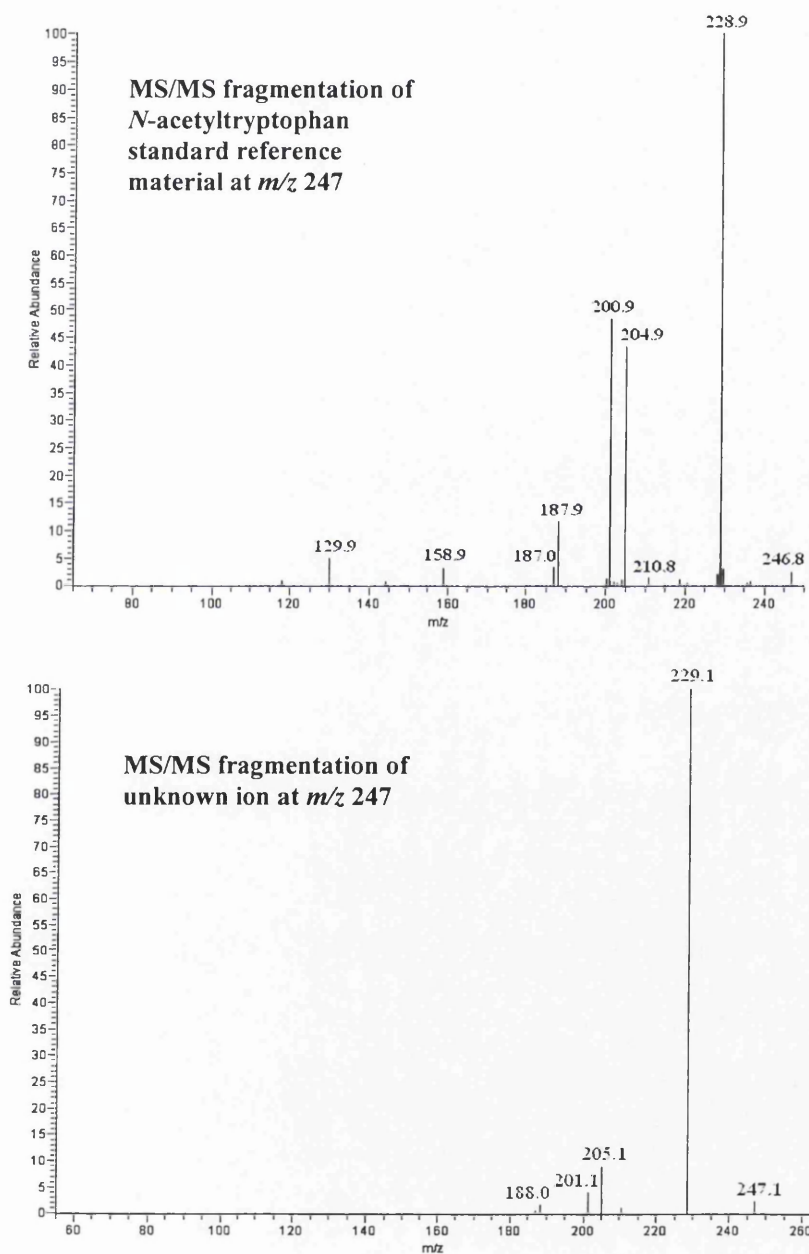


Figure 3.39: MS/MS fragmentation mass spectra of the standard reference material of *N*-acetyltryptophan with parent ion at m/z 247 and an ion of corresponding mass-to-charge within the haemodialysate solution.

A calculated accurate mass was obtained for some of these fragments with comparative elemental formula for the losses described for protonated *N*-acetyltryptophan (see figure 3.41). These when included with the other information of the unknown ion it indicates that it is the uremic toxin *N*-acetyltryptophan.

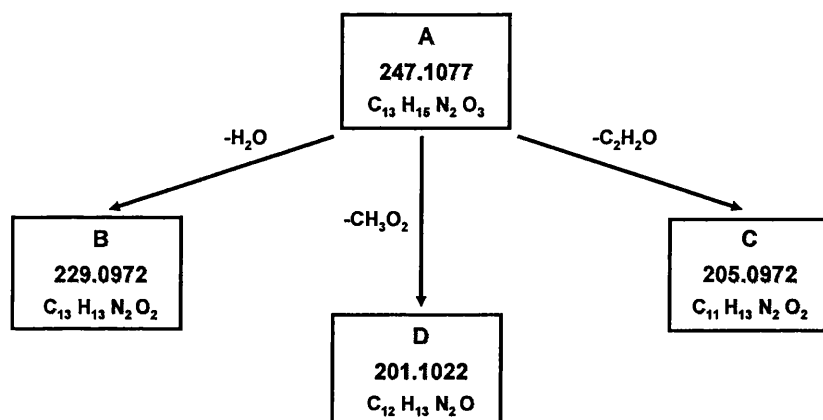
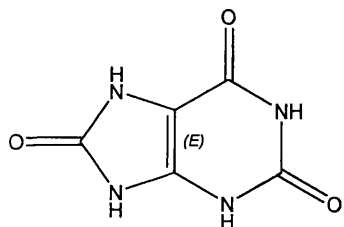


Figure 3.40: Fragmentation reaction scheme of the unknown ion at  $m/z$  247.

### 3.3.11. Uric acid

Uric acid is formed as a consequence of metabolising adenine- and guanine-based purines. At non-toxic levels, this analyte is a potent antioxidant important for maintaining intracellular biochemical interactions<sup>[40]</sup>. However, if allowed to accumulate to toxic levels within the blood it primarily leads to the development of a form of arthritis commonly known as gout. Other affects are often observed in renal or uremic patients and past research has highlighted a link of uric acid accumulation to the more severe illness of stroke in non-insulin dependent diabetes mellitus patients<sup>[41]</sup>.



Molecular Formula -  $C_5H_4N_4O_3$

Molecular Mass - 168.0278

Protonated Molecule  $[M+H]^+$  Formula -  $C_5H_5N_4O_3$

Protonated Molecule  $[M+H]^+$  Mass - 169.0356

Protonated Molecule Measured Mass Error of -2.5ppm

Figure 3.41: Chemical structure of the uremic analyte uric acid.

The fragmentation pattern of this uremic toxin involves losses of carbon monoxide and ammonia which generate the fragments at  $m/z$  141 and  $m/z$  152 respectively (see figure 3.43). These transitions cause the ring systems to open and subsequent fragmentations to occur that will involve cleavages associated with the heteroatoms. An ion of the same mass-to-charge generated a calculated accurate mass that indicated, at an error of -2.5ppm, that the ion has the elemental formula of  $C_5H_5N_4O_3$ . This is consistent with uric acid and the fragmentation pattern of the standard reference material was compared to that of the unknown (see figure 3.42). The mass spectra are highly similar regarding both masses present and the ratio of product ions formed.

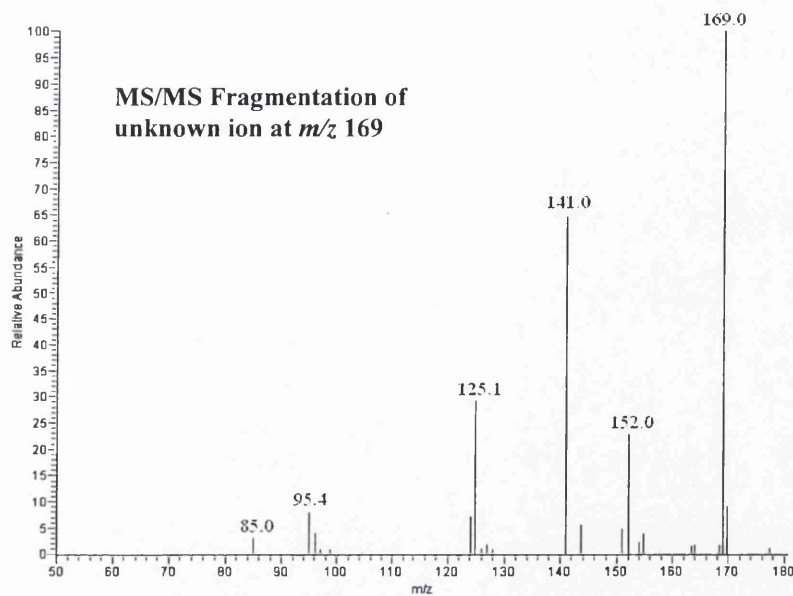
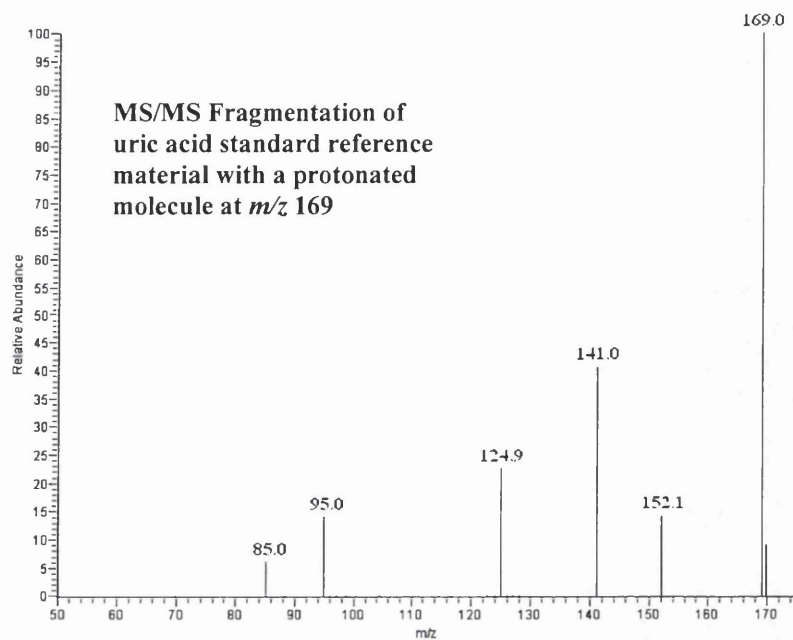


Figure 3.42: MS/MS spectra of uric acid standard reference material with the protonated molecule of  $m/z$  169 and an ion of corresponding mass-to-charge within the haemodialysate.

Accurate mass information was obtained for some of the product ions observed in the unknown MS/MS mass spectrum and their associated elemental formulae are shown in figure 3.43. These product ions have a corresponding elemental composition to those fragments generated from uric acid of the same mass-to-charge. This, with a

low reverse phase retention time of 4.1 minutes and the elemental composition information imply that the unknown present at  $m/z$  169 is the uremic toxin uric acid.

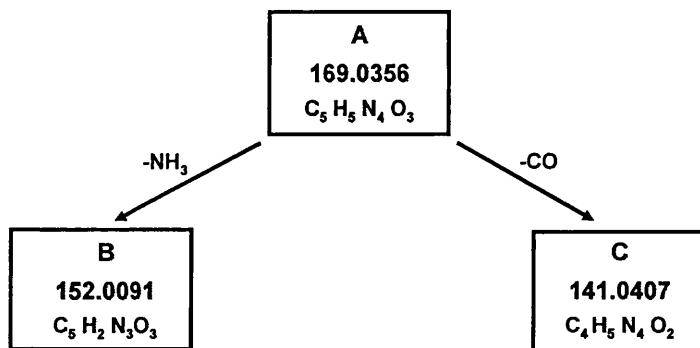
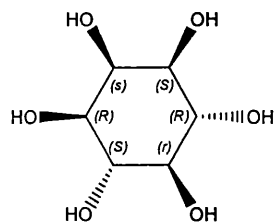


Figure 3.43: Product ion reaction scheme of the unknown ion at  $m/z$  169 suspected to be uric acid.

### 3.3.12. Myo-inositol

This biomolecule is of great importance in maintaining signals within neurones and sodium-potassium ATPase activity through protein kinase C stimulation. This has led to a large proportion of research investigating the role of *myo*-inositol in neurological malfunctions within the brain such as bipolar disorder<sup>[42]</sup>. This biomolecule is also important with diabetic patients, particularly those receiving renal treatment. High levels of glucose, associated with this cohort, can compete with *myo*-inositol to overstimulate the protein kinase C pathway and with long reaching affects. Of particular importance for renal patients is the rapid increase in prostaglandin synthesis as the accumulation of these biomolecules has previously been linked to initial problems of glomerular filtration in early renal patients<sup>[43, 44, 45, 46]</sup>.



Molecular Formula - C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>

Molecular Mass - 180.0628

Protonated Molecule [M+H]<sup>+</sup> Formula - C<sub>6</sub>H<sub>13</sub>O<sub>6</sub>Protonated Molecule [M+H]<sup>+</sup> Mass - 181.0707

Protonated Molecule Measured Mass Error of 0.4ppm

Figure 3.44: Chemical structure of the uremic toxin *myo*-inositol of mass 180Da, and is observed as a protonated ionic species at  $m/z$  181.

The protonated species of *myo*-inositol is observed at  $m/z$  181 and fragments by using collision induced dissociation into several ions as shown in figure 3.46. From the structure it is clear that it is likely to lose water, shown by a difference of 18Th between peaks in the mass spectrum, such as the transitions of  $m/z$  181 → 163 and  $m/z$  153 → 135. An ion observed within the haemodialysate solution at  $m/z$  181 and eluting at appropriate reverse phase solvent composition was suspected of being *myo*-inositol. Additional evidence supporting this is the calculated accurate mass which was obtained within an error of 0.4ppm and indicated that this ion had the elemental formula of C<sub>6</sub>H<sub>13</sub>O<sub>6</sub>. This ion was fragmented and compared to the commercial standard of the uremic toxin. Figure 3.45 illustrates the similarities of the mass-to-charge of ions generated by the fragmentation of both the unknown and standard.

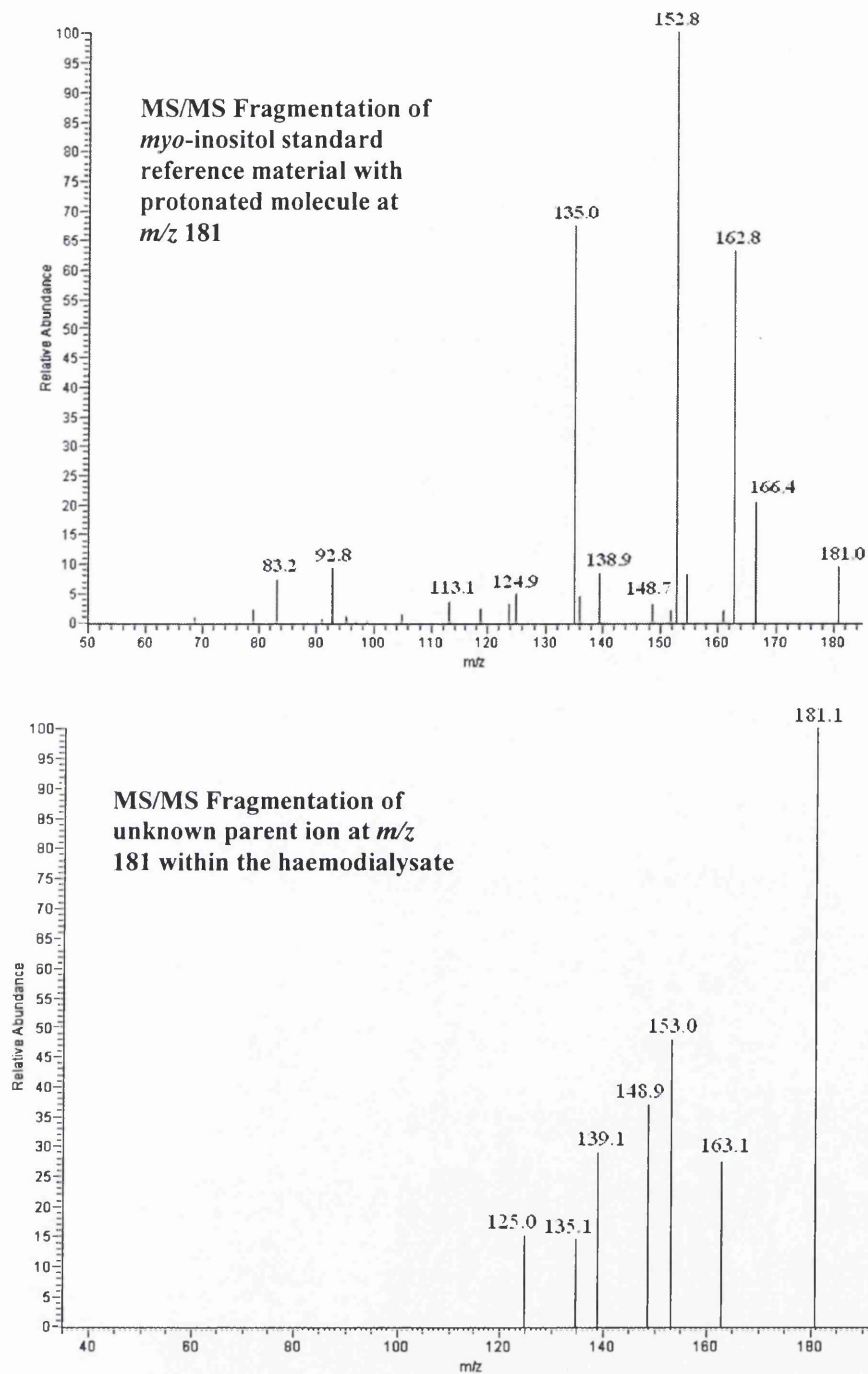


Figure 3.45: MS/MS mass spectra of myo-inositol standard reference material with parent ion at  $m/z$  181 and an unknown ion of the same mass-to-charge present in the haemodialysate solution.

The calculated accurate mass of some of the ions generated in the unknown MS/MS mass spectrum have an elemental formula suitable for the fragments of same mass-to-

charge present in the MS/MS mass spectrum of *myo*-inositol (see figure 3.46). All the mass spectrometric and chromatographic data characterising the unknown ion at  $m/z$  181 indicate that it is the uremic toxin *myo*-inositol.

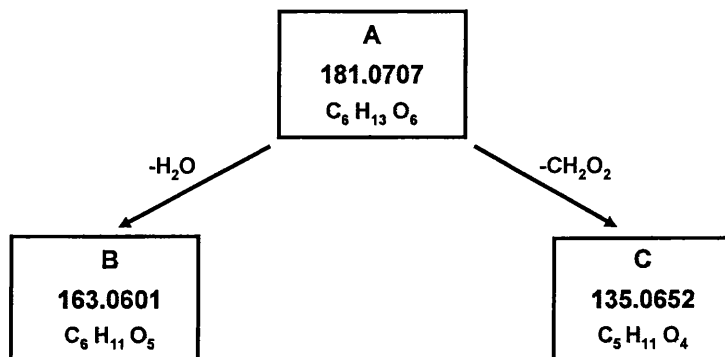


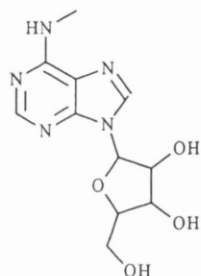
Figure 3.46: Fragment ion reaction scheme of the unknown ion present in the haemodialysate solution of  $m/z$  181.

### 3.3.13. N<sup>6</sup>-methyladenosine

This is one of three isomers of methyladenosine that have similar fragmentation patterns and therefore require separation by chromatography for isolation<sup>[47]</sup>. The levels of these modified nucleosides have been monitored in urine with the potential to predict cancer progression. They are released into the blood from tRNA and are excreted in urine by a healthy kidney prior to becoming toxic to the body due to a lack of degradative or conversion biochemical pathways. In patients with renal insufficiency, this affect is more profound as excretion is limited and levels can accumulate to those associated with uremic toxicity.

An ion of the same mass-to-charge, calculated with an error of 1.9ppm, was discovered within the haemodialysate at a retention time of 43 minutes corresponding to approximately 70% organic elution solvent. The fragmentation data obtained for this ion is indicative of all three isomers showing the protonated molecule at  $m/z$  282

with characteristic fragment ions as a result of losing the sugar moiety and the nucleobase at  $m/z$  150 and 133 respectively. Also present are fragment ions at  $m/z$  106 and 92, and could correspond to a successive loss of neutral ammonia and a secondary amine of formula  $\text{CH}_2\text{NCH}$  from the purine base<sup>[48]</sup>.



Molecular Formula -  $\text{C}_{11}\text{H}_{15}\text{N}_5\text{O}_4$

Molecular Mass - 281.1119

Protonated Molecule  $[\text{M}+\text{H}]^+$  Formula -  $\text{C}_{11}\text{H}_{16}\text{N}_5\text{O}_4$

Protonated Molecule  $[\text{M}+\text{H}]^+$  Mass - 282.1197

Protonated Molecule Measured Mass Error of 1.9ppm

Figure 3.47: Chemical structures of three methyladenosine isomers that are candidates for the ion at  $m/z$  282 within the haemodialysate. Each has a similar fragmentation pattern and require separation according to polarity.

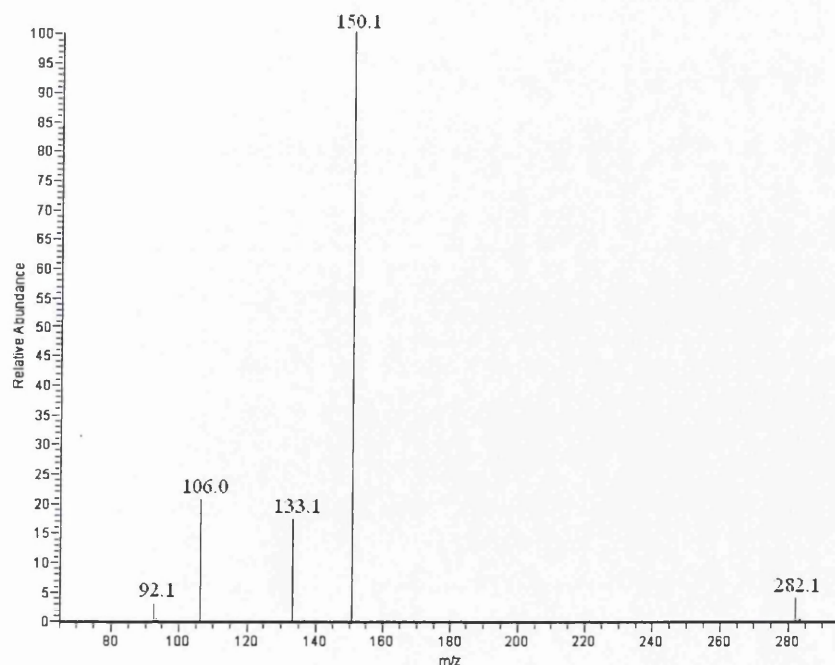


Figure 3.48: Fragmentation MS/MS spectrum of a protonated molecule present in the haemodialysate solution that has a calculated accurate mass corresponding to the

elemental formula of methyladenosine at 282.1213 and a matching fragmentation pattern.

The calculated accurate mass of some MS/MS product ions of the unknown show appropriate elemental formulae for those fragment ions of N<sup>6</sup>-methyladenosine described above (see figure 3.49).

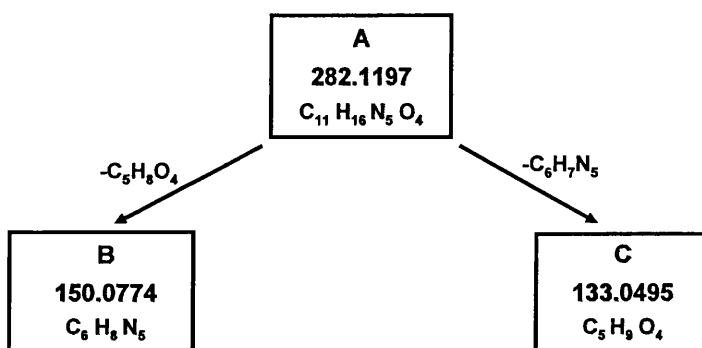


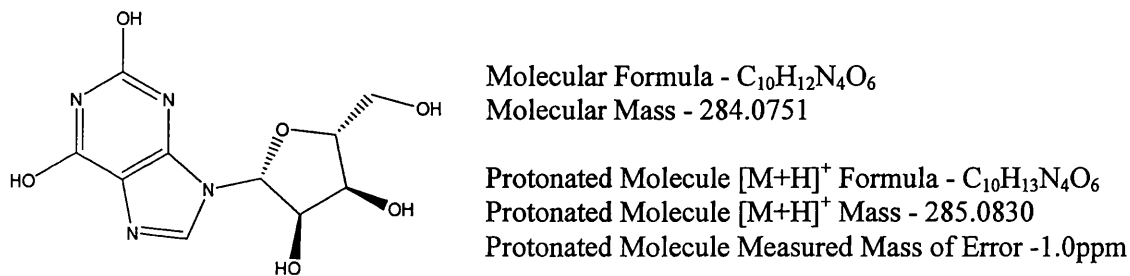
Figure 3.49: Product ion reaction scheme of the protonated unknown ion at  $m/z$  282 present in the haemodialysate.

The similarities of the fragmentation and elemental composition information suggest this ion could be one of three isomers of methyladenosine. However, published literature regarding the chromatographic behaviour of these isomers indicates that this analyte is modified nucleoside N<sup>6</sup>-methyladenosine.

#### 3.3.14. Xanthosine

This is another modified purine nucleoside that is known to accumulate within uremic patients. Previous investigations regarding xanthosine accumulation in uremic patients have involved detection by mass spectrometry but only in serum and urine.

These highlighted that increased levels of this modified nucleoside are observed in the urine and blood of uremic patients whether receiving renal treatment or not<sup>[49]</sup>.



*Figure 3.50: Chemical structure of the modified purine nucleoside xanthosine. When fragmented by collision induced dissociation it forms ions at  $m/z$  153 (the protonated purine base), 136 and 110.*

Following electrospray ionisation xanthosine is observed as a protonated molecule at  $m/z$  285 and fragments primarily into ions at  $m/z$  153, 136 and 110. The first transition from  $m/z$  285  $\rightarrow$  153 accounts for the loss of the ribose sugar from the protonated purine base, and the second from  $m/z$  153  $\rightarrow$  136 involves a loss of ammonia from this base. The full fragmentation mechanism has been elucidated and indicates that the MS/MS mass spectrum should contain the protonated molecule at  $m/z$  285 and the three fragment ions stated above<sup>[47, 48]</sup>.

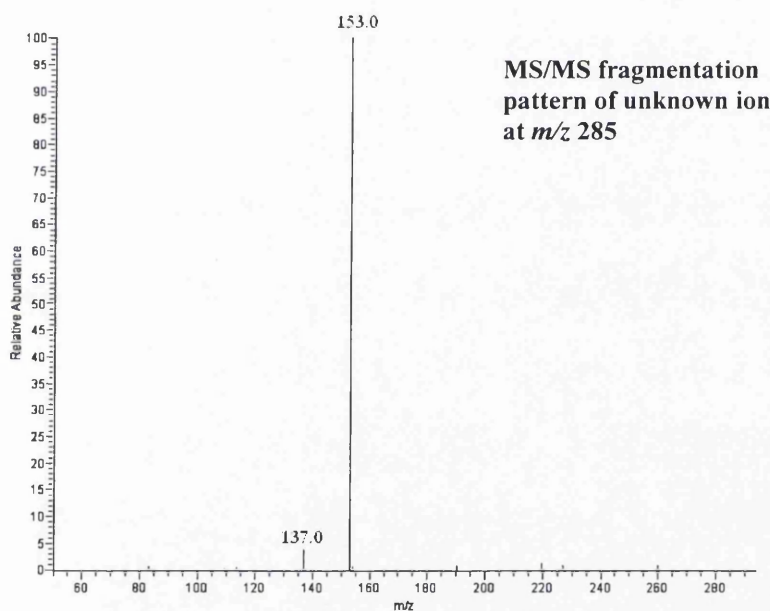
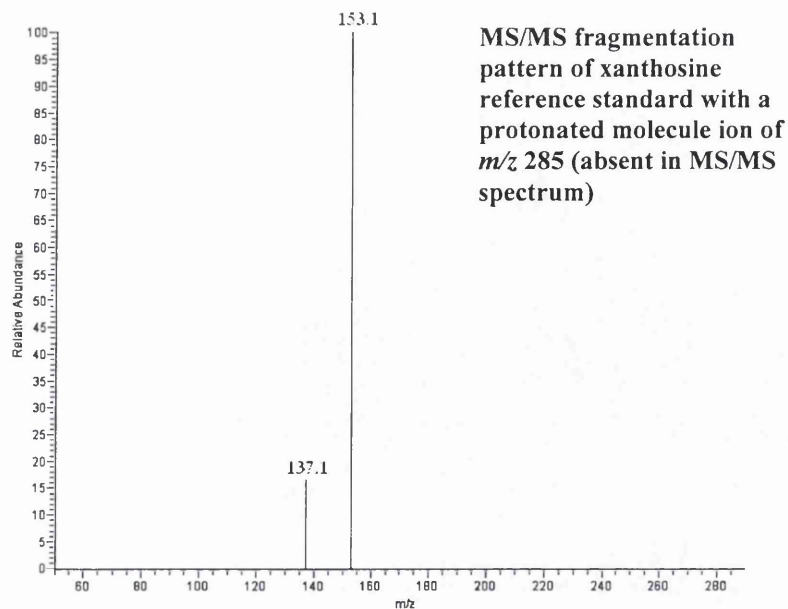


Figure 3.51: Fragmentation mass spectra of xanthosine standard reference material and an unknown ion of corresponding mass-to-charge present in the dialysate. Unfortunately, due to the low abundance of this analyte within the full mass scan mass spectrum only one stage of fragmentation was achieved.

An ion of corresponding accurate mass-to-charge and elemental formula of  $C_{10}H_{13}N_4O_6$  was observed within the haemodialysate solution at an error of -1.0ppm.

The unknown ion was fragmented and the calculated accurate masses determined for two of the fragment ions at  $m/z$  153 and 137. These had comparative elemental formulae for those product ions of xanthosine (see figure 3.52). From a comparison with previously published fragmentation information, and the commercial standard the similarities of MS/MS spectra indicate that this ion at  $m/z$  285 is xanthosine.

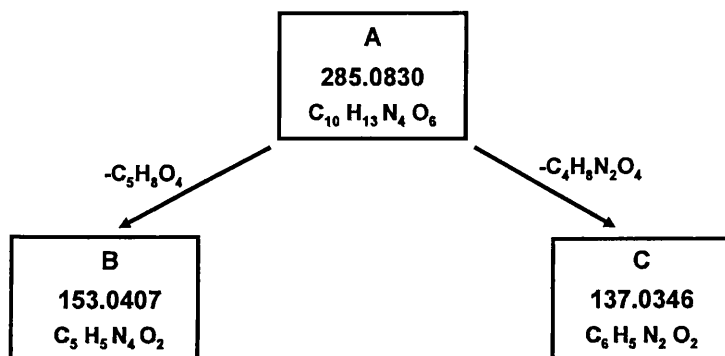
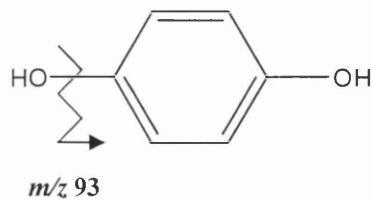


Figure 3.52: Fragment ion reaction scheme of the unknown ion at  $m/z$  285 observed in the haemodialysate solution and suspected to be the uremic toxin xanthosine.

### 3.3.15. Hydroquinone

This uremic analyte is part of a large class of toxins based on their common phenolic functionalities. Past research has indicated that an accumulation of hydroquinone can affect consciousness resulting in lethargy, in extreme cases leading to coma, and cause gastrointestinal bleeding<sup>[50]</sup>. This research has involved the quantitation of these phenolic compounds in serum by gas chromatography-mass spectrometry (GC-MS), but no previously published research has discovered it in the haemodialysis concentrate. This uremic analyte is essentially a double phenolic compound of mass 110Da as shown in figure 3.53.





Molecular formula -  $C_6H_6O_2$   
Molecular mass - 110.0362

Protonated Molecule  $[M+H]^+$  Formula -  $C_6H_7O_2$

Protonated Molecule  $[M+H]^+$  Mass - 111.0441

Protonated Molecule Measured Mass of Error -3.5ppm

Figure 3.53: Chemical structure of the uremic toxin hydroquinone. This is a member of the phenolic class of uremic toxins and has a protonated molecule of  $m/z$  111. Using collision induced dissociation it will fragment into two ions at  $m/z$  93 and 77 (see figure 3.54) accounting for successive losses of hydroxyl groups.

An ion observed in the mass spectrum of haemodialysate of corresponding mass-to-charge was fragmented and compared to the fragmentation pattern of hydroquinone standard reference material.

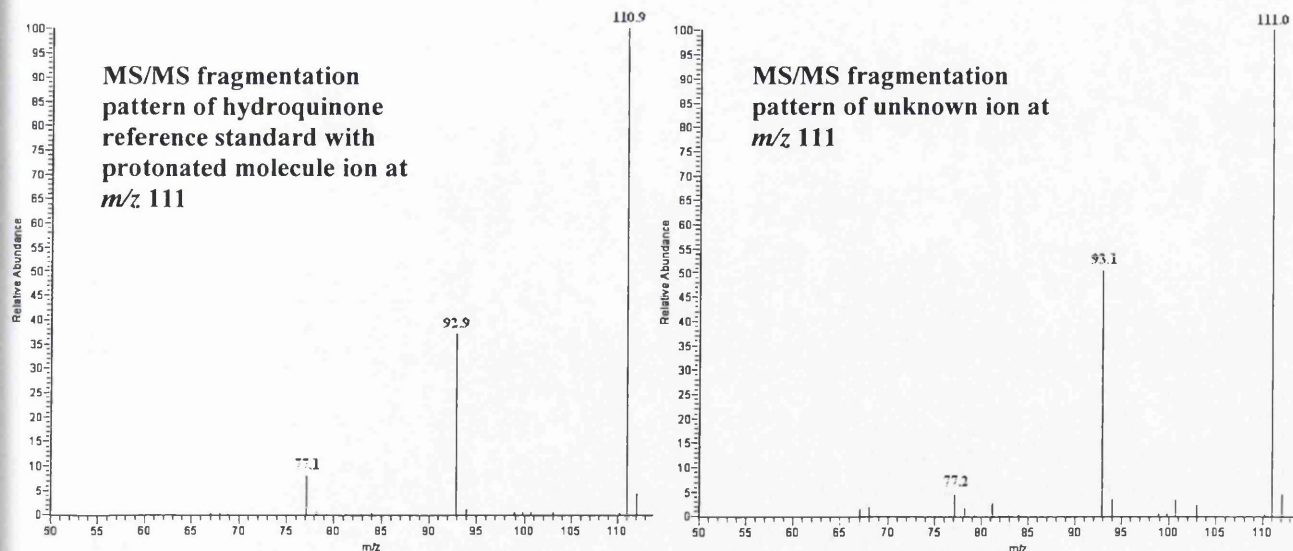


Figure 3.54: MS/MS fragmentation mass spectrum of hydroquinone standard reference material with a protonated molecule of  $m/z$  111 and that of a corresponding ion within the haemodialysate.

Hydroquinone shows very little fragmentation with the loss of the two hydroxyl groups and forming phenoxy and phenyl ring ions at  $m/z$  93 and 77 respectively (see figure 3.54). A visual comparison of the standard and unknown MS/MS spectra with the elemental formula generated from the calculated accurate mass, at an error of -3.5ppm, indicate that this ion is hydroquinone. This proposed identity is further supported by the accurate mass of the unknown fragment ion observed at  $m/z$  93 and has a suggested formula of  $C_6H_5O$  (see figure 3.55). A combination of the mass spectrometric and chromatographic reverse phase retention time, of 5 minutes corresponding to elution conditions of 95% aqueous:5% organic solvent, implies that hydroquinone is present in the haemodialysate solution.

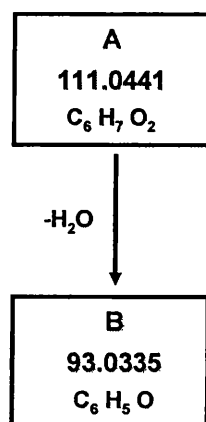


Figure 3.55: Fragment ion reaction scheme of the unknown ion at  $m/z$  111 thought to be the uremic analyte hydroquinone.

### 3.4. Method Validation Summary

Haemodialysate has previously been examined for uremic toxins<sup>[51, 52]</sup>. Of the few discovered the research involved detection by UV chromatography with identification according to wavelength and retention time. This analytical technique has much poorer sensitivity and specificity for analyte identification warranting a method that is

superior and applicable to a wider range of compounds. Modern ionisation modes used in liquid chromatography-mass spectrometry (LC-MS) techniques, such as electrospray (ESI) and atmospheric pressure chemical ionisation (APCI) are considered unsuitable for examining haemodialysate primarily due to the inherent high levels of salt and buffer. A consequence is that research using this technique for detecting uremic toxins has involved other biomatrices such as whole blood, serum, plasma and urine<sup>[49]</sup>. Therefore, the first section of this project was to develop a preparative protocol capable of improving haemodialysate compatibility for modern mass spectrometric analysis as a general screen. This novel methodology has been evaluated according to:

- i. the reproducibility of the chromatographic separations and,
- ii. the capability of the protocol to extract and detect known uremic toxins.

The first chromatographic dimension involved separation by size exclusion and detection by UV spectrophotometry, whilst monitoring both the pH and conductivity of the sample. Each mode of detection showed excellent inter- and intra-reproducibility both in terms of the signals obtained and their associated retention time. Following this assurance, samples were fractionated according to the UV activity of the chromatogram and each fraction subjected to additional separation by reverse phase chromatography as an online procedure into the mass spectrometers. Initial results had a high degree of variability shown by the UV chromatograms and mass spectra generated regarding peak intensity and retention times, and ions observed respectively. This prompted experiments to test sample stability on the autosampler while awaiting overnight analysis over a 16 hour period. These tests illustrated the poor sample stability at benchtop temperatures and the improved

stability of each fraction at a reduced temperature of 10°C during overnight experiments. At these conditions both the reverse phase and mass spectrometric data showed excellent inter- and intra-reproducibility within different aliquots of the same sample.

The second stage of the method validation involved assessing the capability of extracting and identifying known uremic toxins previously characterised in biomatrices other than haemodialysate. Possible identities were obtained by first calculating the accurate mass of an ion of corresponding mass-to-charge to known toxins and generating an elemental formula. This formula was compared to the relevant known toxins and the commercial standards obtained. Each standard was infused into the mass spectrometer and their MS/MS fragmentation patterns generated by collision induced dissociation (CID). The ions suspected to be known toxins were also fragmented and their product ion patterns compared to the standards. Toxin identities were assigned according to a combination of this data and chromatographic relative retention time. This enabled the detection and identification of 15 out of 90 known uremic toxins of varying chemical characteristics (protein bound/non-protein bound) with greater specificity and potential sensitivity than UV spectrophotometry in a sample previously considered unsuitable for modern mass spectrometry techniques. Unfortunately due to time constraints within the project we were unable to carry out any quantitation of these known toxins in order to accurately assess the sensitivity of the technique and the degree of removal by the dialysis procedure. However, the discovery of protein-bound uremic toxins thought unlikely to be present at sufficiently high levels for detection can be used as an approximate measure of sensitivity until the quantitative work is carried out.

The work in this chapter indicates that known toxins can be isolated and detected within haemodialysate by mass spectrometry as an online procedure with reverse phase chromatography, following a simple preparative step using size exclusion chromatography. This novel protocol for haemodialysate can also be applied to identify and characterise new uremic analytes.

## References

1. Dyer, JR., *Applications of absorption spectroscopy of organic compounds*. Prentice-Hall, New Jersey. 1965, p. 4-21.
2. *Section I: Measurement of renal function, when to refer and when to start dialysis*. Nephrology Dialysis and Transplantation (Suppl 7), 2002. **17**: p. 7-15.
3. Yang, BC., Khan, S., and Mehta, JL., *American Journal of Physiology*, 1994. **266**: p. H2212-H2219.
4. Berman, RS., and Martin, W., *British Journal of Pharmacology*, 1993. **108**: p. 920-926.
5. Kock, R., Delvoux, B., Sigmund, M., and Greiling, H., *European Journal of Clinical Chemistry and Clinical Biochemistry*, 1994. **32**(11): p. 837-842.
6. Vanholder, RC., De Smet, RV., and Ringoir, SM., *Clinical Chemistry*, 1992. **38**(8): p. 1429-1436.
7. Gehrke, CW., Waalkes, TP., Zumwalt, RW., Chang, SY., Lakings, DB., Tormey, DC., Ahmann, DL., and Moertel, CG., *Cancer*, 1975. **36**(2): p. 390-398.
8. Liebich, HM., Xu, G., Stephano, C., Lehman, R., Haring, HU., Lu, P., and Zhang, Y., *Chromatographia*, 1997. **45**(1): p. 396-401.
9. Amblard, F., Aucagne, V., Guenot, P., Schinazi, RF., and Agrofoglio, LA., *Bioorganic and Medicinal Chemistry*, 2005. **13**(4): p. 1239-1248.
10. Otto, MJ., *Current Opinion in Pharmacology*, 2004. **4**: p. 431-436.
11. Gerrits GP., Monnens LA., De Abreu RA., Schroder, CH., Trijbels, JMF., and Gabreels, FJM., 1991. *Nephron*. **58**: p. 310-314.
12. Daniewska-Michalska D., Motyl T., Gellert R., Kukulska, W., Podgurniak, M., Opechowska-Pacocha, E., and Ostrowski, K., *Nephron*, 1993. **64**: p. 193-197.
13. Asatoor, AM., *Clinical Chemica Acta*, 1968. **20**: p. 407-411.
14. Schoots, AC., Gerlag, PGG., Mudler, AW., Peeters, JAG., and Cramers, CAMG., *Clinical Chemistry*, 1988. **34**: p. 91-97.
15. Dzurik, R., Laidova, I., Spustova, V., and Opatrny, K Jr., *Nephron*, 1992. **61**: p. 64-67.
16. Niwa, T., Takeda, N., and Yoshizumi, H., *Kidney International*, 1998. **53**: p. 1801-1806.

17. Dudley, E., Tuytten, R., Bond, A., Lemiere, F., Brenton, AG., Esmans, EL., and Newton, RP., *Rapid Communications in Mass spectrometry*, 2005. **19**(21): p. 3075-3085.
18. Godfrey AR., *Haemodialysis International*, 2007. **11**: p. 278-285.
19. Bohlander, JM., Franke, S., Stein, G., and Wolf, G., *American Journal of Physiology – Renal Physiology*, 2005; **289**: p. F645-659.
20. Henle, T., Deppisch, R., Beck, W., Hergesell, O., Hänsch, GM., and Ritz, E., *Nephrology Dialysis Transplantation*, 1999; **14**: p. 1968-1975.
21. Lapolla, A., Fedele, D., Martano, L., Concetta Arico, N., Garbeglio, M., Traldi, P., Seraglia, R., and Favretto, D., *Journal of Mass Spectrometry*, 2001; **36**: p. 370-378.
22. Thornalley, PJ., Battah, S., Ahmed, N., Karachalias, N., Agalou, S., Babaei-Jadidi, R., and Dawnay, A., *Biochemical Journal*, 2003. **375**: p. 581-592.
23. Ahmed, N., Mirshekar-Syahkal, B., Kennish, L., Karachalias, N., Babaei-Jadidi, R., and Thornalley, PJ., *Molecular Nutrition and Food Research*, 2005. **49**: p. 691-699.
24. Teerlink, T., Barto, R., ten Brink, HJ., and Schalkwijk, CG., *Clinical Chemistry*, 2004. **50**(6): p. 1222-1228.
25. Tsai, J-C., Perrella, MA., Yoshizumi, M., Hsieh, C-M., Haber, E., Schlegel, R., and Lee, M-E., *Proceedings of the National Academy of Sciences*, 1994. **91**: p. 6369-6373.
26. McGregor, DO., Dellow, WJ., Lever, M., George, PM., Robson, RA., and Chambers, ST., *Kidney International*, 2001. **59**(6): p. 2267-2272.
27. Holm, PI., Magne Ueland, P., Kvalheim, G., and Lien, EA., *Clinical Chemistry*, 2003. **49**(2): p. 286-294.
28. Depner, TA., *Kidney International*, 1981. **20**: p. 511-518.
29. Gerber, I., ap Gwynn, I., Alini, M., and Wallimann, T., *European Cells and Materials*, 2005. **10**: p. 8-22.
30. Gurreri, G., Ghiggeri, G., Salvidio, G., Garibotto, G., Robaudo, C., and Deferrari, G., *Nephron*, 1986. **42**(4): p. 295-297.
31. De Deyn, PP., Marescau, B., Cuykens, JJ., Van Grop, L., Lowenthal, A., and De Potter, WP., *Clinica Chimica Acta*, 1987. **167**: p. 81-88.
32. De Deyn PP, Marescau, B., D'Hooge, R., Possemiers, I., Nagler, J., and Mahler, C., *Neurochemistry International*, 1995. **27**: p. 227-237.

33. De Deyn, PP., D'Hooge, R., Van Bogaert, PP., Marescau, B., *Kidney International Suppl.*, 2001. **78**: p. S77-S83.
34. Marescau, B., Nagels, G., Possemiers, I., De Broe, ME., Because, I., Billiouw, JM., Lornoy, W., and De Deyn, PP., *Metabolism*, 1997. **46**: p. 1024-1031.
35. May, RC., and Mitch, WE., *Pathophysiology of uremia: The kidney*, 1996. Saunders, Philadelphia. 5<sup>th</sup> Edition: p. 2148-2169.
36. Vanholder, RC, *Advanced Nephrology*, 1997. **26**: p. 143-162.
37. Zhen, Y., Krausz, KW., Chen, C., Idle, JR., and Gonzalez, FJ., *Molecular Endocrinology*, 2007. **21**(9): p. 2136-2151.
38. Saito, A., Niwa, T., Maeda, K., Kobayashi, K., Yamamoto, Y., and Ohta, K., *The American Journal of Clinical Nutrition*, 1980. **33**: p. 1402-1406.
39. Sprince, H., *Clinical Chemistry*, 1961. **7**(3): p. 203-230.
40. Becker, BF., *Free Radical Biological Medicine*, 1993. **14**: p. 615-631.
41. Lehto, S., Niskanen, L., Ronnema, T., and Laakso, M., *Stroke*, 1998. **29**: p. 635-639.
42. Silverstone, PH., McGrath, BM., and Kim, H., *Bipolar Disorders*, 2005. **7**: p. 1-10.
43. Sarubbi, D., McGiff, JC., and Quilley, J., *American Journal of Physiology*, 1989. **257**: p. F762-F768.
44. Kasiske, BL., O'Donnell, MP., and Keane, WF., *Diabetes*, 1985. **34**: p. 360-364.
45. Estmatjes, E., Fernandez, MR., Halperin, I., Camps, J., Gaya, J., Arroyo, V., Rivera, F., and Figuerola, D., *Journal of Clinical Endocrinology Metabolism*, 1985. **60**: p. 1231-1236.
46. Larkins, RG., and Dunlop, ME., *Diabetologia*, 1992. **35**: p. 499-504.
47. Dudley, E., Lemiere, F., Van Dongen, W., Tuttyen, R., El-Sharkawi, S., Brenton, AG., Esmans, EL., and Newton, RP., *Rapid Communications in Mass Spectrometry*, 2004. **18**(22): p. 2730-2738.
48. Kammerer, B., Frickenschmidt, A., Muller, CE., Laufer, S., Gleiter, CH., and Liebich, H., *Analytical and Bioanalytical Chemistry*, 2005. **382**: p. 1017-1026.
49. Niwa, T., *Mass Spectrometry Reviews*, 1998. **16**: p. 307-332.
50. Niwa, T., Aiuchi, T., Nakaya, K., Emoto, Y., Miyazaki, T., and Maeda, K., *Clinical Nephrology*, 1993. **39**: p. 92-96.
51. Fridolin, I., and Lindberg, LG., *Medical and Biological Engineering and Computing*, 2006. **41**(3): p. 263-270.



52. Senflber, FC., Halline, AG., Veering, H., and Dayton, DA., *Clinical Chemistry*, 1976. 22(9): p. 1522-1527.

## **CHAPTER 4:**

# **Identification and Evaluation of Novel Uremic Analytes for Assessing Haemodialysis Adequacy in Patient Specific Samples**

### **4.1. Introduction**

Six haemodialysate samples of 100mL were collected in previously weighed glass amber jars from 8 patients in total. Two patient sets were compared; those who exercised for an hour before dialysis commenced and those who dialysed without prior exercise. This comparison was intended as a pilot study to compare the effect of exercise on the excretion of uremic analytes. It is hypothesised that exercise will increase blood flow, encouraging analytes predominantly present in tissue to pass into blood, thus increasing the speed and quantity of solutes removed during the dialysis session. Samples were obtained at the following time points throughout the haemodialysis session; start ( $t_0$ ), 30 minutes ( $t_{0.5hr}$ ), 1 hour ( $t_{1hr}$ ), 1.5 hours ( $t_{1.5hr}$ ), 2.5 hours ( $t_{2.5hr}$ ) and at session end ( $t_{4hr}$ ). After collection samples were frozen, lyophilized (Edwards, Crawley, UK) and the resulting solid weighed. A tenth of the each solid was then removed and reconstituted in the minimum amount of deionised water, where 1mg requires approximately 1mL of water. These aqueous samples were then filtered and separated by size exclusion chromatography as in section 3.2.2 in Chapter 3.

## 4.2. Comparative LC-MS and LC-MS/MS Investigations of Pooled and Patient Specific Haemodialysate

### 4.2.1. Selecting candidate biomarkers

The validation of the developed analytical methodology (see Chapter 3) has indicated that the interrogation of these samples must involve fresh size exclusion fractions of all samples. The analyses consisted of an online separation, as described in Chapter 2, section 2.2.2, into an LCQ DECA ion trap or LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). A full mass scan from a mass-to-charge of 50-2000 was obtained for each fraction. An in-depth examination of the data showed many ions present at exact or similar relative retention times in both samples with different signal intensities. Blank dialysate was also examined and indicated that these analytes were removed from the patients' bloodstream by dialysis and may exist at elevated levels. These characteristics are considered as two of the six prerequisites for the ideal biomarker in assessing dialysis adequacy. A criterion was applied to these data, due to the large number of candidates generated and excluded ions of poor stability. Therefore candidates chosen had less than a 10% change in peak area for the duration of 16 hours stability test. The resulting ions were selected for additional mass spectrometry experiments. There were 12 ions in total that satisfied these requirements mainly distributed within 2 size exclusion fractions, at retention times of 10-13 minutes and 16-19 minutes. This data, when compared to the retention times obtained for the standards indicate that the candidate ions have molecular weights less than 350Da. However, closer inspection of the mass spectral data shows that some of the candidate ions do have a molecular weight greater than this of up to 875.0Da possibly as a result of aggregate formation.

Fraction	Retention Time (min)	Relative Retention Time (min)	$m/z$ of $[M+H]^+$	% Difference Change 0-16hr
<b>F2 (10-13min)</b>	4.34	0.52	241.0	3.52
	5.93	2.43	214.2	9.39
	26.99	23.18	851.1	7.73
	27.32	23.52	576.0	-9.14
	29.91	26.11	875.0	6.57
	34.73	30.93	764.7	-3.65
	39.37	35.57	275.1	9.40
<b>F4 (16-19min)</b>	7.17	2.58	270.2	6.72
	14.91	10.32	504.7	-9.46
	28.87	24.28	380.1	8.99
	34.86	30.27	646.7	-7.74
			335.0	1.31

*Table 4.1: Possible candidate biomarkers that are thermally stable at 10°C and removed by the dialysis procedure.*

The ions below mass of 400Da were selected for additional mass spectrometric investigations where both their fragmentation patterns and accurate mass, to provide a possible elemental composition, were obtained. This cut-off was applied in order to simplify the number of elemental formulae generated and to maximise the chance of assigning identities within the time allocated for this project.

#### 4.2.2. UV characterisation of novel biomarkers

The online analyses involved analyte detection by both mass spectrometry and the less specific UV spectrophotometry. UV detection is a more amenable technique for use as a general screen within a hospital laboratory, and therefore the capability of novel uremic analytes to produce a discernable UV signal is advantageous. At the maximum absorbance wavelength (242nm) for the dialysate sample two of the six

analytes of masses below 400Da ( $m/z$  270 and 275), displayed a discernable UV signal indicating the presence of a chromophore assignable to a diazoacetic ester, *cis*-crotonic acid, an unsaturated ketone or a ketone monosubstituted benzene ring<sup>[1]</sup>.

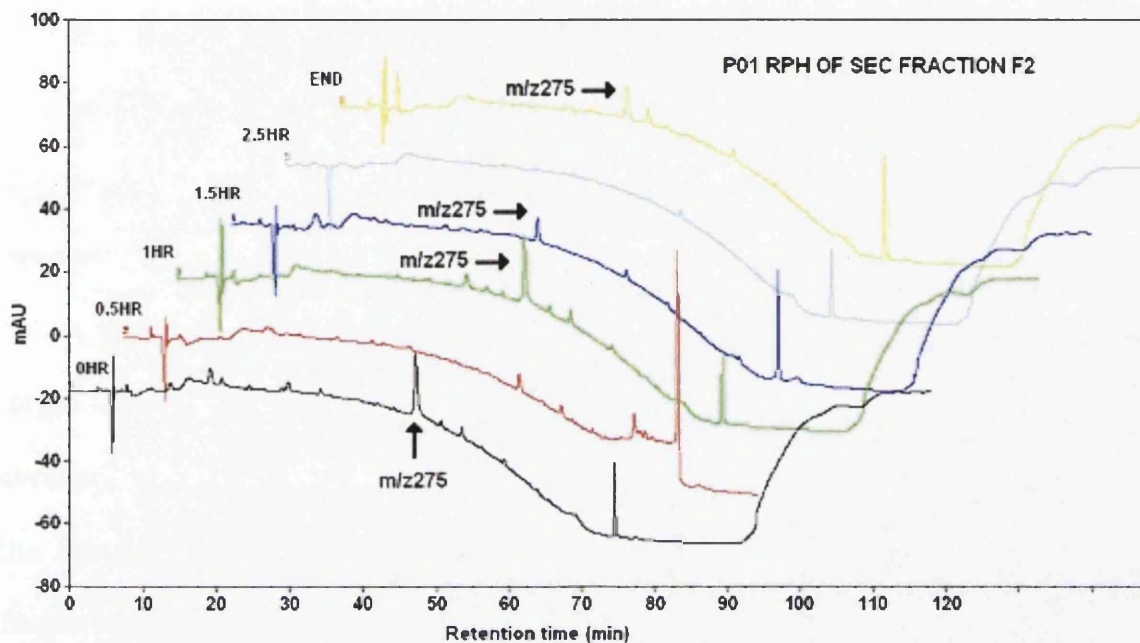


Figure 4.1: UV chromatogram of uremic analyte at  $m/z$  275 within patient P01 showing its removal into the dialysate solution throughout treatment at 0, 0.5, 1, 1.5, 2.5 hours and end of session.

#### 4.2.3. Mass spectrometric structural elucidation

##### 4.2.3.1. Accurate mass

This is essentially the mass of an ion to within four decimal places. Once obtained elemental composition information is suggested as only specific combinations of elements will match the obtained formulae. The suggested elemental formulae are evaluated according to four main parameters as described in Chapter 1. The search and assignment criteria for these analyses are included in Chapter 2, section 2.2.3.2.,

and all ion accurate masses were obtained within an error of 5ppm unless stated otherwise.

#### 4.2.3.2. Structure Investigations Using MS/MS

This was carried out initially using the LCQ DECA ion trap mass spectrometer and then confirmed with the LTQ Orbitrap mass spectrometer. It is required to illustrate that fragmentation variability is not significant between instruments of similar design operated under analogous conditions. The fragmentation patterns were obtained at collision energies necessary for obtaining signals of the product ions using both a target ion search followed by a data dependent analysis approach. Therefore the eventual total ion chromatograms consisted of peaks at the relative retention times for the chosen ions contained in table 4.2 followed by its fragmentation, and further fragmentation of the resulting ions. Comparative fragmentation patterns, obtained with the LTQ Orbitrap were also acquired at low collision energies to obtain information of the product ions while maintaining a signal of the parent ion to monitor any mass drift. By applying the same accurate mass data analysis procedure to the fragment ion, elemental composition information for this ionic species was also obtained. This can aid in the organisation and translation of accurate mass and fragmentation data to obtain the most likely chemical structure for the unknown. Therefore, a parent ion elemental formula was chosen according to the lowest error from the accurate mass and its agreement with the elemental formula obtained for the fragment ions.

SEC Fraction	Parent Ion ( $m/z$ )	MS <sup>2</sup> Ions ( $m/z$ )	MS <sup>3</sup> Ions ( $m/z$ )
F2 (10-13 min)	241.0	222.9, 194.8, 177.8, 153.8, 151.8, 121.8	153.8 → 136.8, 89.9 194.8 → 121.9
	214.2	196.9, 157.8, 155.0, 154.1	196.9 → 179.8, 154.9, 154, 137.9, 125.9
	275.1	263.2, 210.8	
F4 (16-19min)	270.2	270.0, 253.0, 236.1, 211.0, 210.1	253.0 → 234.0, 224.9, 210.0
	180.1 (359.1, 381.1)	180.1, 162.1, 134.1, 105.0, 95.1, 77.1	162.1 → 134.1, 105.0 105.0 → 77.0
	335.1	316.8, 292.9, 275.8, 246.9, 223.4, 205.8, 204.0, 151.8	317 → 299, 257 223 → 206, 147

*Table 4.2: Fragment ions obtained for the candidate biomarker ions using the LCQ DECA ion trap mass spectrometer. Results were acquired at collision energy of 35eV to maximise the fragmentation of the parent ions.*

Following the empirical assignment of an elemental formula, chemical structures may be suggested. Fragmentation patterns can be highly specific to a chemical structure due to the spatial differences of the elements and functional groups. Thus, product ions involving several stages of fragmentation may be used as evidence for analyte identification<sup>[2]</sup>. To minimise man hours deciphering fragmentation patterns of these parent ions a data analysis package, Mass Frontier 3.0 (Thermo Fisher Scientific), was used to generate product ions of all of chemical structures. This software considers numerous likely reactions (shown in the figure 4.2) that can occur after protonation, such as hydrogen rearrangements, charge induction, charge rearrangements (localisation), and loss of neutral moieties like water or carbon monoxide. These

predicted product ions were compared to those obtained using both the ion trap and LTQ Orbitrap mass spectrometers and used to match the appropriate chemical structure to the unknown. Other reactions not included in the software operating parameters, such as homolytic cleavages were considered and manually elucidated for the relevant ions.

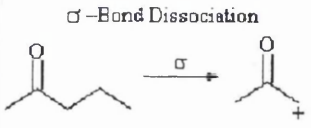
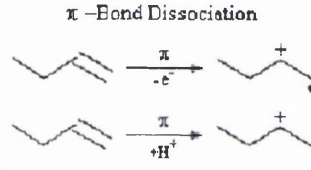
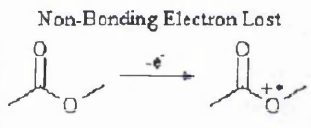
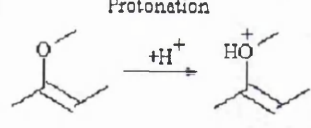
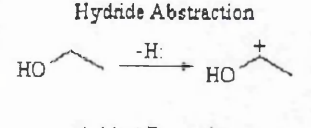
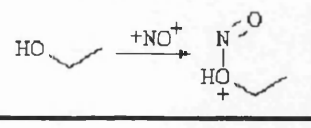
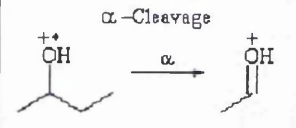
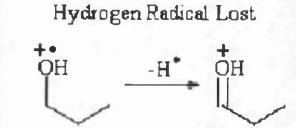
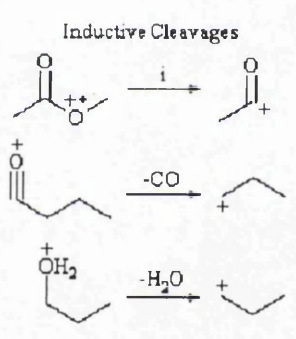
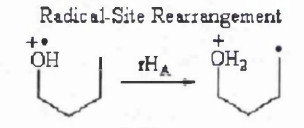
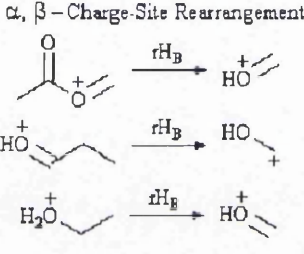
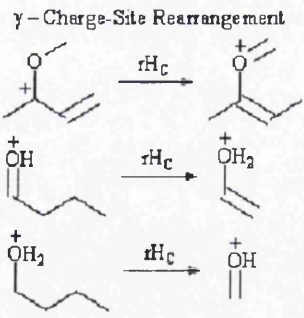
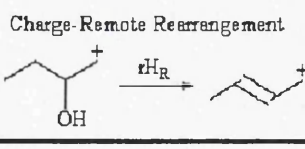
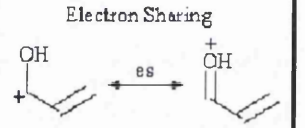
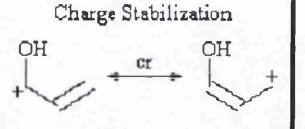
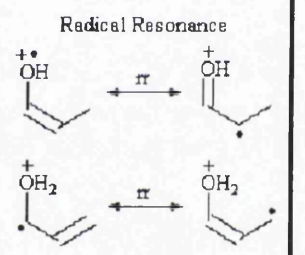
Ionization	Cleavages	Hydrogen Rearrangements	Electron Shifts
<p><math>\sigma</math>-Bond Dissociation</p>  <p><math>\pi</math>-Bond Dissociation</p>  <p>Non-Bonding Electron Lost</p>  <p>Protonation</p>  <p>Hydride Abstraction</p>  <p>Adduct Formation</p> 	<p><math>\alpha</math>-Cleavage</p>  <p>Hydrogen Radical Lost</p>  <p>Inductive Cleavages</p> 	<p>Radical-Site Rearrangement</p>  <p><math>\alpha, \beta</math>-Charge-Site Rearrangement</p>  <p><math>\gamma</math>-Charge-Site Rearrangement</p>  <p>Charge-Remote Rearrangement</p> 	<p>Electron Sharing</p>  <p>Charge Stabilization</p>  <p>Radical Resonance</p> 

Figure 4.2: Summary table of ion reactions considered for the predicted fragmentation patterns reproduced by Mass Frontier 3.0 software.



4.2.4. Discussion of results4.2.4.1. Candidate biomarker 1:  $m/z$  241.0296

Using several different search parameters the most probable elemental composition is shown in the table below.

Elemental Formula of Ion	Accurate Mass of Formula	Error (ppm)	Suitable Isotope Pattern?	DBE	Measured Accurate Mass of Unknown
$C_5H_9N_2O_9$	241.0303	-2.5	✓	2.5	241.0296

*Table 4.3: Accurate mass summary table showing the parameters determining the likelihood of the elemental formula matching the unknown uremic analyte.*

The calculated accurate mass for this ion suggests that it has an elemental formula of  $C_5H_9N_2O_9$ . Using the nitrogen and double bond rules this formula appears to be a protonated even electron ion and will contain either one aliphatic ring and one double bond or two double bonds. The fragmentation of this ion using the LCQ DECA and LTQ Orbitrap mass spectrometers indicate that the chemical structure is capable of successive losses of water, illustrated clearly by the several losses of 18Th between fragments  $m/z$  241 to 223, 223 to 205, and 205 to 187 (see figure 4.3 and 4.7). This is supported by the elemental formula suggested for each fragment ion from their calculated accurate mass (see table 4.4).

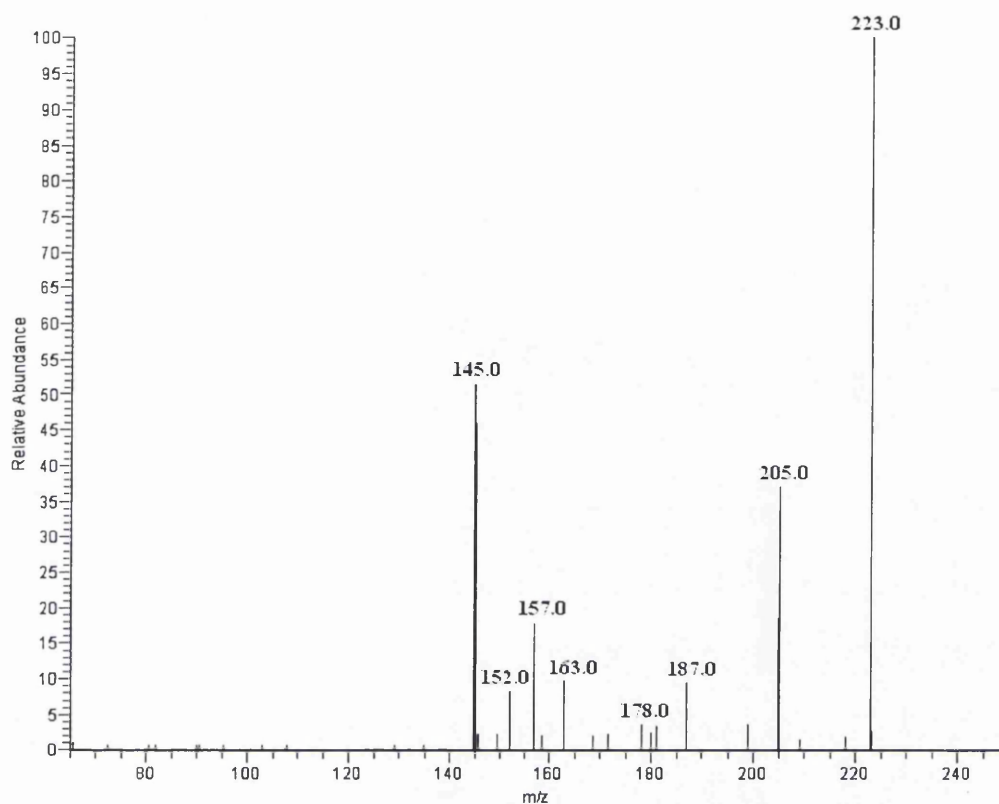


Figure 4.3: Fragmentation MS/MS mass spectrum of novel uremic analyte at  $m/z$  241 obtained using the LCQ DECA at a collision energy of 30eV. Comparative patterns were obtained using the LTQ Orbitrap generating elemental formula from the calculated accurate mass of each fragment ion measured at an error of less than 5ppm unlike  $m/z$  163, 157, 145 and were acquired within 10ppm.

Fragment Ion ( $m/z$ )	Elemental Formula	DBE Value
223	$C_5H_7N_2O_8$	3.5
205	$C_5H_5N_2O_7$	4.5
187	$C_5H_3N_2O_6$	5.5
178	$C_4H_4NO_7$	3.5
163	$C_3HNO_7$	4.0
157	$C_4HN_2O_5$	5.5
152	$C_2H_2NO_7$	2.5
145	$C_3HN_2O_5$	4.5

Table 4.4: Elemental formula obtained from the calculated accurate mass of the fragments produced from the precursor ion at  $m/z$  241.

The data from these mass spectrometers enables a fragmentation reaction scheme to be generated as shown in figure 4.4.

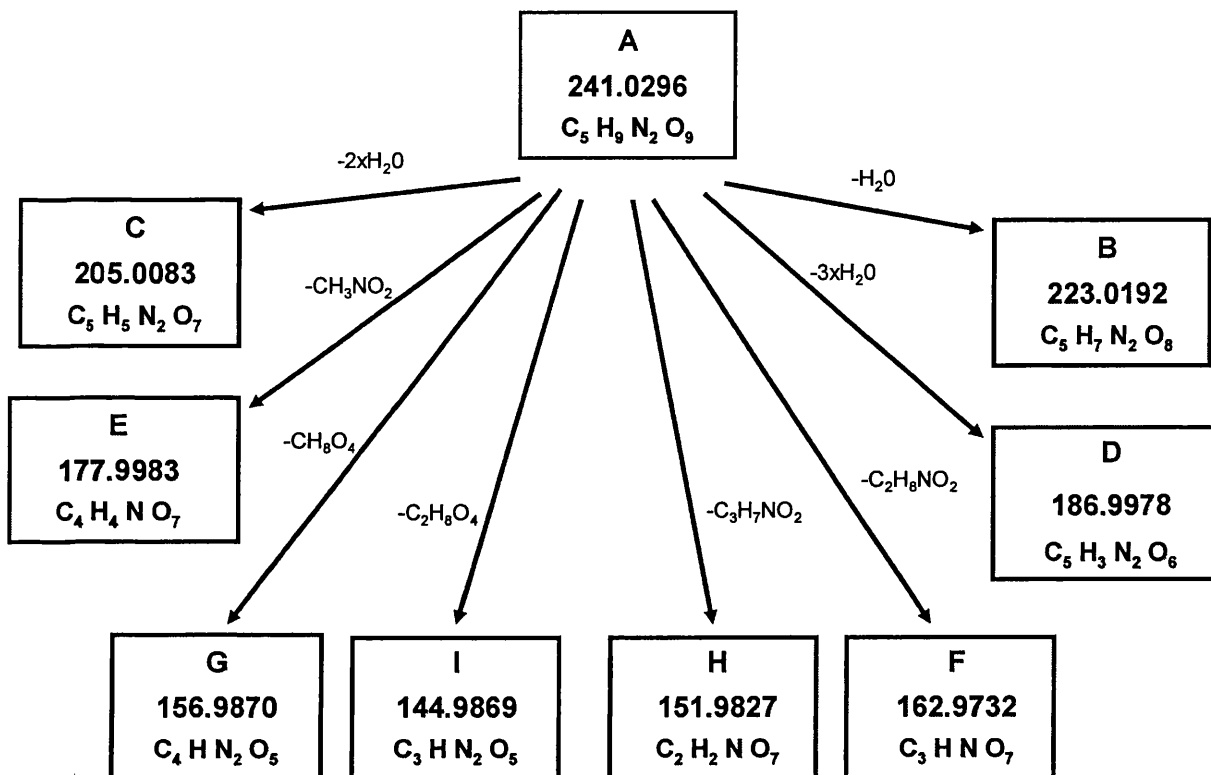


Figure 4.4: Fragmentation map of the novel uremic analyte at  $m/z$  241 showing the elemental formula of each fragment and the resulting loss.

This combined information is specific to a certain chemical structure named 5-(amino-1,2,-dihydroxy-ethyl)-3-nitrosooxy-[1,2,4]trioxine-3,6-diol (Beilstein, CrossFire, MIMAS, Manchester, UK)<sup>[3]</sup>, and is shown in figure 4.5 with all fragmentation mechanisms in Appendix 4.1. According to the calculated accurate mass and elemental formula of the ion, the double bond equivalence value suggests that this structure should contain a total of two double bonds and/or cyclic functionalities.

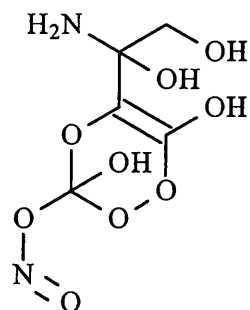


Figure 4.5: Proposed structure of the neutral analyte at with calculated ion accurate mass of  $m/z$  241.0296 and corresponding to the elemental formula of  $C_5H_9N_2O_9$ .

However, this is not apparent for the neutral structure, unless the double bond is present within the aliphatic ring. In solution this analyte will have the structure as shown in figure 4.6 and provides additional evidence regarding the fundamental structure of the ion.

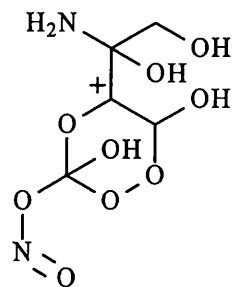


Figure 4.6: Proposed structure of the ion at  $m/z$  241. The protonation of the cyclic ring forms an even electron ion generating the correct DBE value of 2.5 as suggested by the calculated accurate mass and elemental formula.

This structure has several hydroxyl groups capable of losing water as a neutral moiety as successive fragments and accounting for the product ions at  $m/z$  223, 205 and 187 (see figure 4.7).

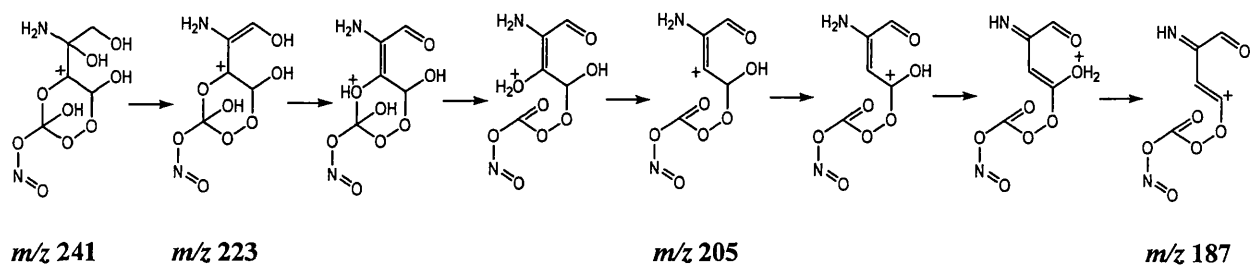


Figure 4.7: Reaction mechanism of transitions involving fragment ions at  $m/z$  223, 205 and 187.

The fragment at  $m/z$  178 is also formed through a number of common neutral losses from the aliphatic section in the upper region of the protonated molecule. For example, an initial ejection of ammonia forming an ion at  $m/z$  224 is followed by a hydrogen rearrangement with a loss of carbon monoxide to  $m/z$  196 and then water to generate the ion  $m/z$  178.

The next product ion at  $m/z$  163 has an elemental formula of  $C_3HNO_7$  which according to the nitrogen rule is an odd electron ion or radical cation. Hence, this may form as a result of a homolytic cleavage at the upper aliphatic section initially generating a loss of  $CH_3OH$  and produces the radical cation by a cleavage between the aliphatic section and the cyclic 6-membered ring (see figure 4.8).

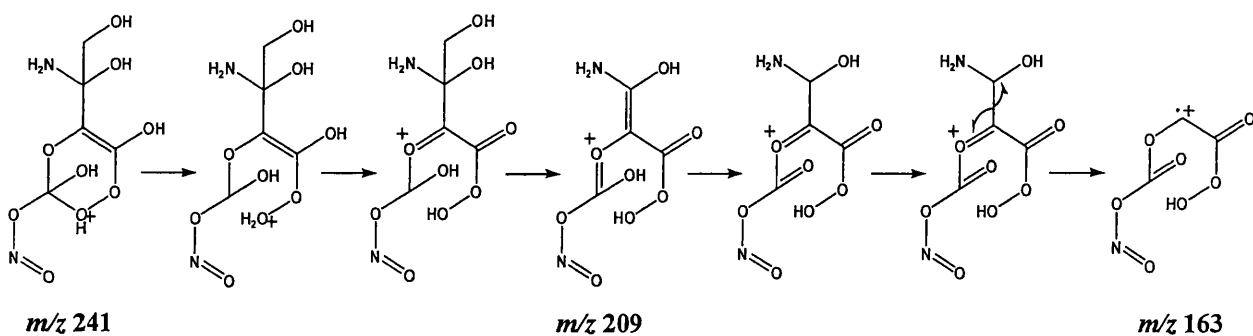


Figure 4.8: Suggested fragmentation mechanism involving an initial heterolytic cleavage to form an ion at  $m/z$  209. The position of the double bond and positive

charge on the saturated ketone group enables this ion to lose the upper hydroxyl amine structure by a homolytic cleavage generating the radical cation at  $m/z$  163.

The next ion present in the MS/MS mass spectrum is at  $m/z$  157 and has the elemental formula of  $C_4HN_2O_5$ . According to the nitrogen rule this is an even electron ion and can be produced as a result of several heterolytic cleavages including ring opening causing losses of water and a loss again involving the upper aliphatic section. This pattern of fragmentation involving this upper section is followed for the remaining product ions primarily due to the high stability associated with the nitroxide-section of the protonated molecule. This functional group enables the formation of the ion at  $m/z$  152. It is generated following the complete loss of the aliphatic section and part of the cyclic ring, unsaturating the *ortho* situated hydroxyl group into a carbonyl group and the observed product ion. The structure of each fragment described above was also confirmed by the DBE value calculated from the accurate mass and elemental formulae (see table 4.4).

An extensive examination of the calculated accurate mass, elemental formula and fragmentation data has shown, with the application of established interpretative rules, that the ion at  $m/z$  241 has the chemical structure shown in figure 4.4. Using the Beilstein database we can assign the name of 5-(amino-1,2,-dihydroxy-ethyl)-3-nitrosooxy-[1,2,4]trioxine-3,6-diol to the novel uremic analyte of mass 240Da.

4.2.4.2. Candidate biomarker 2:  $m/z$  214.1298

The calculated accurate mass obtained for this ion indicates that it has the elemental formula  $C_8H_{16}N_5O_2$ .

Elemental Formula of Ion	Accurate Mass of Formula	Error (ppm)	Suitable Isotope Pattern?	DBE	Measured Accurate Mass of Unknown
$C_8 H_{16} N_5 O_2$	214.1299	-0.0	✓	3.5	214.1298

*Table 4.5: Elemental composition and structural information suggested from the calculated accurate mass for the novel uremic toxin at  $m/z$  214.1298.*

According to the nitrogen rule and a double bond value of 3.5 this molecular ion is a singularly protonated species containing a total of 3 double bonds and/or aromatic rings. This in addition to the fragmentation data can be used to identify the relevant chemical structure for the ion. This structural information was obtained with both the LCQ DECA and LTQ Orbitrap mass spectrometers, with the latter allowing the assignment of fragment elemental formula and subsequent neutral losses (see table 4.6).

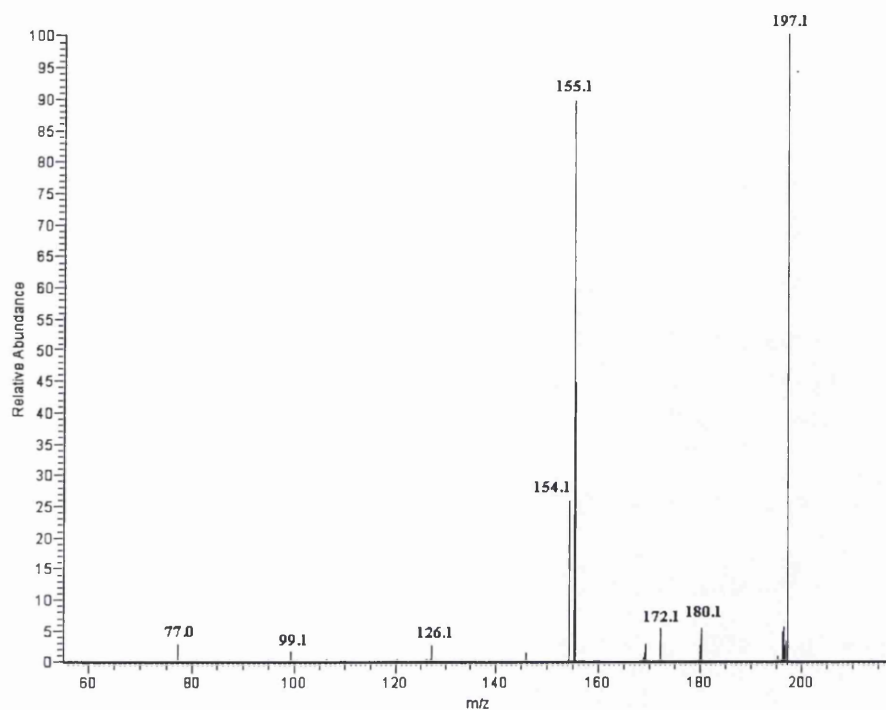


Figure 4.9: MS/MS mass spectrum of novel uremic analyte  $m/z$  214.1298 showing its fragments at a collision energy of 28eV.

The fragmentation data can be arranged to form a degradation reaction scheme of this ion (figure 4.10) and aid in assigning a chemical structure to this unknown.

Fragment Ion ( $m/z$ )	Suggested Elemental Formula	DBE Value
197	$C_8H_{13}N_4O_2$	4.5
180	$C_8H_{10}N_3O_2$	5.5
172	$C_7H_{14}N_3O_2$	2.5
169	$C_8H_{13}N_2O_2$	3.5
155	$C_7H_{11}N_2O_2$	3.5
154	$C_7H_{10}N_2O_2$	4.0
126	$C_6H_{12}N_3$	2.5
99	$C_4H_9N_3$	2.0
77	$C_6H_5$	4.5

Table 4.6: Elemental formula information obtained from the calculated accurate mass of fragment ions generated by collision induced dissociation of the precursor ion at  $m/z$  214. All were measured within an error of 5ppm apart from  $m/z$  77 (-20ppm).



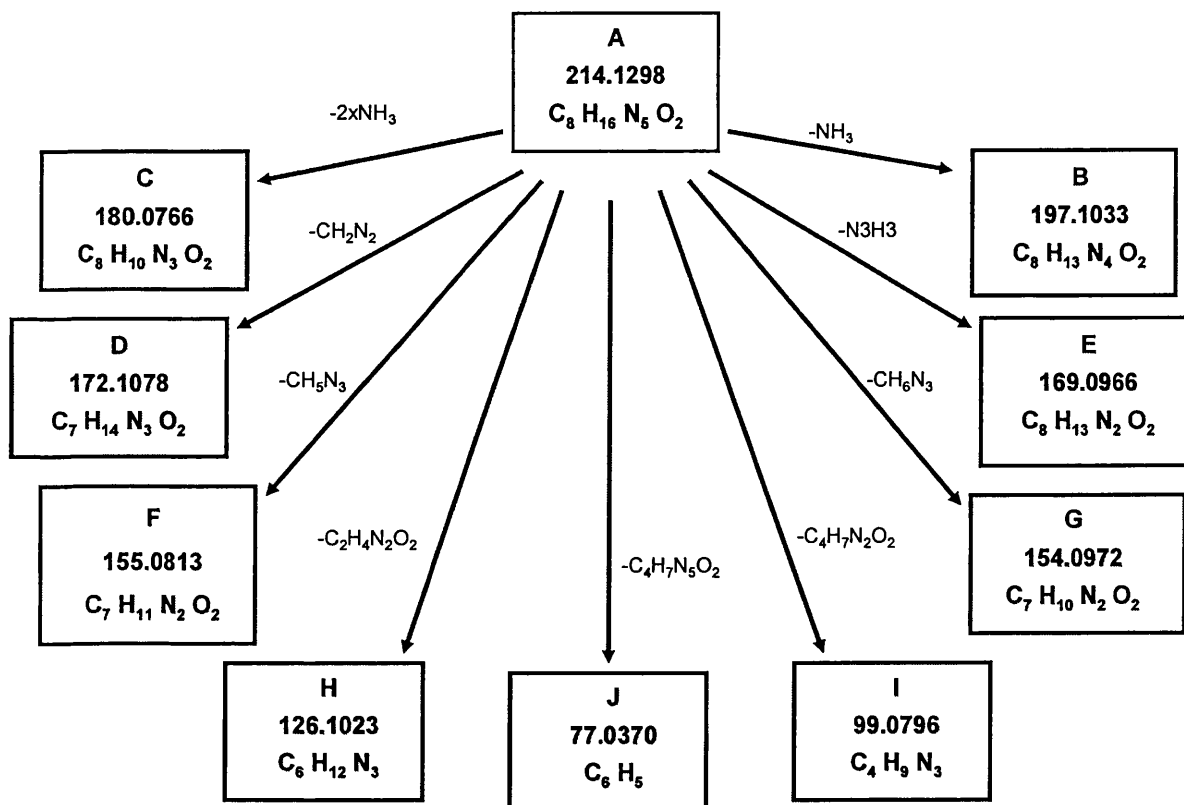


Figure 4.10: Fragmentation reaction scheme of the unknown novel uremic analyte at  $m/z$  214.1298. The initial loss of neutral ammonia and the presence of a phenyl ion at  $m/z$  197 and 77 respectively, are consistent with a primary amine group and an aromatic phenyl ring being present in the structure. The interpretation of this fragmentation information can aid in assigning a chemical structure and therefore an identity to this unknown.

Of the many structures with this elemental formula only a certain configuration will form the relevant fragmentation pattern exactly. An analyte capable of this fragmentation is 2-(5,6-diamino-6-diazenyl-cyclohex-1-enyl)-2-hydroxy-acetimidic acid and has the structure as shown in figure 4.11 with the relevant mechanisms in Appendix 4.2.

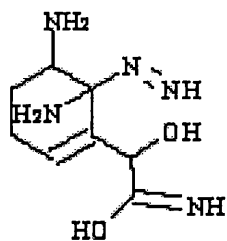


Figure 4.11: Proposed chemical structure of unknown uremic analyte at  $m/z$  214. This is 2-(5,6-diamino-6-diazenyl-cyclohex-1-enyl)-2-hydroxy-acetimidic acid and is novel to all biological systems with no published data regarding its existence.

From the elements present we calculate that this has a total of 4 double bonds and ring structure as a neutral molecule and does not correspond to that obtained for the unknown ion at  $m/z$  214. However, when this structure is charged as  $[M+H]^+$  the additional proton can rearrange within the aliphatic ring and situate itself at the double bond. This reduces the double bond equivalence (DBE) value to that of the ion at  $m/z$  214 of three and we can therefore propose that in solution this analyte actually has the structure shown in figure 4.12.

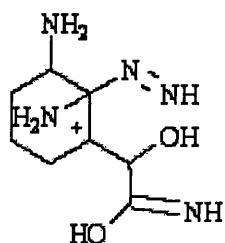


Figure 4.12: Proposed chemical structure of 2-(5,6-diamino-6-diazenyl-cyclohex-1-enyl)-2-hydroxy-acetimidic acid in solution with a DBE value of 3.

The two primary amino groups of this structure account for the initial losses of ammonia forming ions at  $m/z$  197 and 180, and could be successive. Unfortunately,

these fragments could not withstand further fragmentation ( $MS^3$ ) and this proposition will remain a hypothesis until confirmed. The next fragment at  $m/z$  172 is believed to form following the loss of  $CH_2N_2$  involving the rearrangement of the positive charge around the aliphatic ring. This will initially involve the loss of the amino groups as shown in figure 4.13.

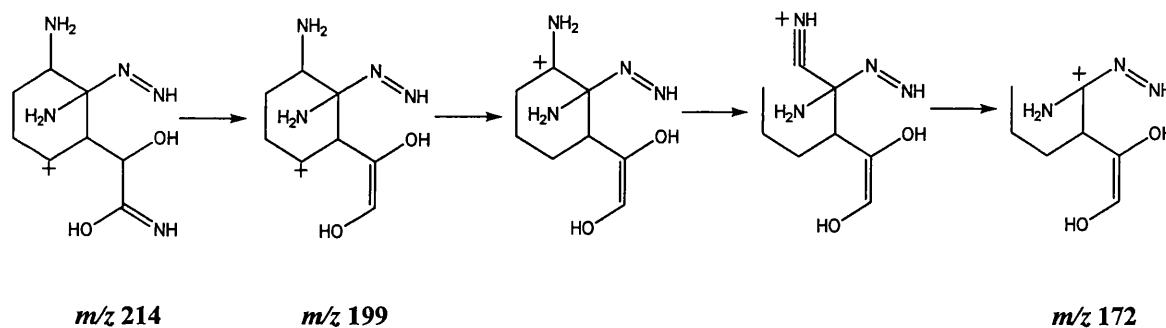


Figure 4.13: Reaction mechanism for the fragmentation of  $m/z$  214 to 172.

The fragment of  $m/z$  169 offers specific information regarding chemical structure as it is formed as a result of losing  $N_3H_3$ . This is indicative of double bonded nitrogen atoms that can be lost in two parts, firstly by releasing ammonia and secondly by the expulsion of a neutral diatomic nitrogen molecule (see figure 4.14).

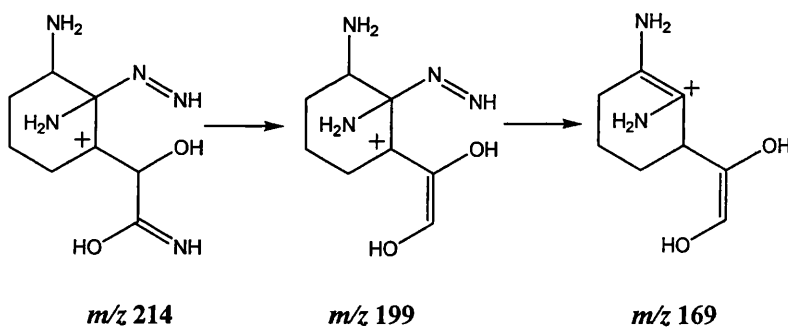


Figure 4.14: Reaction mechanism for the fragmentation of  $m/z$  214 to 169.

Unlike these last two transitions, the formation of the ion at  $m/z$  155 is much more favourable primarily due to the stability of the neutral product formed. This  $\text{CH}_5\text{N}_3$  fragment incorporates or can form due to the double bonded azide structure including the linking carbon and primary amine groups following a number of hydrogen rearrangements and abstractions (see figure 4.15).

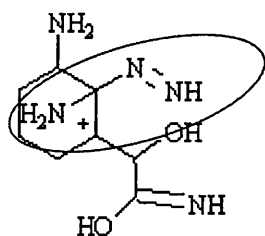


Figure 4.15: Proposed structure of the novel uremic analyte ion at  $m/z$  214 highlighting the group lost to form the fragment observed at  $m/z$  155.

Also present within the MS/MS mass spectrum is a peak at  $m/z$  154. Initially we suspected the loss of a singular proton as this has been observed in structures containing carbonyl groups. However, according to the nitrogen rule this is an odd electron ion indicating the loss of a functional group by a homolytic cleavage. According to the elemental formula suggested this is the radical cation of the azide ion suggested for  $m/z$  155.

The initial loss of ammonia from the protonated molecule producing the ion at  $m/z$  197 can prime the structure to break and form another fragment ion at  $m/z$  126. This is generated through a number of hydrogen rearrangements as shown in Appendix 4.2, losing first the section involving the hydroxyl and amino group, and then a neutral molecule of carbon monoxide. The two fragments below 100Th are

generated as a result of the breaking the aliphatic ring structure. According to the nitrogen rule the ion at  $m/z$  99 is a radical cation formed by a homolytic cleavage involving the amino section of the ring. The final ion present in the MS/MS mass spectrum at  $m/z$  77 is characteristic of a phenyl ring and is confirmed by the elemental formula of  $C_6H_6$ . Additional evidence regarding the suggested chemical structures of the fragment ions was provided by the DBE values obtained for each ion as shown in table 4.6.

An extensive examination of the MS/MS fragmentation pattern and elemental formula has enabled the assignment of a chemical structure to a novel uremic analyte at  $m/z$  214. Using Beilstein<sup>[3]</sup> this compound is named 2-(5,6-diamino-6-diazenyl-cyclohex-1-enyl)-2-hydroxy-acetimidic acid.

#### 4.2.4.3. Candidate biomarker 3: $m/z$ 275.0478

The accurate mass generated for this novel uremic analyte suggested that it is a sulphur containing moiety of formula  $C_{13}H_{11}N_2O_3S$ . This was later confirmed by the presence of the sulphur isotope ( $^{34}S$ ) within the full scan mass spectrum resulting in a peak at 2Th above the molecular ion (see figure 4.16). The information regarding the double bond rule indicates that this is a singly protonated even electron species, containing a total of nine double bonds and aromatic rings.

Elemental Formulae of Ion	Accurate Mass of Formula	Error (ppm)	Suitable Isotope Pattern?	DBE	Measured Accurate Mass of Unknown
$C_{13}H_{11}N_2O_3S$	275.0485	-2.2	✓	9.5	275.0478

Table 4.7: Table containing the calculated accurate mass and corresponding chemical information of the unknown uremic analyte at  $m/z$  275.0478.

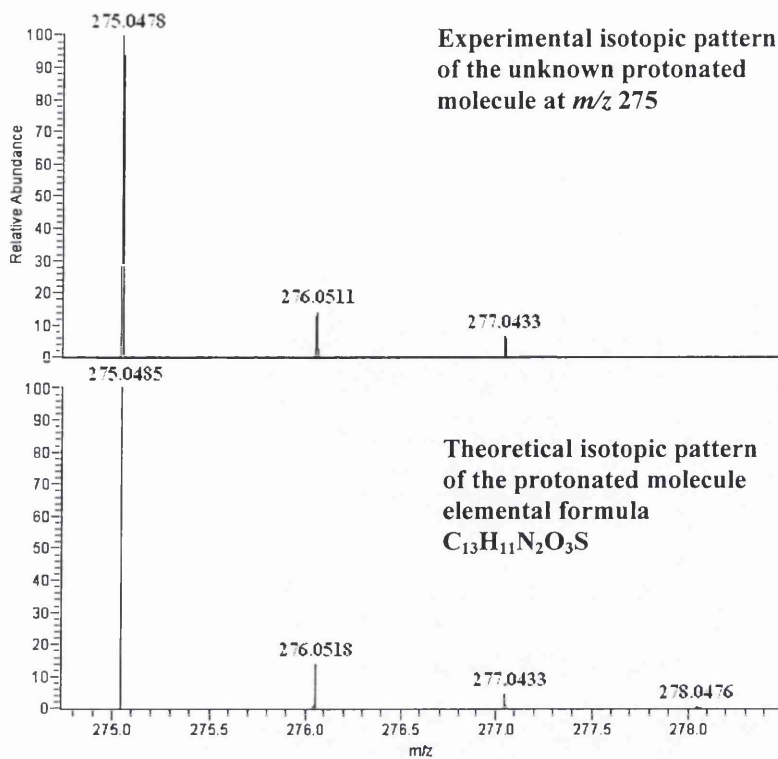


Figure 4.16: Mass spectra of the isotopic pattern of the protonated molecule of the unknown and that expected for the elemental formula  $C_{13}H_{11}N_2O_3S$ .

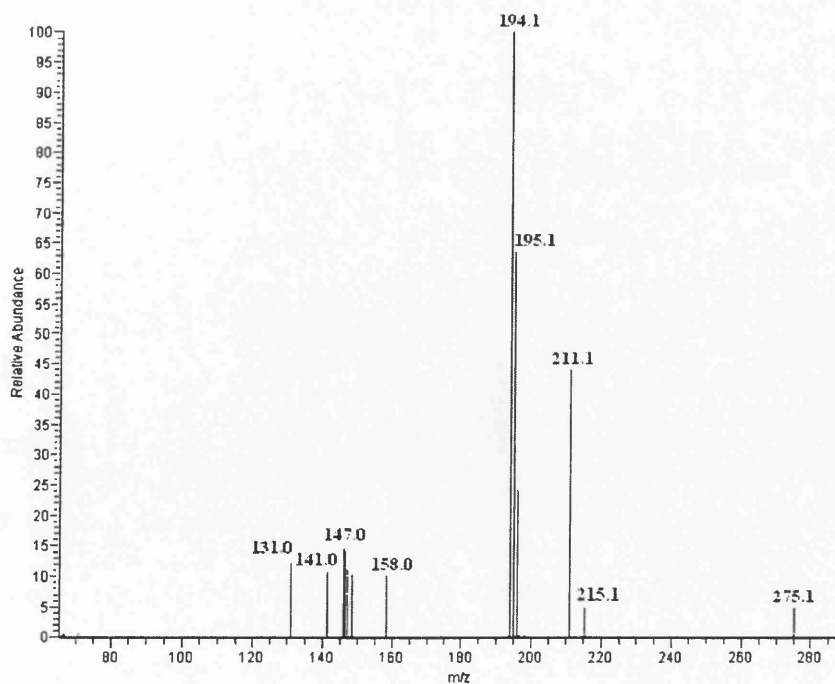
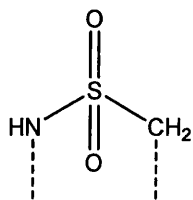


Figure 4.17: MS/MS fragmentation mass spectrum of the unknown ion at  $m/z$  275.0478 obtained at a collision energy of 45 eV.

The MS/MS fragmentation data shown in figures 4.17 and 4.19 indicate that the relevant structure must be capable of losing several distinctive functional groups. The first fragment observed at  $m/z$  215 corresponds to a loss of  $N_2O_2$  and is indicative of nitrogen oxide groups within the analyte. Another characteristic loss is associated with the ion at  $m/z$  211 and appears to be due to a loss of sulphur dioxide. The ejection of this as a neutral moiety is specific to a certain structure; the inclusion of sulphur dioxide situated next to an amino group as part of an aliphatic ring, commonly 6 or 7 membered, and can reform to 5 or 6 membered ring following the loss (see figure 4.18).



*Figure 4.18: Chemical structure of the functionality capable of losing a neutral molecule of sulphur dioxide.*

It appears as though this structure is also capable of losing a sulphonic acid group or  $SO_3H$  and according to the nitrogen rule occurs as a result of a homolytic cleavage. This with the other structural information deduced indicates that nearly all the heteroatoms other than one nitrogen and oxygen atom are located within close proximity of each other. However, the smaller mass-to-charge fragments ( $m/z$  158 and 141) also indicate that this structure has the capability of losing the nitrogen atom present within the ring either as ammonia from the ion at  $m/z$  158 or as part of larger fragment eventually leading to a loss of ammonia.

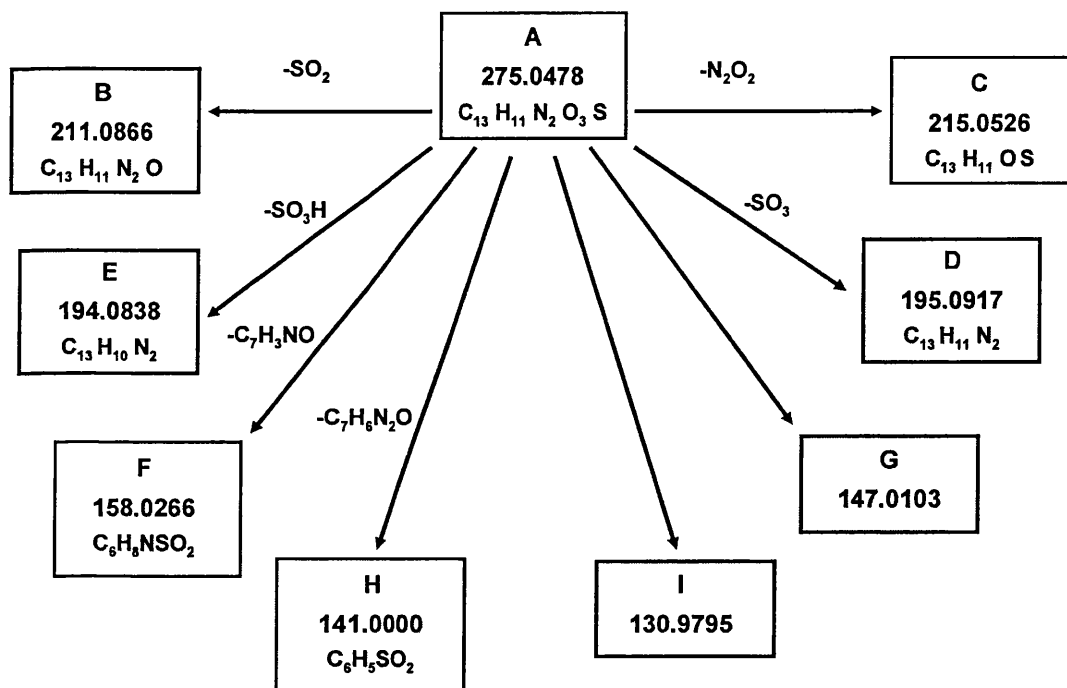


Figure 4.19: Fragmentation map of the unknown ion at  $m/z$  275.0478 showing the initial losses of nitrogen oxide and sulphur dioxide by a subtraction of 61 and 64Th respectively.

Unfortunately, under collision induced dissociation conditions the fragment ions formed have poor gas phase stability in spite of the highly stable protonated molecule, as shown by the high 45eV collision energy required. This results in limited fragmentation information being obtained and insufficient evidence for a complete structural elucidation and assignment. This is planned for future experiments and may be achieved by using other modes of fragmentation or mass analysers capable of fragmentation other than an ion trap. This instrument is associated with unexpected ‘in-trap’ reactions that could initiate further unwanted fragmentation of highly reactive product ions.



4.2.4.4. Candidate biomarker 4:  $m/z$  270.1566

The calculated accurate mass indicates that the most probable elemental formula for the ion is  $C_{11}H_{20}N_5O_3$ .

Elemental Formula of Ion	Accurate Mass of Formula	Error (ppm)	Suitable Isotope Pattern?	DBE	Measured Accurate Mass of Unknown
$C_{11}H_{20}N_5O_3$	270.1566	-0.1	✓	4.5	270.1566

Table 4.8: Elemental formula obtained for the novel uremic analyte at  $m/z$  270.

There are a number of structures that can be represented by this elemental formula; however, when compared to the predicted fragmentation pattern, the most probable is *N*-[2-(7-hydroxy-3-methyl-ocatahydro-imidazo[1,5- $\alpha$ ]pyridine-6-yl)-2-oxo-acetyl]-guanidine<sup>[3]</sup>. From a thorough literature search this appears to be a novel uremic toxin which contains chemical groups (amino and carboxylic acid) common to protein or purine/pyridine metabolism. These functional groups may also result in the short retention time observed during the reverse phase HPLC run and the high UV absorbance at 242 nm associated with its native size exclusion fraction.

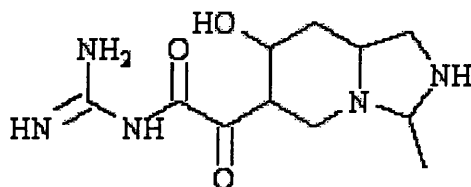


Figure 4.20: Putative neutral structure of the candidate ion  $m/z$  270.1566, named *N*-[2-(7-hydroxy-3-methyl-ocatahydro-imidazo[1,5- $\alpha$ ]pyridine-6-yl)-2-oxo-acetyl]-guanidine<sup>[3]</sup>.

This chemical structure was derived from the fragmentation of this analyte with the suggested elemental formula for each fragment. The multiple stages of fragmentation enable an extensive reaction scheme to be generated and can limit time in elucidating the chemical structure (see figure 4.20). Although, more importantly the small mass fragments, such as those at  $m/z$  97 and 120 can often be vital in ensuring a relatively quick identification. These two particular fragments are known to be characteristic of a 5-membered imadazole ring and so can be included into the eventual structure.

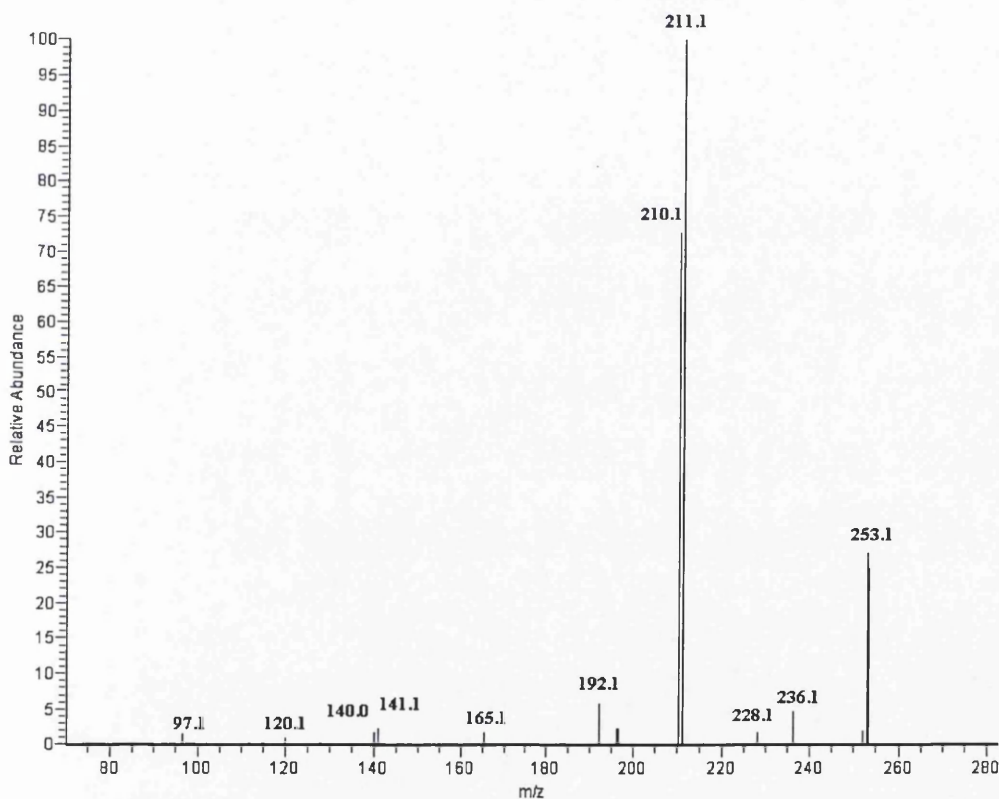


Figure 4.21: MS/MS spectra of candidate biomarker  $m/z$  270 obtained using the LCQ DECA ion trap mass spectrometer at collision energy of 35eV.

In addition to the structure of individual fragment ions, the order of fragmentation and the subsequent neutral group lost is important. This information for the protonated molecule at  $m/z$  270 indicates that the chemical structure must have a primary amine bonded to a carbon atom located next to a carbonyl group with a hydroxyl group in close proximity. This is required in order for the successive losses of ammonia, carbon monoxide, and water to occur as illustrated by the transitions of  $m/z$  228 to  $m/z$  211 to  $m/z$  183 to  $m/z$  165. This pathway is associated with only a few structures and is minimised further by considering the double bond rule. According to the suggested elemental formula from the calculated accurate mass there should be in total four double bonds plus aromatic rings in total. At least one aromatic ring must be present to generate the imadazole ring ion at  $m/z$  97 and two aliphatic double bonds for the carbonyl groups to provide losses of carbon monoxide. This information results in one permutation of these functional groups that correspond to the obtained fragmentation pattern. Elucidated reaction mechanisms of this fragmentation pathway are shown in Appendix 4.

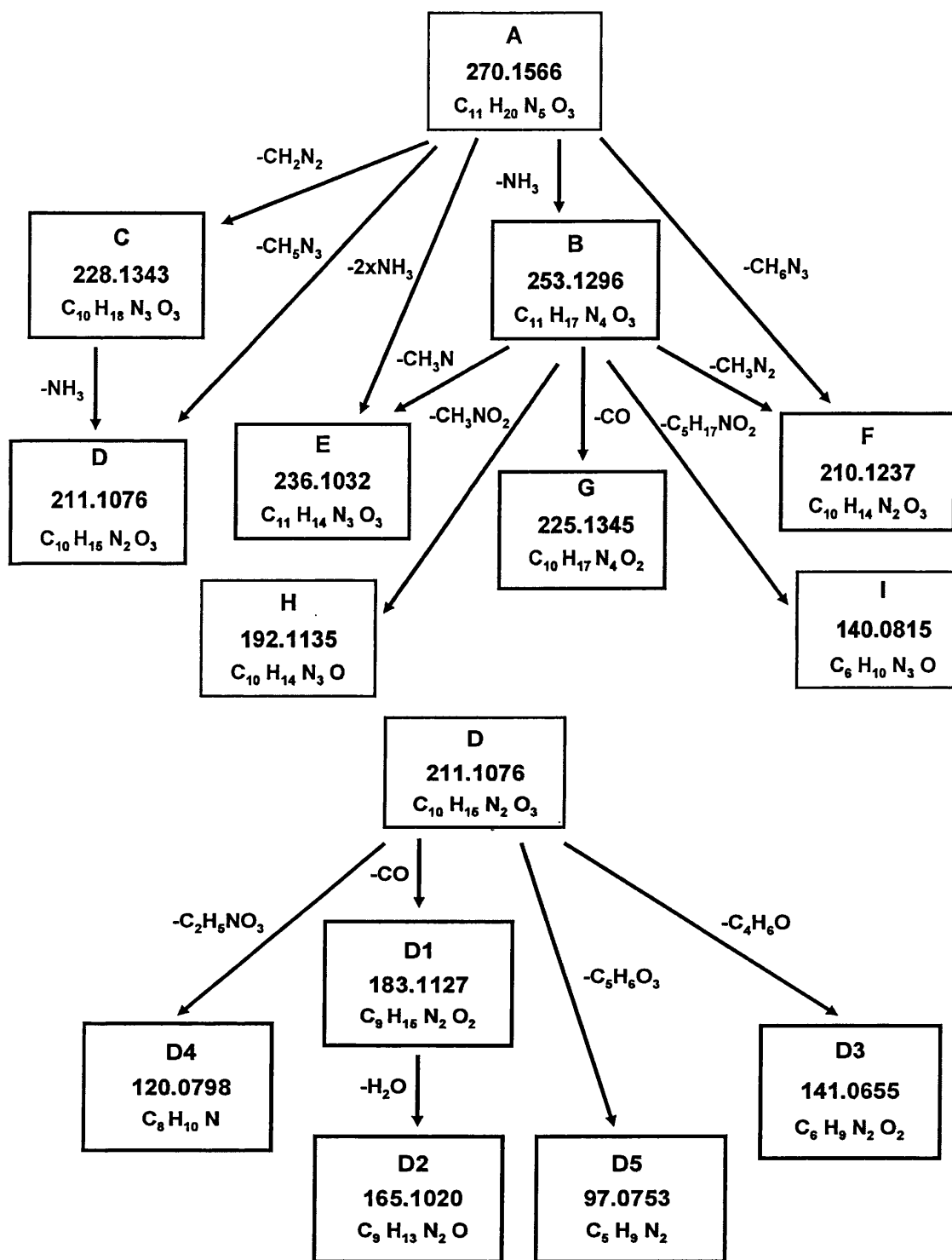
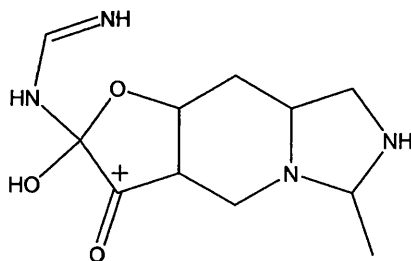


Figure 4.22: Mass fragmentation map of the novel uremic analyte *N*-[2-(7-hydroxy-3-methyl-ocatahydro-imidazo[1,5-*a*]pyridine-6-yl)-2-oxo-acetyl]-guanidine at *m/z* 270.1566.

The majority of fragments shown in the map are formed by relatively straightforward rearrangements, however the formation of two fragments at  $m/z$  225 and 210 are quite unconventional and can also provide additional evidence that the assigned structure is correct. The first fragment is produced as a result of a loss of carbon monoxide from the ion at  $m/z$  253 which is the protonated molecule minus a molecule of ammonia from its aliphatic section. Carbon monoxide would not be lost easily from this structure as both carbonyl groups are embedded within the aliphatic chain. The free rotation of the carbon-nitrogen bond at positions 2 and 3 along the aliphatic chain is required to enable the carbon to swing around to form a bond and a 5-membered ring intermediate with the second carbonyl group. Subsequent rearrangement of both the proton and double bonds places the charge onto the first carbonyl and is then ejected as a neutral molecule of carbon monoxide (see figure 4.23).



*Figure 4.23: Proposed structure of the 5-membered intermediate formed from the ion at  $m/z$  253 that is primed to eject neutral carbon monoxide resulting in the ion  $m/z$  225.*

The second unconventional ion at  $m/z$  210 is formed as a consequence of losing a highly stable fragment as shown in figure 4.25. This fragment of formula

$\text{CH}_6\text{N}_3$  is common to purine bases and is very stable due to the resonance of a double bond between the three amino groups.

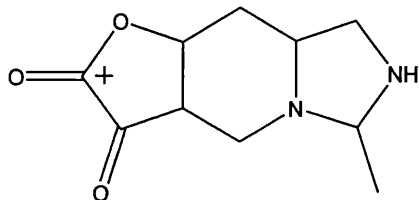


Figure 4.24: Proposed structure of the ion at  $m/z$  210 formed as a result of rotation within the aliphatic section of the molecular ion, with the formation of a 5-membered ring intermediate and the loss of  $\text{CH}_6\text{N}_3$  due to its high stability through resonance of double bonds.

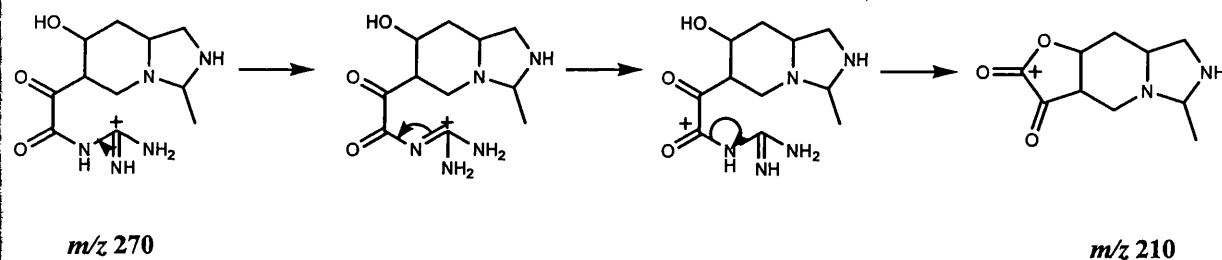


Figure 4.25: Suggested mechanism of forming the ion  $m/z$  210 through the loss of the highly stable neutral  $\text{CH}_6\text{N}_3$ .

The structures of the fragments as described above are also supported by the DBE values obtained for each ion from the calculated accurate mass (see table 4.9).

Fragment Ion ( $m/z$ )	Suggested Elemental Formula	DBE Value
253	$C_{11}H_{17}N_4O_3$	5.5
236	$C_{11}H_{14}N_3O_3$	6.5
228	$C_{10}H_{18}N_3O_3$	3.5
225	$C_{10}H_{17}N_4O_2$	4.5
211	$C_{10}H_{15}N_2O_3$	4.5
210	$C_{10}H_{14}N_2O_3$	5.0
192	$C_{10}H_{14}N_3O$	5.5
183	$C_9H_{15}N_2O_2$	3.5
165	$C_9H_{13}N_2O$	4.5
141	$C_6H_9N_2O_2$	3.5
140	$C_8H_{10}N_3O$	5.5
120	$C_8H_{10}N$	4.5
97	$C_5H_9N_2$	2.5

Table 4.9: Elemental formula information regarding the fragment ions produced from the precursor at  $m/z$  270. All accurate masses were measured to within 5ppm apart from  $m/z$  120 and 97 (-8.1 and -7.5ppm respectively).

An extensive literature search has indicated that this is a completely novel analyte to exist in nature and as the name suggests, it is a likely product of guanidine metabolism. Its presence within uremic patients provides this novel analyte with the potential to be used as a novel marker for assessing dialysis adequacy and the results of its suitability are included in the latter section of this chapter.

#### 4.2.4.5. Candidate biomarker 5: $m/z$ 180.0650 ( $m/z$ 359.1232 and $m/z$ 381.1051)

This ion was observed to co-elute with two other analytes at  $m/z$  359 and 180. It was suggested that the ion at  $m/z$  381 was a monosodiated dimer of the monomer at  $m/z$  180, and the ion at  $m/z$  359 was the singly protonated dimer. This ion at  $m/z$  180 could not be as a result of multiple charging as the mass of the proposed singly charged ion was unsuitable and there was no indication of any half mass isotopes.

This information and the calculated accurate masses suggest that the elemental formula for each moiety is  $C_{18}H_{18}N_2O_6Na$  ( $m/z$  381),  $C_{18}H_{19}N_2O_6$  ( $m/z$  359), and  $C_9H_{10}NO_3$  ( $m/z$  180) as shown in table 4.10.

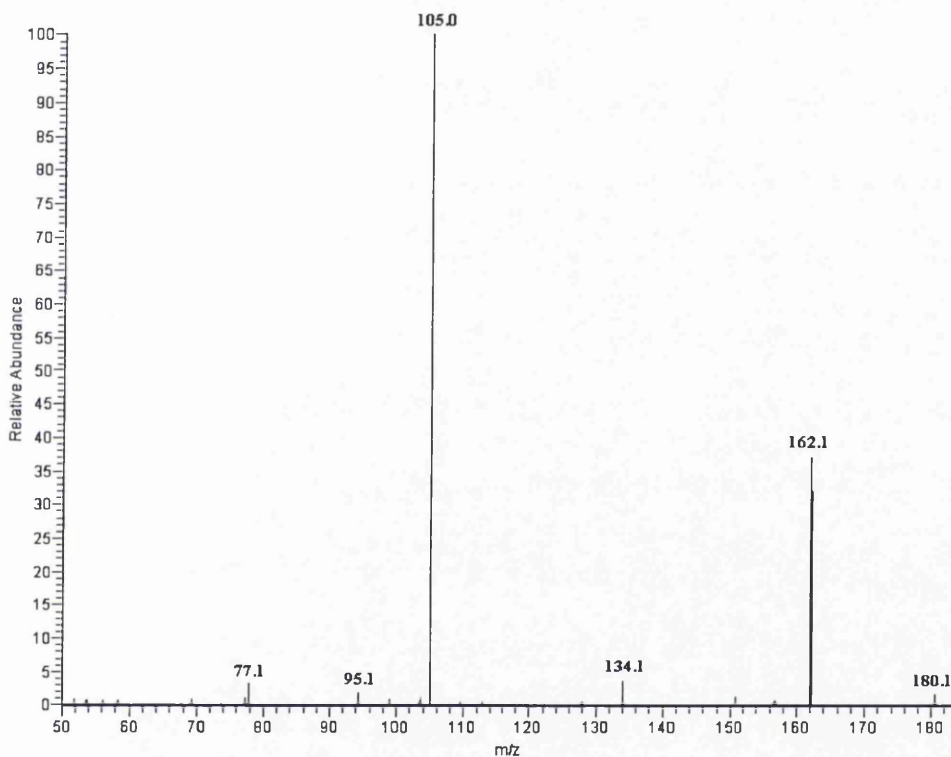
Elemental Formulae of Ion	Accurate Mass of Formula	Error (ppm)	Suitable Isotope Pattern?	DBE	Measured Accurate Mass of Unknown
$C_{18}H_{18}N_2O_6Na$	381.1057	-1.7	✓	10.5	381.1051
$C_{18}H_{19}N_2O_6$	359.1238	-1.5	✓	10.5	359.1232
$C_9H_{10}NO_3$	180.0655	-3.0	✓	5.5	180.0650

*Table 4.10: Accurate mass information and associated elemental formula for the monosodiated dimer, singly protonated dimer, and singly protonated monomer.*

The known uremic toxin Pentosidine ( $[M+H]^+$  of  $C_{17}H_{29}N_6O_4$ ) also has this mass-to-charge as a protonated molecule however the dimer affect and the elemental formula suggest that this is a novel uremic analyte. From these elemental formulae it appears that both dimers are symmetrical species with either a proton or a sodium ion present in the middle as linking group. This is represented by the very little fragmentation of these dimers and the resulting ions as shown in figure 4.27. For example, the monosodiated dimer fragments to form an ion at  $m/z$  202 only and corresponds to the monosodiated monomer. This is consistent with the protonated dimer which fragments to the protonated monomer only. Hence, any in-depth structural information regarding the position of functional groups must be deduced from the fragmentation of the monomer. Again, most information regarding structure can be obtained from the small mass fragments and those at  $m/z$  77, 95 and 105 indicate that



this analyte contains a homocyclic 6-membered aromatic ring with both hydroxyl and carbonyl groups directly attached. These relatively polar groups with a non polar aromatic functionality are further supported by the elution time of the reverse phase chromatographic separation corresponding to mobile phase conditions of 50:50 aqueous:organic.



*Figure 4.26: Fragmentation MS/MS mass spectrum at collision energy of 35eV of the singly protonated monomer at m/z 180 showing the characteristic fragments of an aromatic ring and in particular a phenoxy ring at m/z 77 and 95 respectively. The base peak of the spectrum at m/z 105 is also typical of an aromatic ring with a carbonyl group attached and is a highly stable ion accounting for its relative abundance.*

The fragmentation map also shows a double loss of carbon monoxide suggesting the presence of two carbonyl groups that can be removed along with the hydroxyl and amino groups to leave the single phenyl ring. This information and the order of

fragmentation implies that this structure is an aromatic ring with hydroxyl group and an aliphatic chain containing two carbonyl and an amino group attached.

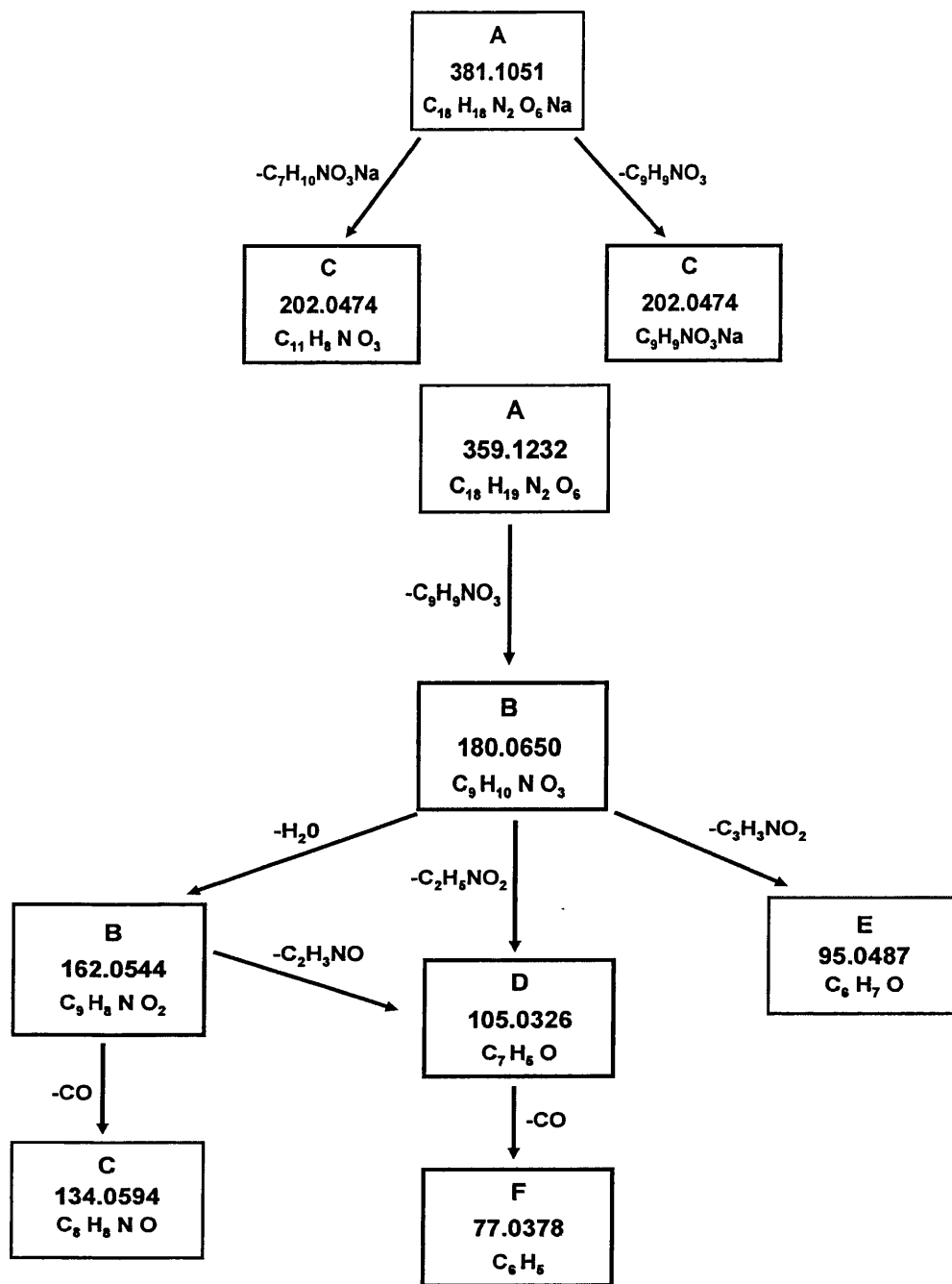
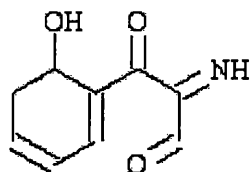


Figure 4.27: Fragmentation map of the monoprotonated dimer; initially showing the formation of the monomer and subsequent monomer fragments.

One permutation of these functional groups that matches the fragmentation pattern completely has the name 3-(6-hydroxy-cyclohexa-1,3-dienyl)-2-imino-3-oxopropionaldehyde<sup>[3]</sup>. This structure is consistent with the chromophoric groups required to generate a high absorbance signal at 242 nm associated with the size exclusion fraction analysed.



*Figure 4.28: Proposed neutral structure of  $m/z$  180 and corresponds to the name 3-(6-hydroxy-cyclohexa-1,3-dienyl)-2-imino-3-oxopropionaldehyde.*

The mechanisms involved in the fragmentation of the monomer consist of several common rearrangement reactions all reliant on the resonance of charge within the structure as shown in Appendix 4 section 4.5. The DBE value implies that the structure of the ion should have a total of five double bonds plus ring structures. If the DBE number is determined from the neutral proposed structure, then it should have a value of six. However, when charged in solution this DBE value can provide some evidence of the native structure. For example, the structure must lose a double bond through the resonance effect, and the charge to residing on the aliphatic ring as shown in the proposed solution structure (see figure 4.29).

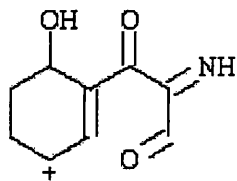


Figure 4.29: Structure of 3-(6-hydroxy-cyclohexa-1,3-dienyl)-2-imino-3-oxopropionaldehyde when charged in solution.

The application of the DBE calculation to the individual fragments obtained for this monomer analyte, as shown in Appendix 4.5, provides additional evidence to support these structures and the overall structure of the monomer (see table 4.11).

Fragment Ion ( $m/z$ )	Suggested Elemental Formula	DBE Value
162	$C_9H_8NO_2$	6.5
134	$C_8H_8NO$	5.5
105	$C_7H_5O$	5.5
95	$C_6H_7O$	3.5
77	$C_6H_5$	4.5

Table 4.11: Elemental formula information obtained from the calculated accurate mass of each fragment. All accurate masses were measured within an error of 5ppm apart from  $m/z$  105 and 77 which were acquired within 10ppm.

A thorough literature search has shown that this is not only a novel uremic analyte but also novel regarding its existence in biological systems. However, it does have structural similarities to an intermediary metabolite of benzoylglycine. Hence, it is possible that this structure is formed as a result of incubation of benzoylglycine with other uremic analytes at relatively high levels. The suitability of this analyte in monitoring dialysis adequacy in comparison to the current markers, urea and

creatinine, has been assessed in patient specific samples and the results are discussed in the latter part of this chapter.

#### 4.2.4.6. Candidate biomarker 6: $m/z$ 335.0545

This ion was observed in size exclusion fraction 4 (as expected for this mass-to-charge) with a calculated accurate mass of 335.0545 (see table 4.12). This accurate mass indicates that this analyte has the elemental formula of  $C_{11}H_{15}N_2O_8S$  and is supported further by a corresponding isotopic pattern as shown by figure 4.30.

Elemental Formula of Ion	Accurate Mass of Formula	Error (ppm)	Suitable Isotope Pattern?	DBE	Measured Accurate Mass of Unknown
$C_{11}H_{15}N_2O_8S$	335.0544	-0.3	✓	5.5	335.0545

Table 4.12: Accurate mass data and suggested elemental composition of the candidate ion  $m/z$  335.0545.

There are no published uremic analytes that would be observed at  $m/z$  335 and indicates that this is a novel uremic analyte. This putative elemental formula information was searched in published literature for a corresponding chemical structure, although, this proved unsuccessful. This does not render the ion unsuitable in assessing dialysis adequacy as it may function as a diagnostic marker regarding patient health.

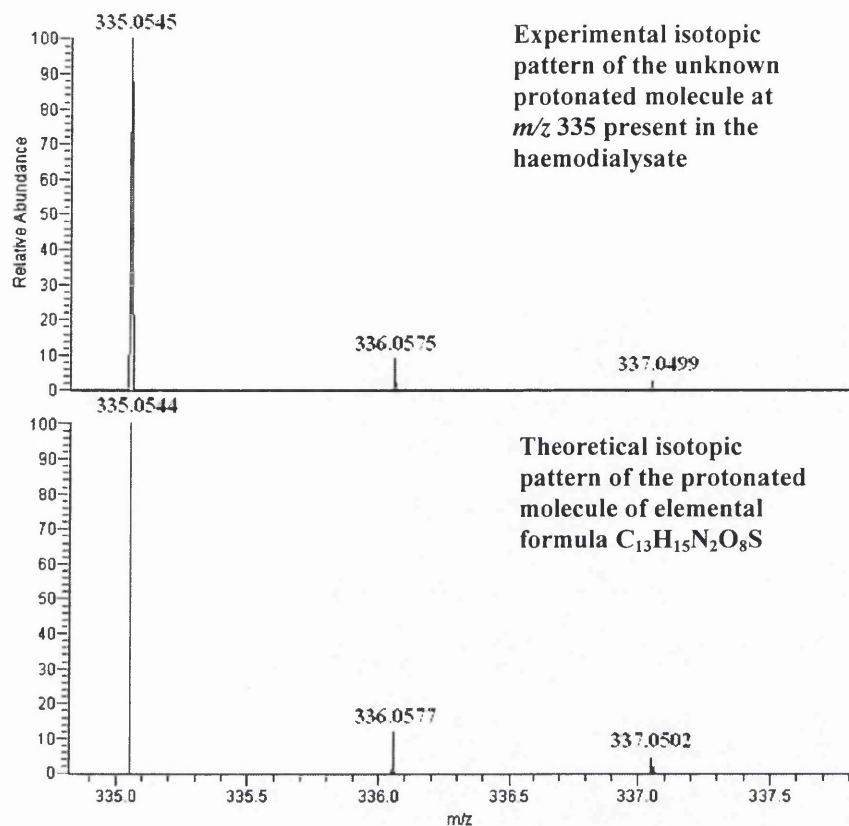


Figure 4.30: Mass spectra of the isotopic patterns of the unknown ion at  $m/z$  335 observed within the haemodialysate solution and the molecular ion of formula  $C_{13}H_{15}N_2O_8S$ .

The fragmentation MS/MS mass spectrum shows a characteristic double water loss from the protonated parent molecule. This is indicative of the presence of two hydroxyl groups attached to carbon atoms and is supported by the calculated accurate mass and associated elemental formula of the relevant fragment ions (see figure 4.32). Unfortunately, limited elemental formula information was acquired for the product ions shown in the MS/MS mass spectrum and results in an increased difficulty in assigning a chemical structure to the analyte. However, the product ion data can be examined for any common neutral losses and relevant functional groups proposed for the chemical structure. For example, the ionic species at  $m/z$  299 of formula

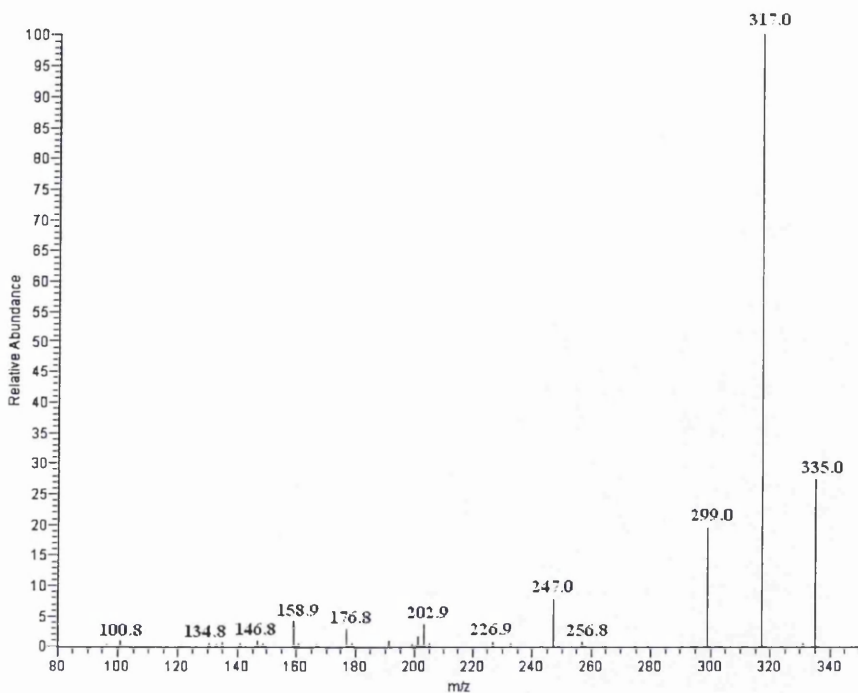


Figure 4.31: MS/MS fragmentation mass spectrum of the candidate ion at  $m/z$  335 using collision energy of 35eV and showing comparable product ions of  $m/z$  317, 299, 247, 227, 203, 177, 159 and 147.

$C_{11}H_{11}N_2O_6S$  is capable of forming two ions at  $m/z$  219 and 201 corresponding to losses of 80 and 98 Da respectively. From the starting elemental formula these ions could be generated due to the neutral loss of  $SO_3$  (-80 Da) and sulphuric acid,  $H_2SO_4$  (-98 Da). The fragmentation scheme (see figure 4.32) also shows losses of 42 and 72 Da from the ion at  $m/z$  299. These could be as a result of losing groups such as  $C_2H_2O$  and  $C_3H_8O_4$  respectively, which are common to precursor ions containing a sugar residue. This is further supported by the presence of the ion at  $m/z$  203 in the MS/MS mass spectrum which could correspond to a loss of  $C_5H_8O_4$  or a ribose sugar moiety. If this premise and an  $SO_3$  or sulphuric acid loss is correct then the remaining product ion will have a formula of  $C_6H_6NO_2$  which is indicative of the biomolecule nicotinamide.

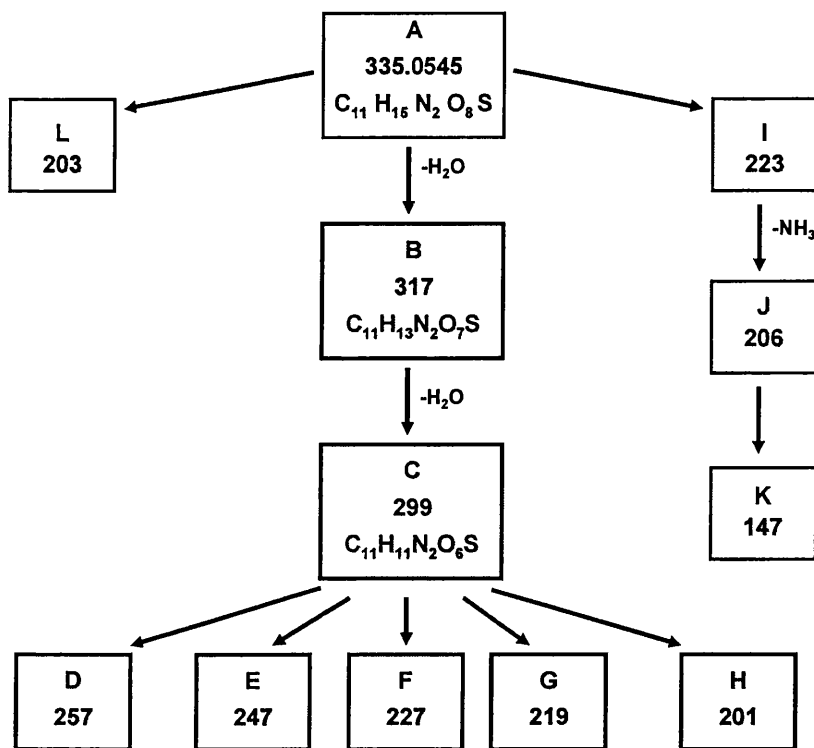


Figure 4.32: Product ion reaction scheme of the novel uremic analyte at  $m/z$  335.

This has an important biochemical role within the body as part of a larger molecule of nicotinamide adenine dinucleotide phosphate or NADP. Therefore, a possible structure of this unknown analyte could be the ribonicotinamide moiety bonded to a sulphate rather than the conventional phosphate group. However, unlike structure of NADP, the nicotinamide and ribose units for the unknown must be joined differently. This is required for the loss of the ribose sugar to occur and the nicotinamide and sulphate sections to be retained. The losses of sulphuric acid and possible ammonia (shown by the transition of  $m/z$  223 to  $m/z$  206) can also provide information about the chemical structure. For example, the  $\text{SO}_3$  group must be bonded to the oxygen atom of the nicotinamide structure and the nitrogen atoms are not included in linking the individual molecular moieties. Hence, these inferences suggest that a possible



structure of the unknown ion at  $m/z$  335 is as shown in figure 4.33 and the reaction mechanisms for the relevant transitions are shown in Appendix 4.5.

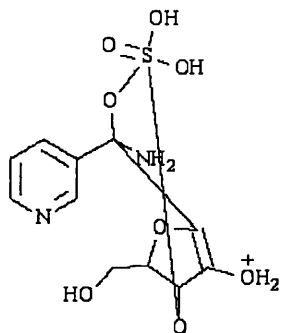


Figure 4.33: Possible structure of the unknown ion at  $m/z$  335.

However, for this unconventional biomolecular structure to be confirmed further fragmentation studies with product ion accurate mass and elemental formula are required.

#### **4.3. Excretion Kinetics of Urea and Creatinine with Novel Uremic Analytes in Patient Specific Samples**

Haemodialysate samples were obtained from two patient cohorts at time points 0, 0.5, 1, 1.5, 2.5 hours and at end of session to provide sufficient coverage of analyte excretion in the minimum number of samples. The samples collected were intended as a pilot study to monitor analyte excretion in patients receiving the conventional dialysis treatment and those who exercised for an hour in addition to this treatment. Past research has shown that exercise can increase the efficiency of dialytic removal

for urea and creatinine and is equivalent to extending the treatment time by 15-20 minutes<sup>[4, 5]</sup>. We hypothesise that exercise will improve the initial rate of analyte excretion and the total amount during the session. Hence, any compartment excretory behaviour, such as drop in the amount excreted from emptying the blood (first compartment) would be observed at an earlier stage of the dialysis treatment. In addition to the conventional biomarkers of urea and creatinine, the novel uremic analytes identified were monitored within the haemodialysate solution. Therefore, we characterised the excretory behaviour of these novel analytes within both patient sets and evaluated their performance against urea and creatinine as indicators of dialysis adequacy and solute removal. Urea and creatinine levels were determined using enzymatic tests (F. Hoffmann-La Roche, Switzerland) by Morrison Hospital Swansea NHS Trust Pathology department. In order for these results to be comparative we have normalised the data acquired for each patient and expressed it as a percentage of the maximum excretion point within the dialysis session. Each of the time point results were averaged separately creating two patient groups (control who received dialysis only and exercise, those who exercised prior to treatment) to compare. These mean time point results were then normalised to the maximum point of excretion observed within this data set and again expressed as a percentage. These data can now be represented as an excretory profile comparing the mean results for each patient cohort (see figure 4.34). The standard error of the mean patient data sets were obtained from repeated time point measurements and then multiplied by a scaling factor to account for the last stage of data normalisation. These are included on each excretory profile to illustrate the range of results acquired for each time point. Further information regarding this statistical process is shown in Appendix 4 for urea.

### 4.3.1. Excretion during haemodialysis

#### 4.3.1.1. Urea

Urea measurements were obtained within the range of 0.5-62.9mmol/L with the highest concentration present in the 0 hour time point sample regardless of the patient cohort. Generally for both control and exercise patients' urea is excreted to the highest degree during the initial stages of haemodialysis. However, for control patients urea excretion is sudden at the 0 hour time point which then plateaus from 0.5 hour to 1 hour and is then repeated for the remainder of the session (see figure 4.34). This is indicative of compartment excretion kinetics in which the circulatory system (first compartment) is emptied of analytes that are replaced from the tissues (second compartment) and then removed again. This effect is not as pronounced for those patients who have exercised prior to starting dialysis treatment. Figure 4.34 shows that this patient cohort has a lower starting urea concentration which decreases relatively rapidly until 1 hour and then gradually removed until the end of the session.

Urea is a common degradative product often formed at the end of metabolism. This information could indicate that exercise causes a decrease in the activity of metabolic pathways forming this uremic analyte due to the recycling of the starting material into a beneficial biological moiety. In addition, exercise seems to be associated with a greater level of variation in data obtained at each time point, when compared to control patients. This may be related to patient exertion during the exercise regimen and could suggest that more stringent requirements such as target heart rate should be included in future clinical trials.

Mean excretion of urea shown as a percentage of the largest concentration excreted at a timepoint throughout the hemodialysis session

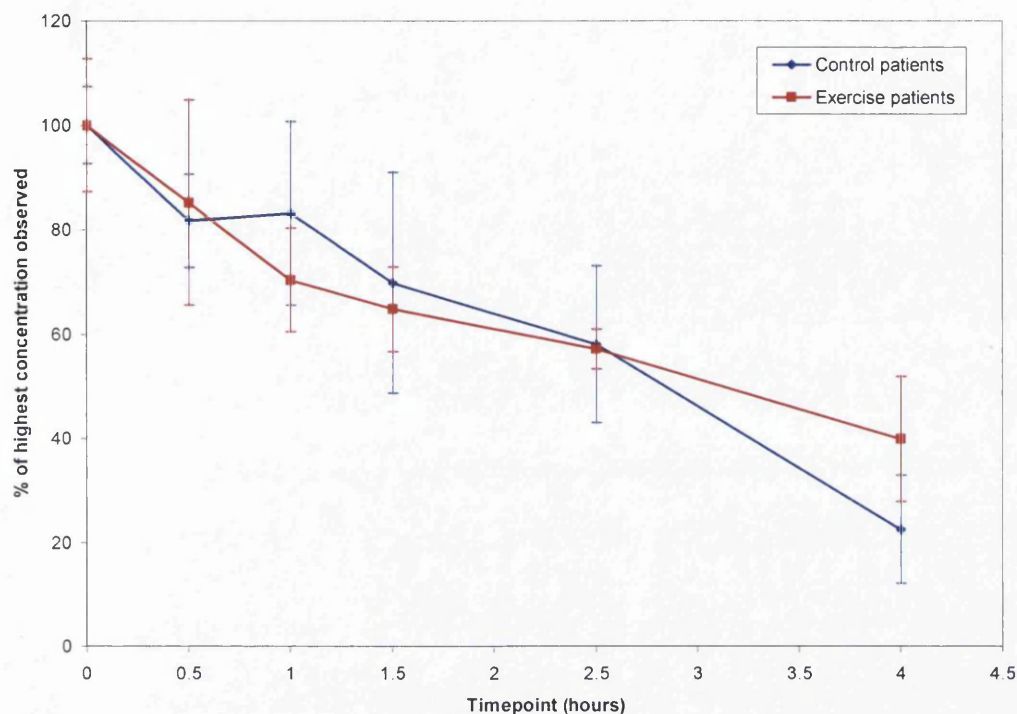


Figure 4.34: Excretory profile of the uremic toxin urea for patients receiving haemodialysis treatment only (control,  $n = 5$  patients  $\times$  2 measurements) and haemodialysis patients who exercised prior to initiating treatment ( $n = 3$  patients  $\times$  2 measurements).

In summary it appears that the highly polar urea analyte is sufficiently removed during the haemodialysis treatment for both control and exercise patients. This is shown by the difference in percentage excretion at the start of dialysis to the end of the session which, correspond to concentrations considered to be within a 'normal' range.

#### 4.3.1.2. Creatinine

Creatinine, like urea, is excreted to the greatest extent during the initial stages of haemodialysis treatment. Throughout the dialysis session creatinine is excreted at levels of 18-39 $\mu$ mol/L and the highest concentration is detected at 0 hour and the lowest 5 minutes prior to disconnection. Again, like urea, creatinine has slightly different excretory profiles for the control patients and those who have exercised before commencing haemodialysis treatment. Excretion in control patients shows subtle compartment kinetics; rapid initial removal (0 hour), reaching equilibrium or constant removal (0.5 to 1 hour), followed by a decrease in excretion (1 to 1.5 hours). This constant removal and subsequent decrease is then repeated until the end of the dialysis session (see figure 4.35). Patients who exercise prior to treatment have an altered excretory profile for creatinine. Initial losses of creatinine appear to be lower than control patients and overall show a more gradual removal throughout the dialysis session. A significant difference between both patient sets is that a lower level of creatinine is present in the control samples at the end of dialysis than those who have exercised. Creatinine serum levels can be affected by several factors, ranging from muscle mass to quantity of protein consumed in the patients' diet. However, creatinine clearance is understood not to suffer from these effects. Therefore it is unlikely that this elevated concentration at the end point is caused by the addition of muscle from the exercise regime as this would cause a decrease in creatinine clearance. A possible explanation for this rise in creatinine excretion could be an increase in protein metabolism, generating greater quantities of creatinine within the blood for removal.

Mean excretion of creatinine shown as a percentage of the largest concentration excreted at a timepoint throughout the hemodialysis session

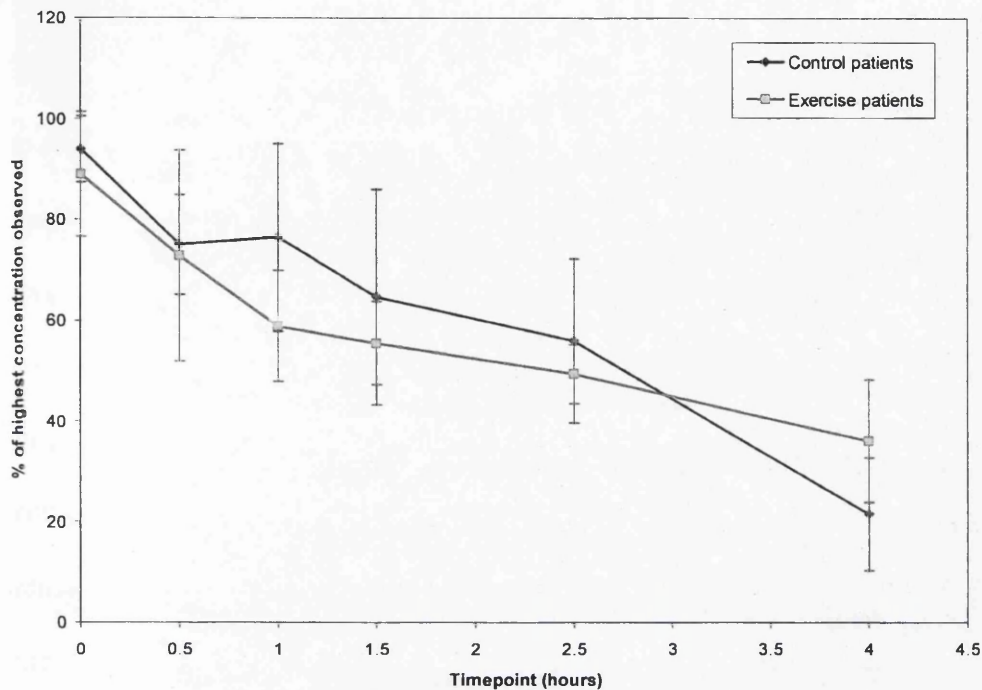


Figure 4.35: Excretory profile of the uremic toxin creatinine for patients receiving haemodialysis treatment only (control,  $n = 10$ ) and haemodialysis patients who exercised prior to initiating treatment ( $n = 6$ ).

To summarise, both excretory profiles for creatinine are highly similar, each illustrating a good level of removal for this polar analyte represented by the low end points observed. According to these data points creatinine will be removed to the extent of 9% or 24% of the maximum excreted in control and exercise patients by the end of the treatment session, all of which correspond to concentrations considered within 'normal' range.

4.3.1.3. Candidate biomarker 1: 5-(amino-1,2,-dihydroxy-ethyl)-3-nitrosooxy-[1,2,4]trioxine-3,6-diol (*m/z* 241.0296)

This polar analyte (RP-HPLC relative retention time of 0.52 minutes) unlike urea and creatinine has the highest point of excretion at 0.5 hours in control patients. There are subtle differences observed for the excretion of this analyte between these and exercise patients and all profiles were obtained with good observed reproducibility (see figure 4.36). The excretory profile of control patients shows a gradual decrease in percentage excretion following the profile maxima and an end point value that is comparable to the start of treatment. This could indicate that some of the analyte is removed by dialysis but not with great efficiency. The incorporation of exercise into the treatment regime appears to improve the effectiveness of dialysis in extracting this solute from the patient. Following the maxima at 0 hours this profile rapidly decreases from 0.5 to 1.5 hours at which it peaks again at 2.5 hours and decreases until the end of session. However, unlike control patients, the end of session percentage excretion is considerably lower and observed at no greater than 50% of the profile maximum. This could suggest that exercise increases the rate at which this solute is removed during dialysis or perhaps the passage of analyte from the tissues to the circulatory system ready for removal.

Mean excretion of uremic analyte at m/z 241 shown as a normalised percentage of the largest peak area obtained at each timepoint throughout the hemodialysis session

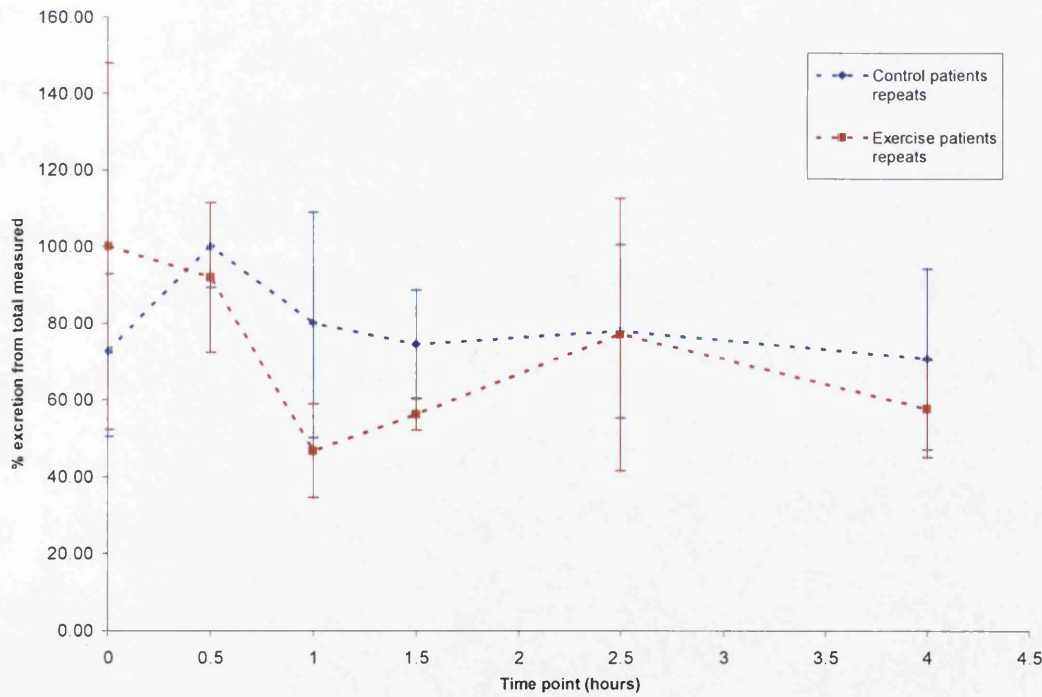
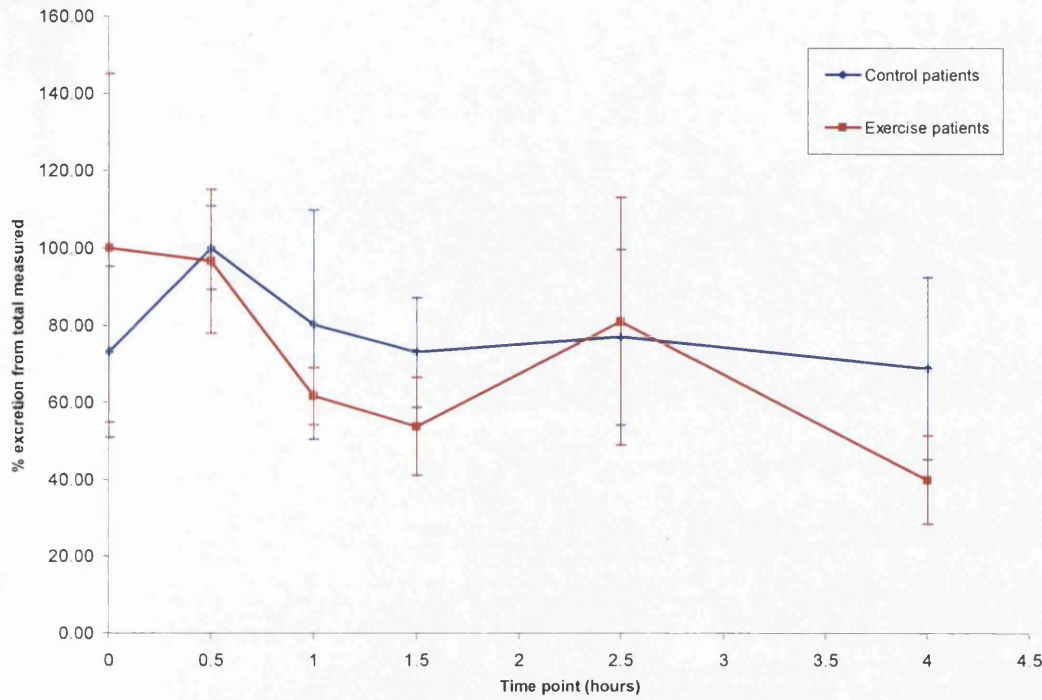


Figure 4.36: Excretory profiles of the novel uremic analyte observed at m/z 241 obtained for control and exercise patients (n = 10 and 6 respectively) including repeat analysis profile of the same sample.



Also, if this analyte has a role in metabolism, exercise could alter the amount formed through the degree of exertion by the patient. This latter suggestion is supported by the larger degree of variation observed in the data for exercise patients and could be related to differences in patient exertion during the regimen. As stated previously, this would be confirmed or refuted by imposing more stringent guidelines to the exercise regime for future studies. However, in order to confirm the exact cause of variation in excretion absolute levels need to be quantitated and would require the synthesis of the reference material for an internal standard.

4.3.1.4. Candidate biomarker 2: 2-(5,6-diamino-6-diazenyl-cyclohex-1-enyl)-2-hydroxy-acetimidic acid ( $m/z$  214.1298)

This polar analyte shows a good level of removal during the haemodialysis procedure regardless of patient set, illustrated by the difference in percentage excretion observed when treatment commenced and finished. The profiles were obtained with good reproducibility for both control and exercise patients and show elements of compartment excretion achieving an apparent high degree of removal, although there are some slight variations. The profile for control patients could suggest that a large amount of analyte is removed from the circulatory system but gradually replenished by the tissues at a relatively slower rate than which it is removed by the dialysis treatment (see figure 4.37). In order for this to be confirmed a standard reference material would need to be synthesised to quantitate the levels removed during the session. Exercising before receiving the same treatment causes the initial excretion level to increase and drop to 70% excretion at 1.5 hours. This is followed by a subtle rise in excretion possibly due to the replacement of analyte from the tissue

Mean excretion of uremic analyte at m/z 214 shown as a normalised percentage of the largest peak area obtained at each timepoint throughout the hemodialysis session

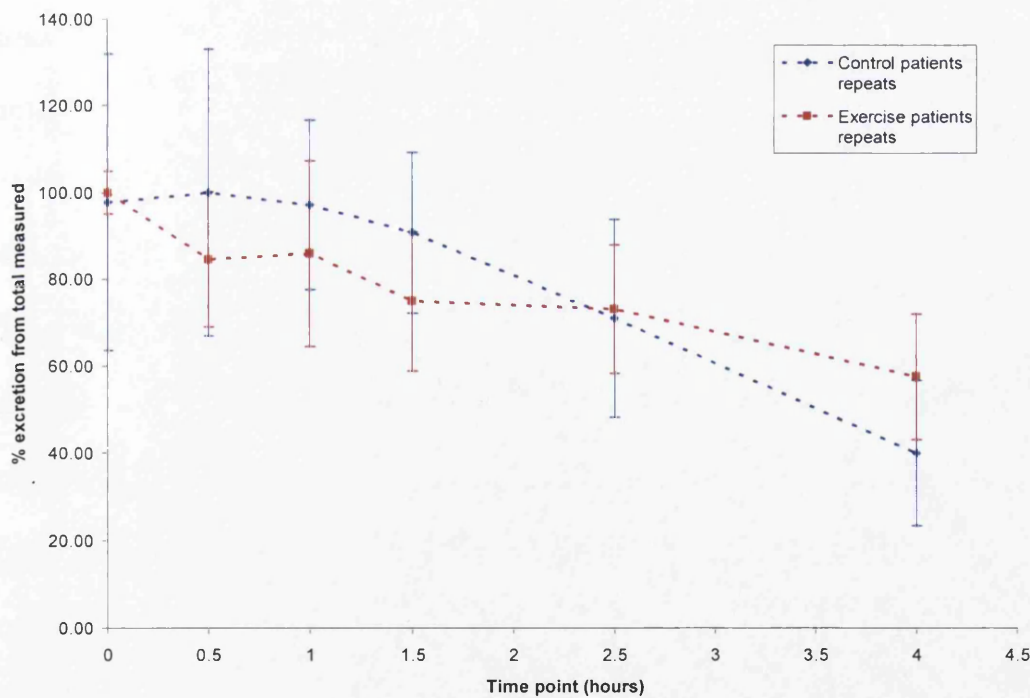
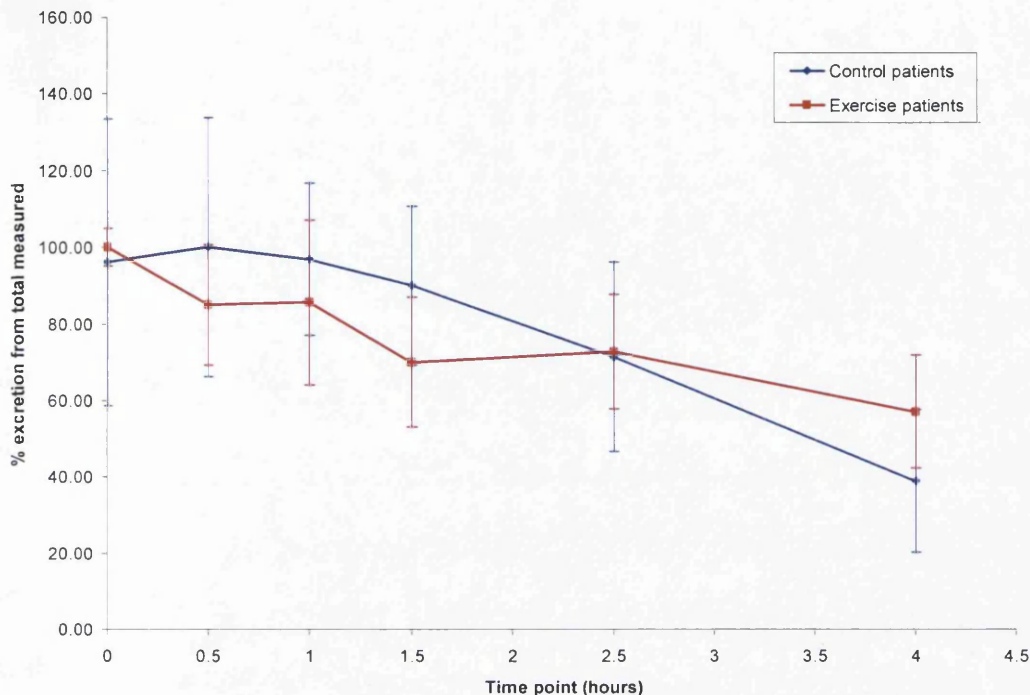


Figure 4.37: Excretory profiles of the novel uremic analyte observed at m/z 214 for control patients and those who have exercised prior to commencing treatment (n = 10 and 6 respectively), including profiles obtained from repeat analysis.

compartment into the blood stream, and then decreases until the end of the session (see figure 4.37). Another difference with respect to the two patient cohorts is the excretion level at the end of dialysis. Exercise patients appear to have a slightly higher degree of removal and could indicate that this has increased metabolism and the level present in the blood, creating a larger concentration gradient. The effects of exercise are also apparent by the degree of variation within the results. Again, a greater level of variation is observed when the patient exercises and may be related to the amount of effort exerted by the patient prior to commencing treatment.

The profiles of both patient sets show good removal of analyte indicated by the relatively large differences observed between the time of maximum excretion and the end point of treatment. Therefore, we can suggest that overall the conventional markers, urea and creatinine, would be representative of the excretion of this polar solute.

#### 4.3.1.5. Candidate biomarker 3: $m/z$ 275.0478

This relatively non-polar analyte with a RP-HPLC relative retention time of 35.6 minutes has a very different excretory profile obtained with a good degree of reproducibility when compared to urea and creatinine, regardless of the patient cohort (see figure 4.38). The profiles for both control and exercise patients show a poor level of removal clearly illustrated by the high degree of excretion associated with the end of session samples. These could indicate that for similar non-polar analytes dialysis needs to be carried out for much greater lengths of time for sufficient removal. Exercise appears to have a considerable impact on the excretion of this solute shown by a large peak at 1 hour. Levels of excretion seem to remain relatively constant until 2.5 hours from which it increases gradually until the end of the session.

Mean excretion of uremic analyte at  $m/z$  275 shown as a normalised percentage of the largest peak area obtained at each timepoint throughout the hemodialysis session

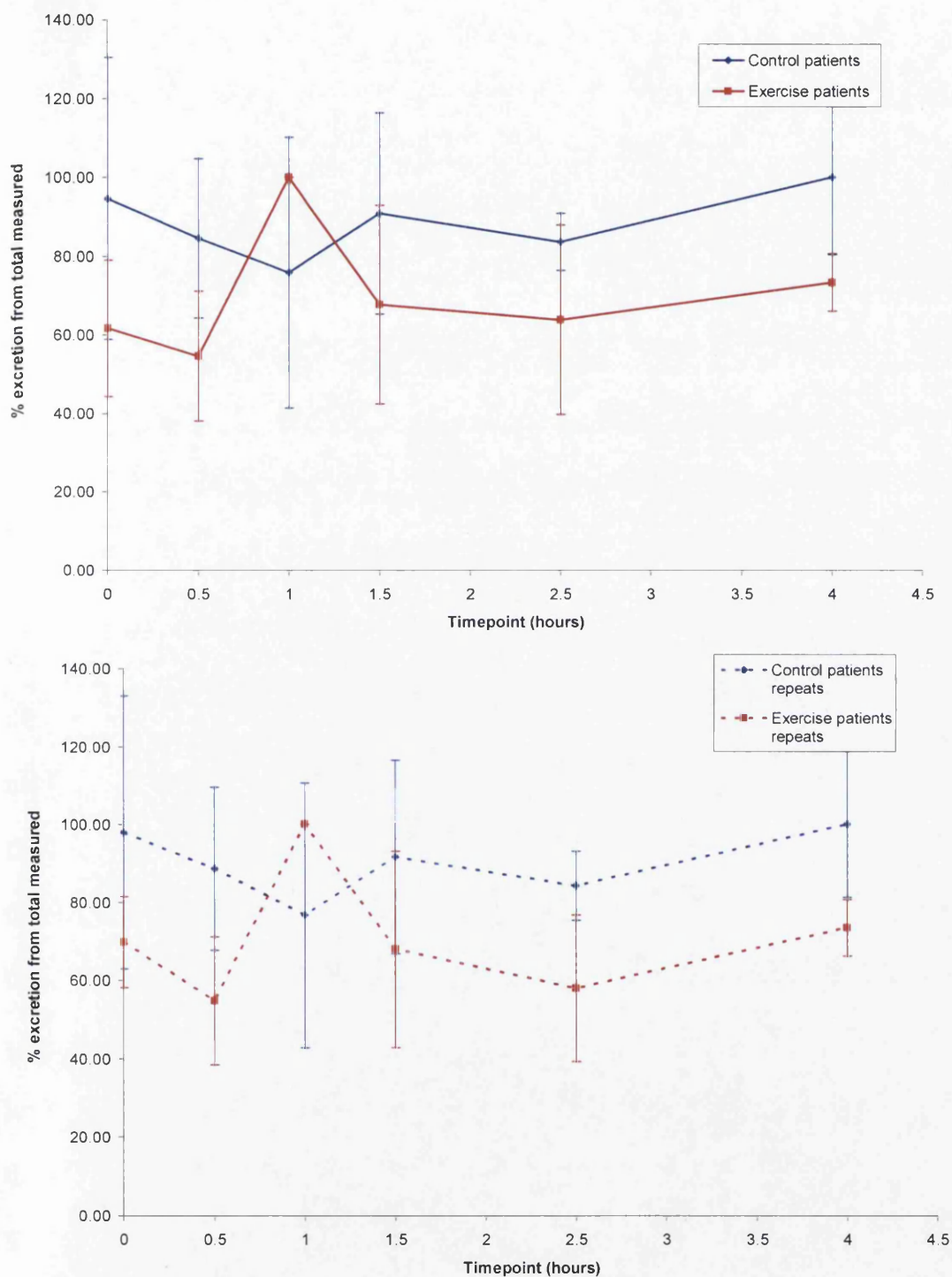


Figure 4.38: Excretory profiles of the non-polar novel uremic analyte observed within the haemodialysate at a relative RP-HPLC retention time of 35.6 minutes and at  $m/z$  275. This figure includes both first pass and repeat analysis data showing comparable profiles ( $n = 10$  and 6 for control and exercise patients respectively).

This could indicate the level of analyte within the circulatory system was increased by exercise and removed from 0.5-1.5 hours, with a slow rate of transport into the blood from the tissues. The high degree of variation in excretion for this data set could also suggest that patient exertion during the exercise regime should be monitored and maintained more closely.

In summary, this non-polar analyte appears to have very poor removal in patients undergoing conventional dialysis treatment, and is illustrated by the maximum percentage excretion observed at the end of the session. This alone provides sufficient evidence that the current biomarkers, urea and creatinine, are not representative of all uremic solutes in assessing dialysis adequacy. The behaviour of this analyte may offer an insight of how haemodialysis may be developed to increase efficiency of removal. For example, exercise does appear to improve removal with the greatest level of excretion at 1 hour into the session. This could be due to a number of reasons including an increase in the level of analyte in the circulatory or first excretory compartment, ready for removal. However, high levels of excretion at the end of session also imply that exercise alone can not result in sufficient removal of this non-polar analyte. An extraction reliant on osmotic diffusion of non-polar analytes into an immiscible aqueous solution of dialysate is unfavourable due to its low solubility level. This could result in the poor degree of removal displayed by these results and hence, improvements to the haemodialysate technique by enhancing this solubility should be investigated.

4.3.1.6. Candidate biomarker 4: *N*-[2-(7-hydroxy-3-methyl-oxo- $\alpha$ -hydro-imidazo[1,5- $\alpha$ ]pyridine-6-yl)-2-oxo-acetyl]-guanidine (*m/z* 270.1566)

This polar analyte shows a good and reproducible profile of removal in both control and exercise patients. Excretion in control patients shows the maximum at 0.5 hours which gradually decreases until the end of the treatment session. The profile indicates that a high proportion of this analyte has been removed by the procedure with the end of session sample showing no more than 13% excretion of the maximum value. The inclusion of exercise prior to dialysis appears to cause a greater amount of initial excretion when compared to controls, and decreases rapidly at 1 hour. The subsequent time point is a near equivalent to the excretion maximum for the profile resulting in a sudden increase in removal followed by a gradual decrease until the end of session. However, unlike the control patients the level of excretion is much higher at the end point sample perhaps indicating that the level of this analyte has been increased as a product of metabolism.

This solute appears to have a relatively good removal by haemodialysis and exercise may cause an increase in the level excreted at the start of treatment. However, the overall higher levels of this analyte observed in this patient set could indicate that exercise actually increases the amount present in the blood, implying that it could be a product of a metabolic pathway that increases with oxygen intake. This is consistent with the identity assigned and implies that it is related to purine or pyrimidine metabolism. The good removal and its similarities to urea and creatinine excretion suggest that these biomarkers could be used to monitor the excretion of this novel polar analyte at *m/z* 270.

Mean excretion of uremic analyte at m/z 270 shown as a normalised percentage of the largest peak area obtained at each timepoint throughout the hemodialysis session

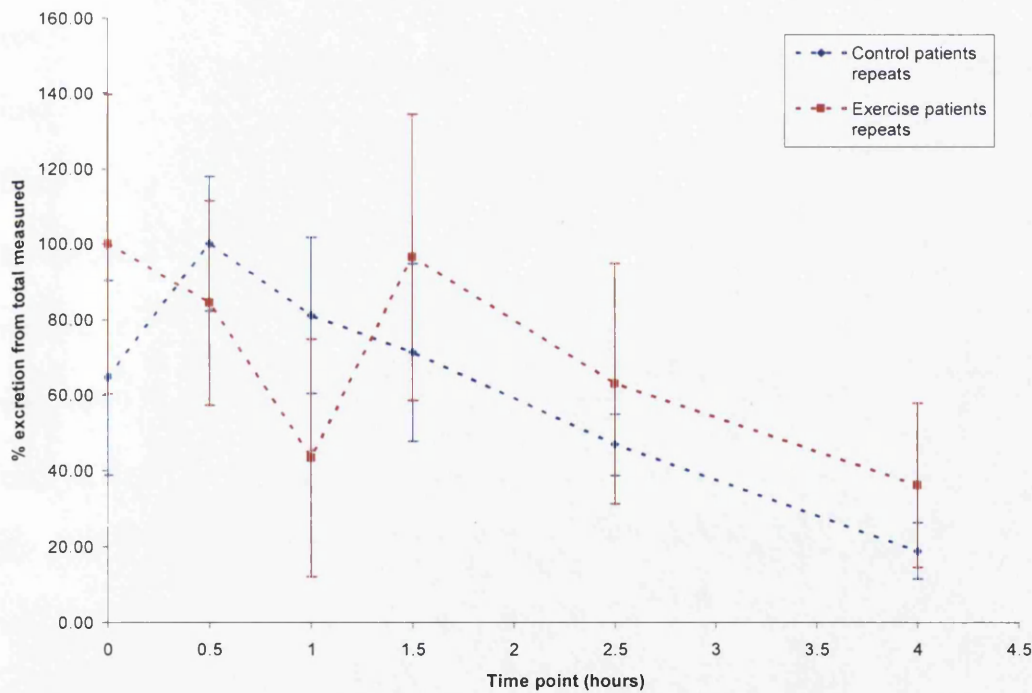
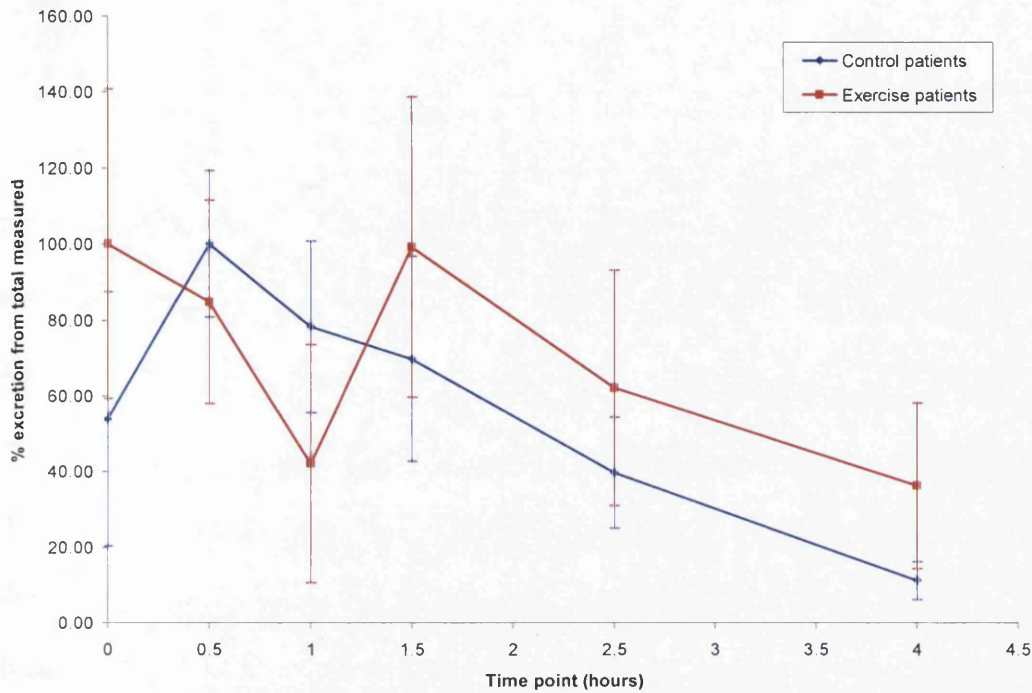


Figure 4.39: Excretory profiles of the novel uremic analyte at m/z 270 obtained during an initial and repeated analytical runs for control and exercise patients ( $n = 10$  and  $6$  respectively).

4.3.1.7. Candidate biomarker 5: 3-(6-hydroxy-cyclohexa-1,3-dienyl)-2-imino-3-oxopropionaldehyde ( $m/z$  180.0650 [ $m/z$  359.1232 and  $m/z$  381.1051])

This analyte ion can exist as its native protonated monomer, a protonated dimer or as a monosodiated dimer and the latter two may form as a result of the electrospray ionisation process. The relatively non-polar monomer analyte is observed within the RP-HPLC run at a relative retention time of 24.3 minutes. The individual ions have almost identical excretory profiles that are highly reproducible and were combined for comparison to urea and creatinine as one uremic solute. In control patients they appear to be removed poorly until 1.5 hours into the dialysis session, and then decreases rapidly until the end of treatment. This could imply that there is rate limiting step at the beginning of dialysis for these analytes. A possible explanation could be due to compatibility or solubility issues of the non-polar analytes passing into a primarily aqueous dialysate matrix or a slow rate of diffusion of the analytes from the tissue to circulatory compartments. Exercise appears to improve initial levels of excretion resulting in the maximum excretion at 0 hour. This declines quite rapidly, with a subtle increase at 1 to 1.5 hours and a gradual decrease in excretion until the end of the session. Exercise also seems to increase the amount present at the end of dialysis with a greater end point percentage excretion. A possible explanation for this is an increase in metabolic rate and production of the solutes within the tissues which gradually pass into the circulatory system.



Mean excretion of uremic analyte at m/z 180 (incl. m/z 359 and 381) shown as a normalised percentage of the largest peak area obtained at each timepoint throughout the hemodialysis session

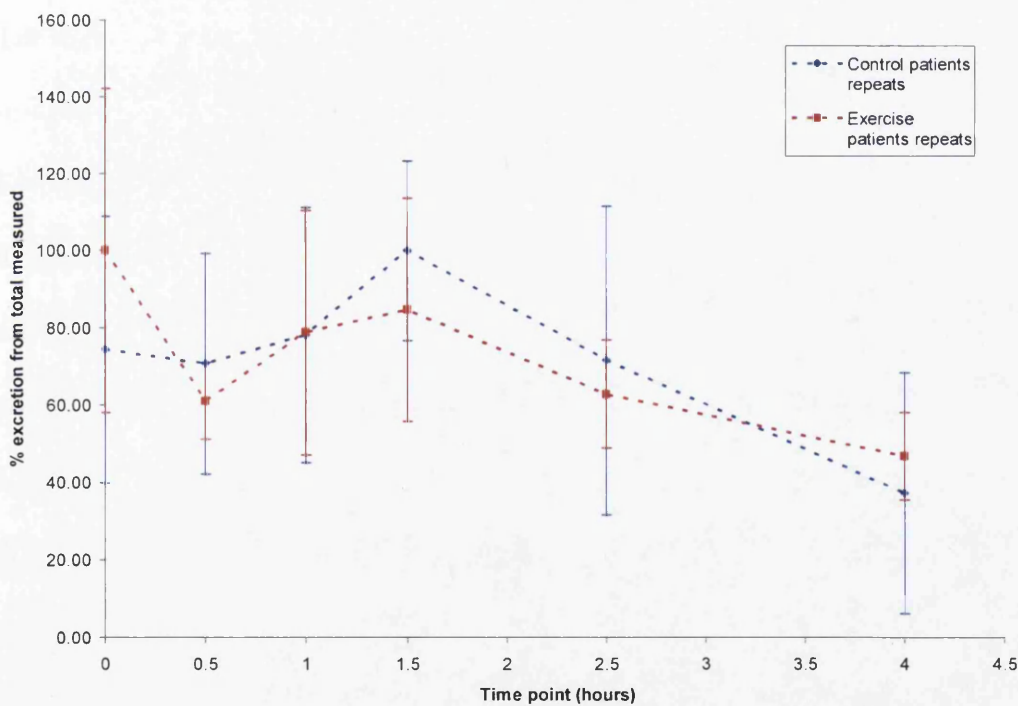
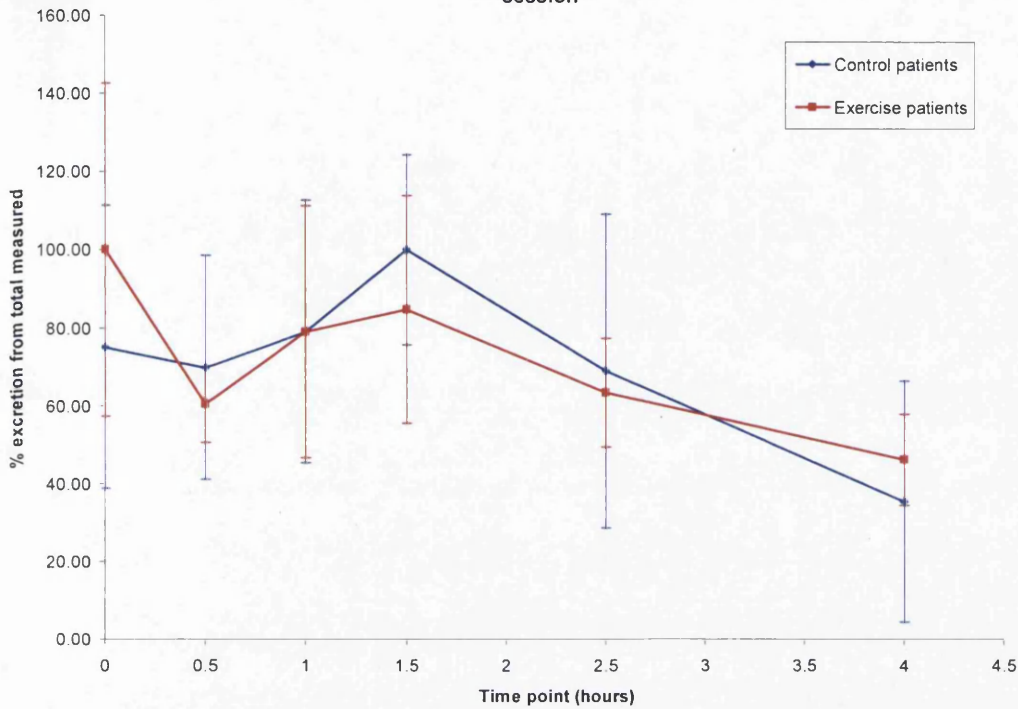


Figure 4.40: Excretory profiles for initial and repeat analysis of the novel uremic analyte present in haemodialysate obtained from control patients and those who have exercised prior to dialysis treatment (n = 10 and 6 respectively).

In summary, this relatively non-polar novel uremic analyte appears to have a delayed removal during dialysis, with maximum excretion observed at 1.5 hours. The addition of exercise to treatment appears to improve the initial extraction of analytes perhaps by increasing the amount present in the blood stream through an increased metabolic rate or transport from the tissues. The first point is supported by the proposed monomer identity assigned, indicating that this analyte may be related to benzoylglycine and the glycine metabolic pathway. However, the disparities between this profile, urea and creatinine suggest that these conventional markers aren't completely representative of the excretion of these novel uremic analytes.

#### 4.3.1.8. Candidate biomarker 6: $m/z$ 335.0545

This non-polar uremic solute generates highly reproducible excretory profiles and is insufficiently removed throughout haemodialysis regardless of the patient cohort. Evidence for this is provided by the maximum point of excretion present at the end of dialysis treatment. Patients undergoing dialysis alone appear to have poor initial removal possibly due to incompatibilities of solute and dialysate polarity or solubility, and peaks at 1 hour showing maximum excretion. This decreases slightly, but continues to increase to the end of the dialysis session, indicating an insufficient duration of dialysis treatment. Exercising prior to treatment seems to decrease the amount excreted at the start of dialysis. The percentage excretion then rapidly decreases and follows a similar pattern to that exhibited by control patients. However, unlike the control patients, the maximum level of excretion is observed with the end of session sample and a rapid increase in the level of this analyte is present within the patients' blood following 1.5 hours of dialysis treatment.

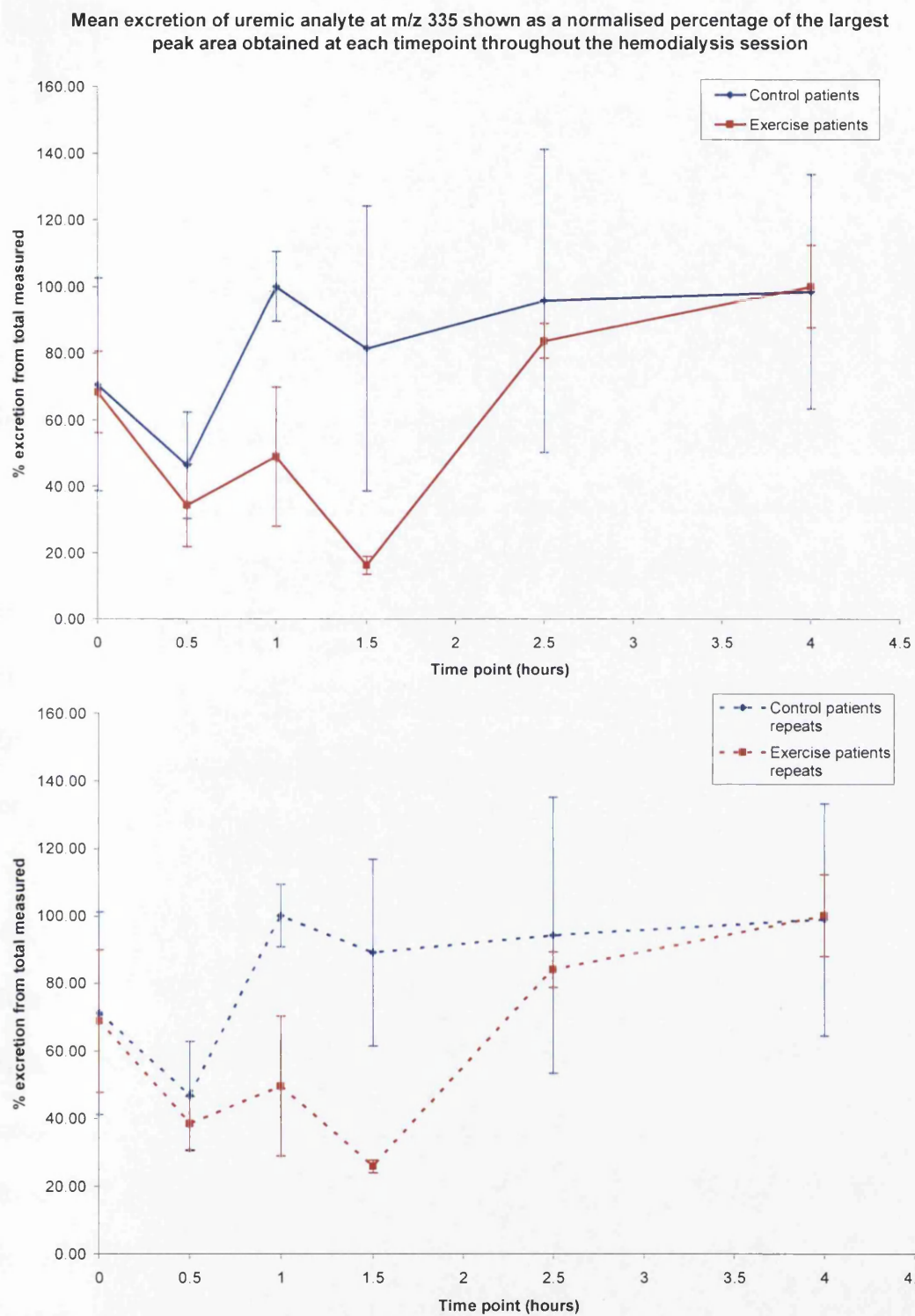


Figure 4.41: Excretory profiles obtained over two analytical runs of the novel non-polar uremic analyte observed at a RP-HPLC retention time of 30 minutes and at m/z 335 ( $n = 10$  and  $6$  for control and exercise patients respectively).

The proposed identity of this analyte (shown in section 4.2.4.6) suggests that it is formed as a result of malfunctioning metabolic processes. If this is correct exercise could be causing a decrease in this malfunction, activating the correct metabolic pathway and resulting in a decline of the uremic solute in the blood for removal. After more than 1.5 hours of inactivity, this faulty process may take precedence again causing the level of the uremic solute to increase in the blood stream. This may also be an underlying mechanism within control patients and could account for the greater level of excretion observed at all time points when compared to the exercise group.

In summary this non-polar analyte has very poor removal from the patient regardless of treatment. Exercise does affect excretion but greater levels appear to be removed from the control patients possibly due to greater amounts originally present. In comparison to profiles obtained for urea and creatinine, the excretion of these conventional markers does not appear to be representative of the solute at  $m/z$  335.

#### 4.3.2. Suitability of use as a biomarker

It is quite clear from the excretory profiles shown that chemical property of an analyte will dictate its removal during haemodialysis treatment. The clearance of urea and creatinine are the benchmark measurements by which adequacy of all haemodialytic treatment is monitored. However, these polar uremic solutes are thought not to be representative of some toxins retained in the body following treatment. Evidence for this has been given by this current work indicating that despite 'normal' urea and creatinine levels, other non-polar analytes (at  $m/z$  275, 180 and 335) are not removed sufficiently. This warrants the use of a selection of uremic analytes of a range of polarities to be monitored and assess the efficacy of haemodialytic treatment.

#### 4.3.3. Effect of exercise on haemodialysis excretion and its future in treatment

The inclusion of exercise prior to commencing haemodialysis treatment unlike the literature<sup>[4, 5]</sup> appears to have little affect on conventional markers, urea and creatinine, generating similar excretion profiles to the control group. The primary effect on the excretion of the novel polar uremic analytes is an increase in initial removal followed by a more profound compartment profile. The greatest impact on excretion was observed with non-polar analytes. Exercise appears to lower the amount excreted throughout the whole of the dialysis session and perhaps encouraging metabolism to function correctly decreasing the level of these analytes in the blood. A common characteristic observed with all analytes monitored was a greater level of variation in the data acquired for each time point. This may be due to disparities between patient exertion whilst exercising and highlights the requirement of more stringent guidelines during the exercise regimen. These could include maintaining the patients' heart rate at a specific level to ensure comparable results and minimising intra-time point or inter-patient variability.

#### 4.4. Summary of Identifying New Uremic Analytes and Their Application to Haemodialysis Treatment

Previous work carried out in Chapter 3 has shown that this methodology is capable of isolating and identifying analytes within the haemodialysate matrix. This has been applied with the aim of discovering novel uremic analytes that are observed to have a high degree of thermal stability at a reduced temperature of 10°C. A range of mass

spectrometric experiments has enabled for the vast majority of novel analytes a near complete elucidation of their chemical structure. These included extensive fragmentation investigations of the analyte and the acquisition of the elemental formulae for both the protonated molecule and the product ions. For the latter the LTQ Orbitrap proved vital particularly for the analyte at  $m/z$  214. The high mass resolution capabilities of this instrument showed a peak within 0.8 Th that have could provided an incorrect elemental formula and identity assignment by causing the ion peak to shift in mass-to-charge. There are two analytes that have unconfirmed complete structures and would require additional work to be carried out using the LTQ Orbitrap mass spectrometer. Unfortunately this was no longer at our disposal and will have to remain as future work until the instrument is available.

Following analyte identification we have examined the excretion of these novel uremic analytes within haemodialysate and evaluated their suitability as markers in assessing dialysis adequacy. Ultimately we have found that the chemical polarity of uremic solutes has a significant impact on their extraction from the body. Non-polar analytes proved particularly difficult to remove possibly due to their insolubility with the aqueous sampling matrix, the dialysate concentrate. Therefore we have highlighted the limitations of using urea and creatinine to monitor the excretion of this type of solute and the need to develop haemodialysis to enhance non-polar solute removal. Our suggestions regarding this may include using a non-polar additive within the dialysate concentrate, such as an inert surfactant or attaching a non-polar polymer to the external side of a conventional dialysis membrane. Exercise has shown to have some beneficial effects on solute removal during dialysis although, it's true affect is not completely understood and warrants further work. As a short term measure we believe that it is essential for future treatments to be monitored not

only using polar analytes such as urea and creatinine, but to use a solutes of a range of polarities, ensuring an unbiased measurement of dialysis adequacy.

## References

1. Dyer, JR., *Applications of absorption spectroscopy of organic compounds*. Prentice-Hall, New Jersey. 1965, p. 4-21.
2. Beynon, JH., Morgan, RP., and Brenton, AG., *Philosophical Transactions of the Royal Society*, 1979. **A293**: p. 157-166.
3. [www.mimas.ac.uk/crossfire/](http://www.mimas.ac.uk/crossfire/)
4. Kong, CH., Tattersall, JE., Greenwood, RN., and Farrington, K., *Nephrology Dialysis and Transplantation*, 1999. **14**: p. 2927-2931.
5. Parsons, TL., Toffelmire, EB., King-VanVlack, CE., *Archives of Physical Medicine and Rehabilitation*, 2006. **87(5)**: p.680-687.



## **CHAPTER 5:**

### **Characterisation of Novel MALDI**

### **Matrices for MALDI Tissue Imaging.**

#### **5.1. Matrix-assisted Laser Desorption/Ionisation (MALDI) Imaging Mass Spectrometry (MIMS)**

Imaging mass spectrometry is a developing technology at the forefront of protein analysis, using MALDI time-of-flight-mass spectrometry (MALDI-ToF-MS) to both profile and map proteins, and their related biomolecules, primarily from thin previously frozen tissue sections. When tissues are removed from the body degradative enzymes initiate a process of autolysis. Hence, to maintain integrity the tissues need to be processed immediately either by freezing or using a fixative. Current MIMS techniques use sections cut from frozen tissue rather than tissue that has been in contact with the fixative which has proven to be problematic. Common fixatives include paraffin wax which infiltrates the tissue at high temperatures after a process of dehydration. A possible effect of paraffin on tissue ionisation is the reduction of the number of potential ionic functional groups that can result in a mass-to-charge signal. Unlike the freezing process using liquid nitrogen, a fixative is employed for archiving tissue biopsies as it is carried out at lower cost than freezing and can still maintain the morphology of the tissue for later microscopic analysis. Therefore, developments in imaging mass spectrometry for paraffin sections could enable this technique to become commonplace for detecting disease states, such as cancer progression, or the interface of healthy and diseased regions in tissue biopsies.

We have investigated several different parameters of tissue preparation for MIMS applicable to:

1. paraffin-embedded,
2. deparaffinised and,
3. frozen sections.

These include investigating novel lipophilic fluorescent compounds for use as MALDI matrices, with the effects of solvent composition, and drying temperature. In order to improve the compatibility of paraffin based sections for MIMS technology antigen retrieval buffers were also examined to enhance protein/peptide signals.

#### 5.1.1. Tissue preparation

Current methodologies for protein and peptide analysis involve several forms of sample preparation<sup>[1]</sup> and are outlined below:

##### 5.1.1.1. Tissue blotting on polymer membranes

This involves transferring proteins and peptides from freshly microtomed tissue sections, by a blotting action, onto an organic polymer membrane previously mounted on a MALDI sample plate. Commonly used membranes consist of carbon-filled polyethylene as it is capable of maintaining a good electrical potential between the sample plate and the electrode within the MALDI ionisation source. Adsorption of proteins and peptides to the membrane occurs through electrostatic and hydrophobic interactions. The blotted areas are then washed with deionised water to remove residual tissue fragments, blood and most importantly salts. A matrix solution, such as sinapinic acid, is then applied to these washed regions whilst wet and air dried

ready for analysis. This protocol is generally used for the analysis of proteins and is capable of detecting masses of up to 100kDa. Past studies have shown that it is also capable of achieving reproducible results between tissue sections and is applicable to different organs such as lung, heart, pancreas, epididymis, brain and kidney<sup>[2, 3]</sup>.

#### 5.1.1.2. Tissue imaging using laser capture microdissection (LCM)

This preparative technique initially involves dissecting cells identified from a previously stained and fixed tissue section. Selected regions for dissection are irradiated with an infrared laser transferring the cells to an organic membrane or film, commonly consisting of ethylene-vinyl acetate, and matrix solution is applied as a droplet using a narrow capillary. This membrane is then attached to the MALDI sample plate using double-sided carbon-filled polyethylene conductive tape ready for analysis. A suggested advantage of this protocol is the ability to obtain a good quality protein signal from small amounts of cellular material<sup>[4]</sup>. Although, this method is not necessarily suitable in determining the spatial distribution of proteins *in-situ* and the non-conductive properties of the organic film can reduce the resolution of signals obtained with the ToF mass analyser.

#### 5.1.1.3. Direct MALDI imaging

Tissue sections are produced using a cryotome at -20°C and placed immediately onto a MALDI sample plate. Matrix solution is deposited over the whole slice of tissue, dried in a desiccator and then mass analysed. There have been many protocols designed for direct matrix application with the intention of achieving a homogenous coating of matrix crystal yet maintaining the spatial arrangement of the proteins and peptides. The first is the 'sliding drop' method and involves loading the tissue with a

large volume of matrix solution which is then spread over the whole section. It has the capability of obtaining relatively good signals for proteins and peptides but is prone to disrupting the spatial arrangement of these biomolecules. The conventional MALDI dried-droplet spotting method was also developed for tissue imaging<sup>[2]</sup>. This used a narrow capillary to apply small volumes of matrix solution, usually 100-200nl, at regular intervals across the tissue section. The solution is dried naturally or by a desiccator and then mass analysed in a raster pattern. Another popular approach involves loading approximately 0.5mL of matrix solution in a pneumatic airbrush and sprayed at pressures of 1-1.5bar onto a tissue section previously placed on a MALDI sample plate. This matrix coating must be applied within 10cm of the section and allowed to dry prior to the next matrix application, with successful imaging often achieved after 10 repeat applications. Following this, the crystallised matrix layer is resolubilised by spraying a solvent mixture to redistribute any heterogeneously arranged matrix crystals. This technique can prove problematic where poor crystal distribution and density on the tissue section can result in a reduction of protein signals above masses of 20kDa. Attempts to overcome this problem have involved an electrospray apparatus in which an electric current is placed on the needle of a pneumatic sprayer. Again, this technique is also susceptible to problems in matrix distribution and to achieve a homogenous layer the sprayer must be placed at an optimum distance from the tissue section. This is to ensure the matrix does not arrive at the tissue either dehydrated, resulting in an amorphous matrix deposition, or, overly wet causing the resolubilisation of the matrix solution.

The modified dried droplet methodology is advantageous when compared to the other preparative techniques as it is capable of determining the spatial arrangement of biomolecules with minimal sample preparation steps. There is also

the possibility that the detected biomolecules are in fact subcellular since the initial sectioning step could expose those present in the cytoplasm of the cells<sup>[1, 2, 3, 4]</sup>.

### 5.1.2. Tissue analysis by MALDI mass spectrometry

Following matrix application the sample plate is placed within the MALDI ionisation source and mass spectra acquired by a raster motion over the MALDI plate eventually obtaining data as x and y-coordinates of the plate. Ions obtained for a specific area of tissue is known as a '*profile*', whilst a complete collection of ions can be plotted as a 2D ion density map to give the mass image of the tissue<sup>[2, 5]</sup>. This MIMS technology is intended to provide information regarding spatial distribution of biomolecules, such as peptides and proteins, and has the potential to show any variations that may correspond to disease states. Hence, there is a definite interest in identifying the protein and peptide complement of these tissue profiles and images for use as a biomarker(s). These may then be used to outline the stage of a particular disease, the boundary of healthy and diseased tissue within a biopsy and to highlight the need to initiate or suspend a course of treatment.

Current methods of protein and peptide identification by mass spectrometry rely on database searching of both the molecular species and fragment ions. Searching according to mass-to-charge of the molecular species can encounter problems since many of the databases include the precursor molecule and do not account for post-translational modifications. Therefore, a more accurate identification process is required and is achieved by matching peptide fragments obtained after digestion with endoproteases, such as trypsin, with those expected in the database. However, for this to have sufficient accuracy the target protein must be isolated from the tissue section. An approach frequently employed is high-performance liquid chromatography

(HPLC) and can separate proteins into fractions ready for analysis as an on-line process with electrospray ionisation (ESI) or off-line using MALDI mass spectrometry. In addition to tryptic digestion, isolated proteins can be fragmented and identified by mass spectrometry alone. MALDI time-of-flight (MALDI-ToF) instruments can only achieve some degree of fragmentation through a process of post-source decay (PSD) and sequentially stepping the voltage applied to the electrostatic reflectron lens. Hence, the most commonly used instruments for structure elucidation are interfaced with ESI using a quadrupole ion trap and quadrupole time-of-flight (Q-ToF) mass analyser. Although, unlike ESI, MALDI may be used for the direct analysis of tissue and is not as susceptible to ionisation suppression, making it better suited to identifying biomolecules *in situ*. The use of a tandem time-of-flight (ToF-ToF) mass spectrometer is likely to become an important part of MALDI imaging. This incorporates the benefits offered from both techniques described above<sup>[6]</sup> and acquires data without contamination from different proteins in close spatial proximity as observed with poorly separated tryptic peptide ions.

#### 5.1.2.1. Matrix-assisted laser desorption/ionisation time-of-flight/mass spectrometry (MALDI ToF-MS)

This analytical technique has revolutionised the analysis of large biomolecules, such as proteins<sup>[7, 8, 9]</sup>. A laser at a UV wavelength of 337nm, is used to irradiate the sample mixed with UV absorbing additive, known as a matrix. The energy absorbed by the matrix doped sample causes ultra rapid heating and volatilization resulting in the ionisation of sample analytes through several processes<sup>[10, 11]</sup>. The resulting gas phase ions then pass into the mass analyzer for mass measurement by recording their time-of-flight using the relationship shown overleaf.

$$t = d \sqrt{\frac{m}{2zeV}}$$

where,  $t$  = time of flight,  
 $d$  = distance travelled,  
 $m$  = mass of ion,  
 $V$  = accelerating voltage and  
 $z$  = number of charges on the ion.

Several types of lasers are available for producing UV radiation, each with specific characteristics and are chosen for optimum performance for a particular application. For example, a Nitrogen laser can be obtained at relatively low cost but has a limited repetition rate of 10Hz with a total number of shots of approximately  $10^7$  before a new tube is required. However, for imaging a Nd<sup>+</sup>/YAG tripled laser is preferable which can work at 100Hz or more repetition rate with a total number of shots of approximately  $10^9$ .

The selection of suitable matrices is important and needs to match the different type of target molecule such as, synthetic polymers or biomolecules<sup>[9, 12]</sup>. In addition to matrix choice other parameters such as, solvent composition and the inclusion of additives, such as trifluoroacetic acid (TFA), facilitate and improve ionisation of the target molecule. MALDI spectra generally consist of protonated molecular species of the sample components and matrix. However, additional peaks may be observed as a consequence of sample degradation during laser application, and adduct formation, from the presence of sodium and even potassium ions in biological samples. These considerations must be accounted for during spectral interpretation and have particular importance when applying a database search, such as those used in protein

identification. A major upshot of this technique is the ability to identify very large molecular weights of up to 500kDa<sup>[8]</sup> and when used for bioanalysis traditionally involves an extracted sample. This extraction or preparative process however, can often result in a loss of information of the sample and therefore, a preferred method would be an analysis *insitu*. Prior to MIMS proteins were removed from their native tissue samples and either analysed as a whole molecule or as peptides through use of enzymes, for example, the endoprotease, trypsin. The peptide ion pattern found were then searched using a protein database such as SwissProt (UniProt Knowledgebase, Geneva, Switzerland) and likely protein identities ranked according to criteria such as the number of matching peptide ions.

#### 5.1.2.2. Fundamentals of MALDI Matrices for identifying new matrices for imaging mass spectrometry

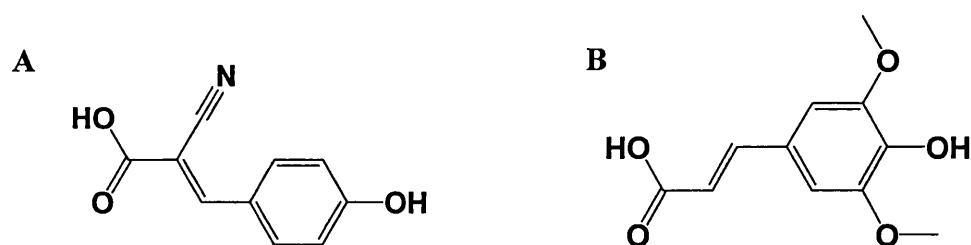
Laser desorption ionisation (LDI) included a material later known as a matrix after 1985 when Karas and co-workers attempted to investigate the ionisation mechanisms involved in films of UV absorbing amino acids<sup>[7]</sup>. Previous LDI methods were poor at ionising biomolecules and the inclusion of a matrix enabled this application by performing the following functions<sup>[5]</sup>:

- a. absorbs laser energy and then intermolecularly transfers energy to the analyte leading to thermal desorption (and excitation of the matrix and analyte),
- b. protects analyte from in-source decomposition,
- c. enhances ion formation following matrix photoexcitation through a process of proton transfer to the analyte species and
- d. prevents formation of sample aggregates.



Despite this knowledge the complete nature of ionisation processes in a matrix are not fully understood. One essential requirement for a good matrix is that it generates a high analyte ion yield. The choice of matrix is critical in achieving good ionisation as different matrices are suited to specific classes of molecules and mass ranges, and few so far have been found suitable for high molecular weight compounds (>25kDa). For example,  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) is suited to ionising proteins and peptides below 10kDa, while sinapinic acid (SA) is used for ionising proteins above 10kDa. Presently a large number of matrices have been identified and cover a whole spectrum of biomolecules, making MALDI suitable for wide ranging analysis.

A large proportion of these matrices are derivatives of both benzoic acid and cinnamic acids. All these compounds have a common chromophore, essential for absorbing the 337nm UV radiation, and were recognized at the early stages of matrix development as being good for ionising proteins<sup>[13]</sup>.



*Figure 5.1: Chemical structures of two commonly used cinnamic acid derivatives, A =  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA), B = sinapinic acid (SA). Unlike CHCA, SA does have a phenolic hydroxy group in the ortho-position, therefore some other functionality of CHCA, such as the carboxylic acid group in conjunction with the alkyl double bond, must contribute to the ionisation of biomolecules. This disparity in chemical structure indicates that other groups may account for the ionisation of molecules of differing masses.*

In particular the most successful of these matrices contain a phenolic hydroxy group in an *ortho* position to a carbonyl functional group. These functional groups when exposed to UV radiation undergo an intramolecular proton transfer from the carboxylic acid moiety. This movement is supported by the reduction in acidity and deficiency in proton affinity (PA) of the carboxyl group in the subsequent metastable excited structure. Proton affinity of the functional groups of both matrix and analyte is thought to be an important parameter in determining successful ionisation of the analyte. For example, a matrix with a low PA when compared to the analyte will undergo a highly probable proton transfer and ionise the analyte since it will occur as an exothermic reaction. Unlike the corresponding *meta* and *para* isomers, the *ortho* metastable excited structure has a relatively long lifetime ( $\sim 3\text{ns}$ )<sup>[12]</sup>. This, in addition to the acidic phenolic hydrogen atoms and disparity in proton affinities, could account for the protonation of the analyte via the matrix and the better performance of the *ortho* isomers in UV MALDI. Hence, in the search for a new improved MALDI matrix proton affinities and this *ortho* structure may prove to be a fundamental requirement.

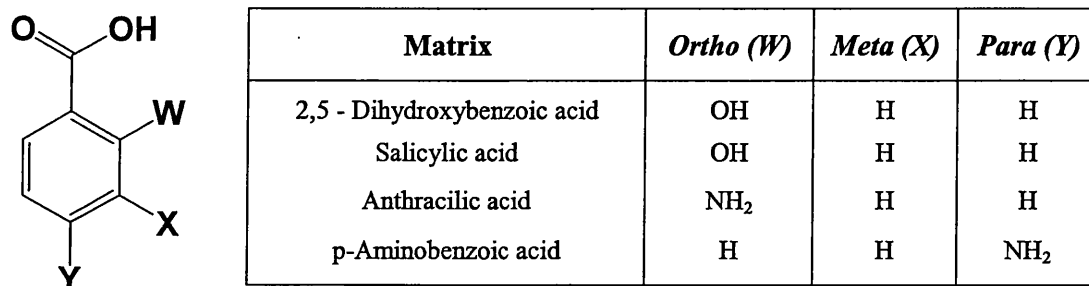


Figure 5.2: Hydroxycarbonyl structure consistent with some conventional matrices, *W* = *ortho*-position, *X* = *meta*-position, *Y* = *para*-position.

MALDI matrices may be classified according to their ground state configuration and whether they are solid, liquid, and liquid/solid two-phase matrices. Solid matrices are most commonly used in MALDI experimental protocols since they tend to be less expensive and can involve minimal sample preparation. A major problem encountered with solid matrices is the existence of 'hotspots' in the crystal lattice. These are areas of crystallisation that are capable of intense ionisation which are often unidentifiable by optical microscopic study and only become apparent from MALDI spectra of that region. If this is an undesirable characteristic for the analysis liquid matrices are often chosen as a replacement since the surfaces can be continuously re-solubilised reducing the probability of a 'hotspot'<sup>[11]</sup>. Liquid/solid two-phase matrices consist of an absorbing solid material mixed with a non-absorbing liquid. Since not all solid matrices are available in liquid form, this method was developed to maintain the performance of the solid matrices with minimising the likelihood of 'hotspots'<sup>[11]</sup>.

MALDI matrices also have an important role in subsequent fragmentation of an ion. They can impart varying amounts of internal energy into a neutral molecule and it is protonated following ion molecule collisions in the MALDI gas plume. Internal energy effects are evident by fragmentation including metastable decay, known as post-source decay (PSD). Matrices are considered 'hot' or 'cold' for high or low amounts of fragmentation, respectively. For example, sinapinic acid is considered to be a 'hotter' matrix than 2,5-dihydroxybenzoic (DHB) particularly for glycoproteins<sup>[14, 15, 16]</sup>, and yet DHB causes more fragmentation than 3-hydroxypicolinic acid (3-HPA) when used for the analysis of oligodeoxynucleotides. There are several theories proposed to explain this variation in post-source decay, ranging from differences in sublimation temperature of the matrices<sup>[17]</sup>, the impact of

internal temperature on sublimation<sup>[18]</sup>, and possible hydrodynamic affects during expansion of the MALDI gas plume<sup>[19]</sup>.

At present there are only empirical guidelines for what compounds will function as good matrices derived from experimental observations. However, matrix compounds in general have high absorption at the chosen laser wavelength, good solubility in the solvent preparation, the potential to form a lattice structure, can sublime and ionise, and have good vacuum stability<sup>[20]</sup>. Matrices are commonly identified as a consequence of screening a large variety of compounds. Understanding the process of MALDI ionisation would allow greater specificity in selecting new matrices and despite many suggested mechanisms only those described previously have been confirmed. These known mechanisms can not fully explain the phenomena observed in MALDI and it is believed that many others, such as disproportionation reactions and multiphoton ionisation have a significant contribution to the total ionisation process<sup>[11]</sup>.

## **5.2. Investigating Novel Matrices for Tissue Imaging**

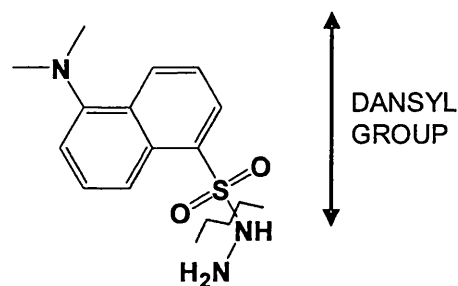
Existing MIMS protocols have several limitations and disadvantages; they are only applicable to frozen tissue sections that are rarely available and current matrices are inadequate in accessing certain biomolecules such as membrane bound proteins<sup>[21]</sup>. Thus, there is a distinct need to develop the existing selection of matrices to include a matrix that has the capability to ionise protein from the more common paraffin embedded tissue sections and those proteins that are currently not amenable by MIMS.

### 5.2.1. Discussion of novel matrices

A total of five novel compounds were chosen as matrices with the following desirable characteristics:

- a. absorbs UV energy at the laser wavelength of 337nm,
- b. has a low molecular mass (<500Da),
- c. fluoresces and,
- d. has a hydrophobic/lipophilic nature.

These properties were chosen with the intention of selecting a matrix that will ionise the biomolecular complement of the tissue, producing a low background signal at higher mass, and be capable of accessing membrane-bound proteins, located by both mass spectrometry and fluorescence microscopy. Four of the chosen novel matrices from commercially available materials were dansylated compounds, the first dansylhydrazine, has previously been used in improving the detection and quantitation of glycoproteins<sup>[22]</sup> and hormones<sup>[23, 24]</sup>. The three remaining dansyl compounds dansylcadaverine, dansyl-DL- $\alpha$ -aminocaprylic acid, and 11-(dansylamino)undecanoic acid (Sigma Aldrich, Poole, UK) are fatty acid analogues previously used to enhance detection of non-polar polymers by fluorescence detection<sup>[25]</sup>. The dansyl chromophore is capable of absorbing energy at UV wavelengths and emitting at fluorescence wavelengths possibly sufficient for the MALDI ionisation of biomolecules forming a molecular ion. This chemical moiety in the form present in these novel matrices is also known to covalently react with carbonyl groups resulting in an amide bond. Thus, these dansyl matrices may have the potential to ionise biomolecules such as peptides and proteins through a process of adduct formation.



*Figure 5.3: Dansylated compound indicating by the section above the red line the dansyl functional group. It is this group that accounts for both the UV absorbing and fluorescence properties of some of the novel matrices.*

The fifth matrix, fluorescamine or Fluram (Sigma Aldrich) was chosen since it will only absorb at the laser wavelength, 337nm, when bound to a protein or peptide. This has the potential to limit the background signal from ionising the matrix when examining peptides at relatively low mass ranges. This compound contains a number of regions of *pi*-electrons required for the chromophore structure in the form aromatic rings and carbonyl groups as shown in figure 5.4.

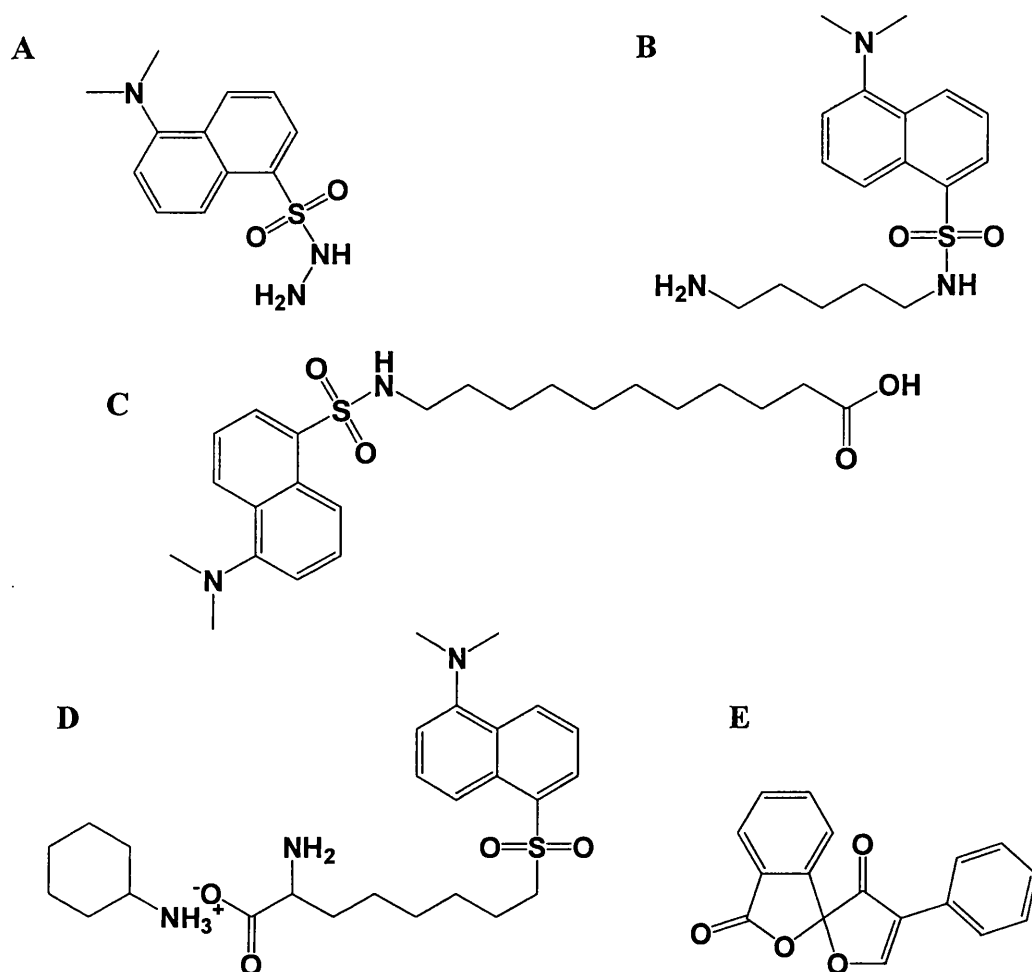


Figure 5.4: Chemical structures of the chosen novel MALDI matrices for tissue imaging. A = dansylhydrazine, B = dansylcadaverine, C = 11-(dansylamino)undecanoic acid, D = dansyl-DL- $\alpha$ -aminocaprylic acid and E = fluorescamine.

#### 5.2.1.1. Fluorescence microscopic study of novel matrices

The absorption of energy, in the form of a photon, will elevate a molecule to an excited energy state. This energy can be lost by radiationless processes, collisions with surrounding molecules followed by radiation emission, called fluorescence, between the excited and ground electronic states. Fluorescence is a radiational emission or energy transition involving electrons of the same spin quantum numbers,

such as an excited singlet to a ground singlet state ( $S_1 \rightarrow S_0$ ), and results in the emission of light at a higher wavelength. Fluorescence emission has a relatively short lifetime of  $10^{-8}$  to  $10^{-4}$  seconds, is structurally dependent and not always observed with chromophores as an excited molecule is also likely to emit this energy through a radiationless process of internal conversion. Chemical groups capable of absorbing this energy and emitting it as fluorescent light are known as fluorophores, and these are generally unsaturated aromatic or carbonyl groups. One advantage of measuring fluorescence emission is that it has high sensitivity as only a small proportion of chemical structures contain fluorophores, and improves the signal-to-noise ratio by limiting the background signal. Thus, even at relatively low concentrations of a novel matrix, its location on the tissue can be detected even if it is not visually observed.

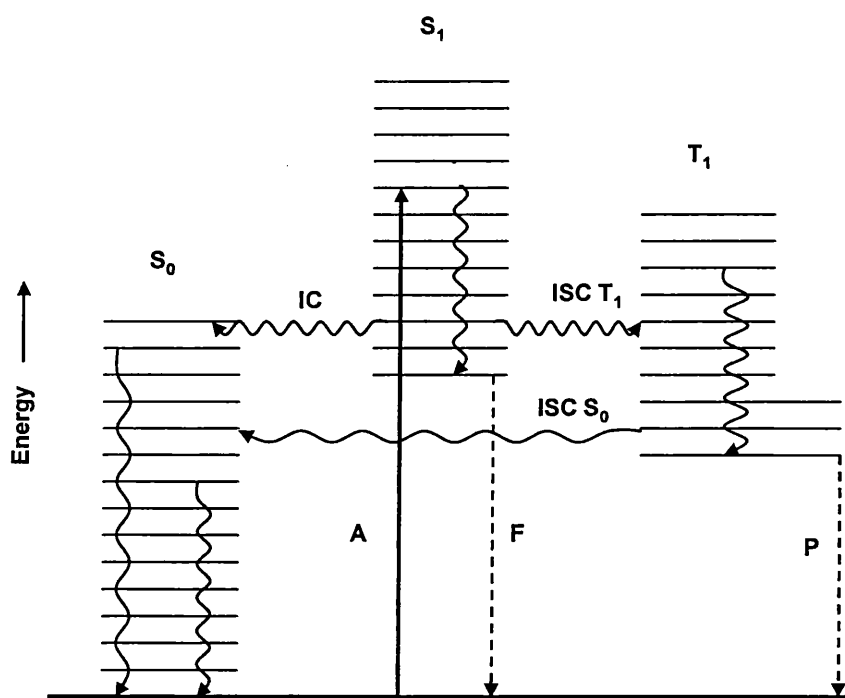


Figure 5.5: Energy diagram of the ground and excited singlet states ( $S_0$  and  $S_1$ , respectively) and a triplet excited state ( $T_1$ ) of a molecule. Each arrow shows an energy transition within the excited molecules. A – absorption, F – fluorescence, P – phosphorescence, ISC – intersystem crossing, and IC – internal conversion. The



curly arrows indicate a form of radiationless energy transition or emission through relaxation. The dashed arrows illustrate luminescent transition where energy is emitted from an excited state as light either in the form of fluorescence or phosphorescence.

#### 5.2.1.2. Application of novel compounds as MALDI matrices

From the chemical structures of these novel matrices it is apparent that they do not contain the suspected ionising moiety of a phenolic hydroxy group *ortho* to a carbonyl functional group. However, the dansyl matrices do have very basic amino groups, capable of accepting a proton and carboxylic acid groups, with the potential to donate a proton. Thus, despite the absence of the *ortho* structure these matrices could still have the potential to ionise target analytes by proton transfer between the analyte and matrix and function in both positive and negative ionisation modes.

Novel Matrix	Solvent	$\lambda_{\text{ex}}$ max	Reaction Characteristics	$\lambda_{\text{ex}}$ max	$\lambda_{\text{em}}$ max
Dansylhydrazine	MEOH	336	Saccharide adduct	336	487
Dansylcadaverine	MEOH	335	Reaction product	338	500
Dansyl-DL- $\alpha$ -aminocaprylic acid	MEOH	335	Protein adduct	-	523
11-(dansylamino)undecanoic acid	MEOH	335	Protein adduct	-	505
Fluram	ETOH/ ACETONE	234	Amine/protein adduct	334 (ETOH) /360 (ACETONE)	455

Table 5.1: Absorption data of the chosen novel matrices giving the excitation and emission wavelengths ( $\lambda_{\text{ex}}$  and  $\lambda_{\text{em}}$ ) of both the molecular species and any adducts formed.

The dansylation reaction is commonly employed in biochemistry to make a biomolecule, and in particular, proteins, peptides and amino acids, amenable by fluorescent detection<sup>[26, 27, 28]</sup>. For proteins and peptides, this reaction occurs with the  $\epsilon$ -amino group of lysine, the hydroxyl group of tyrosine and the N-terminal amino group. It has previously been used in conjunction with MALDI mass spectrometry to improve the sequence coverage of tryptic peptides for protein identification. Park and co-workers found a 5-fold improvement in sensitivity of peptides containing a C-terminal arginine upon dansylation in spite of a 2-fold dilution, and the added advantage of detecting, by a mass shift of 233.051, low molecular weight peptides often hidden by matrix peaks<sup>[29]</sup>. It was therefore thought that these compounds could have the potential to both ionise the target analytes of proteins and peptides, and bind to them for use as a fluorescent tag.

### 5.2.2. Investigating sample preparation protocols for novel matrices

#### 5.2.2.1. Practical considerations of solvent composition

From previously published work it is apparent that sample preparation, in particular, solvent composition and hence matrix solubility is an important factor in matrix crystallisation and overall ionisation of the target analytes<sup>[30, 31]</sup>. The solvents selected for the novel matrices are those which enable absorption at the laser wavelength. Hence, for the dansylated derivatives methanol was preferred, and for Fluram, both ethanol and acetone were selected. The Fluram-acetone solution when mixed with proteins, peptides or amino acids absorbs UV radiation at a wavelength at approximately 360nm (as indicated by table 5.1). Despite the disparity between this and the laser wavelength, this solvent mix was still investigated as it has been used as the preferred option for using Fluram as a spray solution<sup>[32]</sup>. Therefore it was

believed that this solution utilized in conjunction with an electrospray-type application may warrant investigation.

In order to determine the suitability of these solutions four different organic solvent compositions in aqueous were studied, 100%, 70%, 50%, and 30% as a 10mg/mL solution with each novel matrix. The latter two were selected as the conventional protocols for both CHCA and SA use 50% and 30% acetonitrile, respectively. All the 100 and 70% organic solutions (methanol, acetone and ethanol) were deemed unsuitable for MALDI since upon application to the sample plate the solution spread uncontrollably into other sample wells. Hence, 50 and 30% organic solutions were selected for further evaluation with biomolecular standard reference materials.

#### 5.2.2.2. Testing novel matrices for protein and peptide analysis

##### 5.2.2.2.1. Protein and peptide mixes

Sequazyme calibration protein and peptide mixes, and BSA solutions (Applied Biosystems, Foster City, CA, USA) were used to test the performance of the novel matrices for MALDI analysis. These are originally freeze-dried mixes and required reconstitution using the supplied standard diluent in the Sequazyme kit (Applied Biosystems). These calibration solutions were diluted 1/25 (1 $\mu$ L calibration standard in 24 $\mu$ L matrix) with each novel and conventional matrices (CHCA, DHB, SA).

Conventional matrix solutions were made as follows:

- $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) – 10mg/mL in 50:50 acetonitrile:0.1%TFA,
- 2,5-dihydroxybenzoic acid (DHB) - 10mg/mL in 50:50 acetonitrile:0.1%TFA,
- sinapinic acid (SA) – 10mg/mL in 30:70 acetonitrile:0.3%TFA

The calibration standard-matrix mixes were spotted according to the dried droplet protocol by depositing 1  $\mu$ L of the mix on a MALDI plate in duplicate. In addition to these mixes 'blank' matrix solution was also spotted to highlight common matrix peaks and any interferences present.

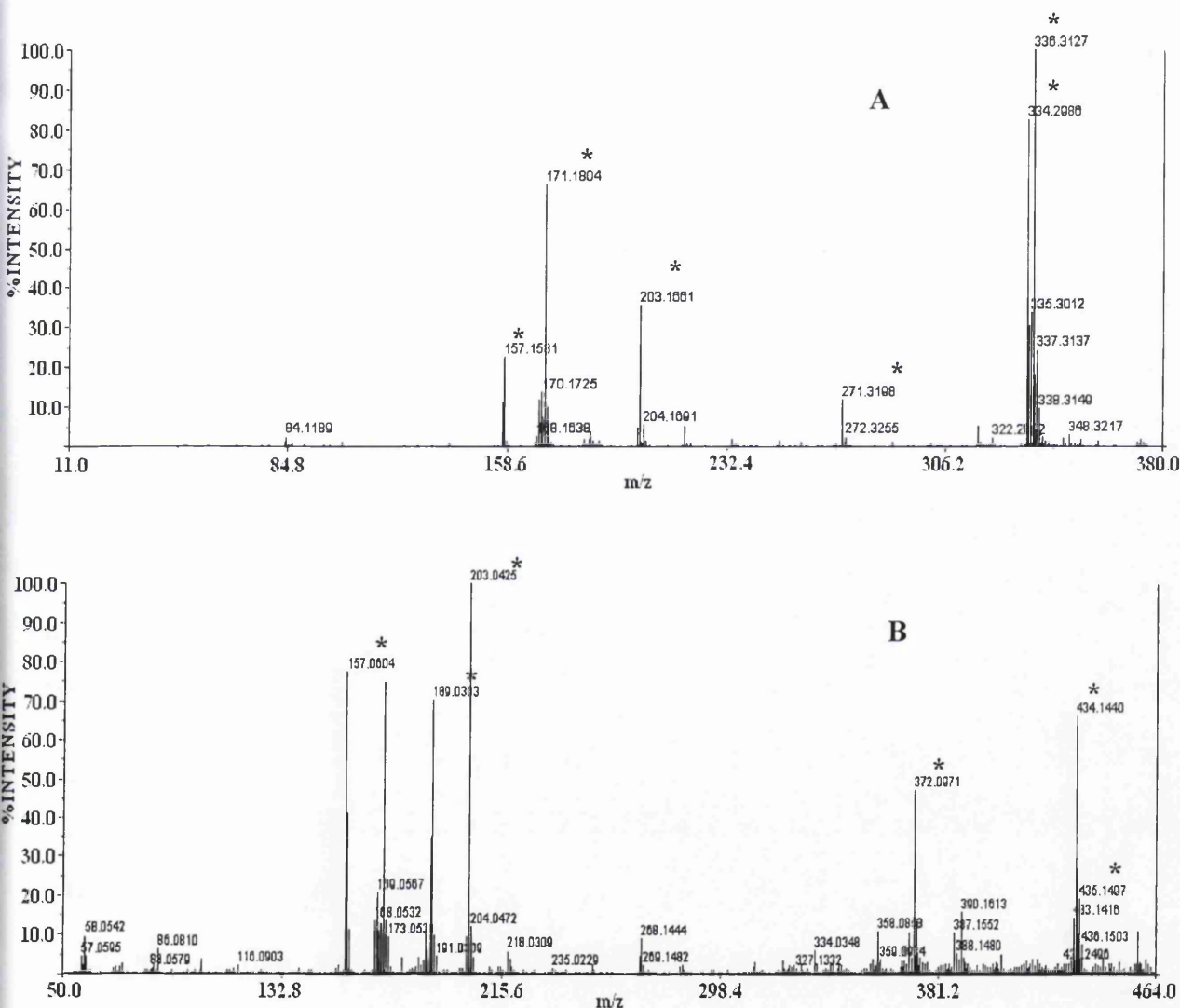


Figure 5.6: Mass spectra of the novel matrices; **A** = Dansylcadaverine (DC), molecular weight of 335, and **B** = 11-(dansylamino)-undecanoic acid (DUA), molecular weight of 435, displaying common ions (\*) listed in table 5.2.

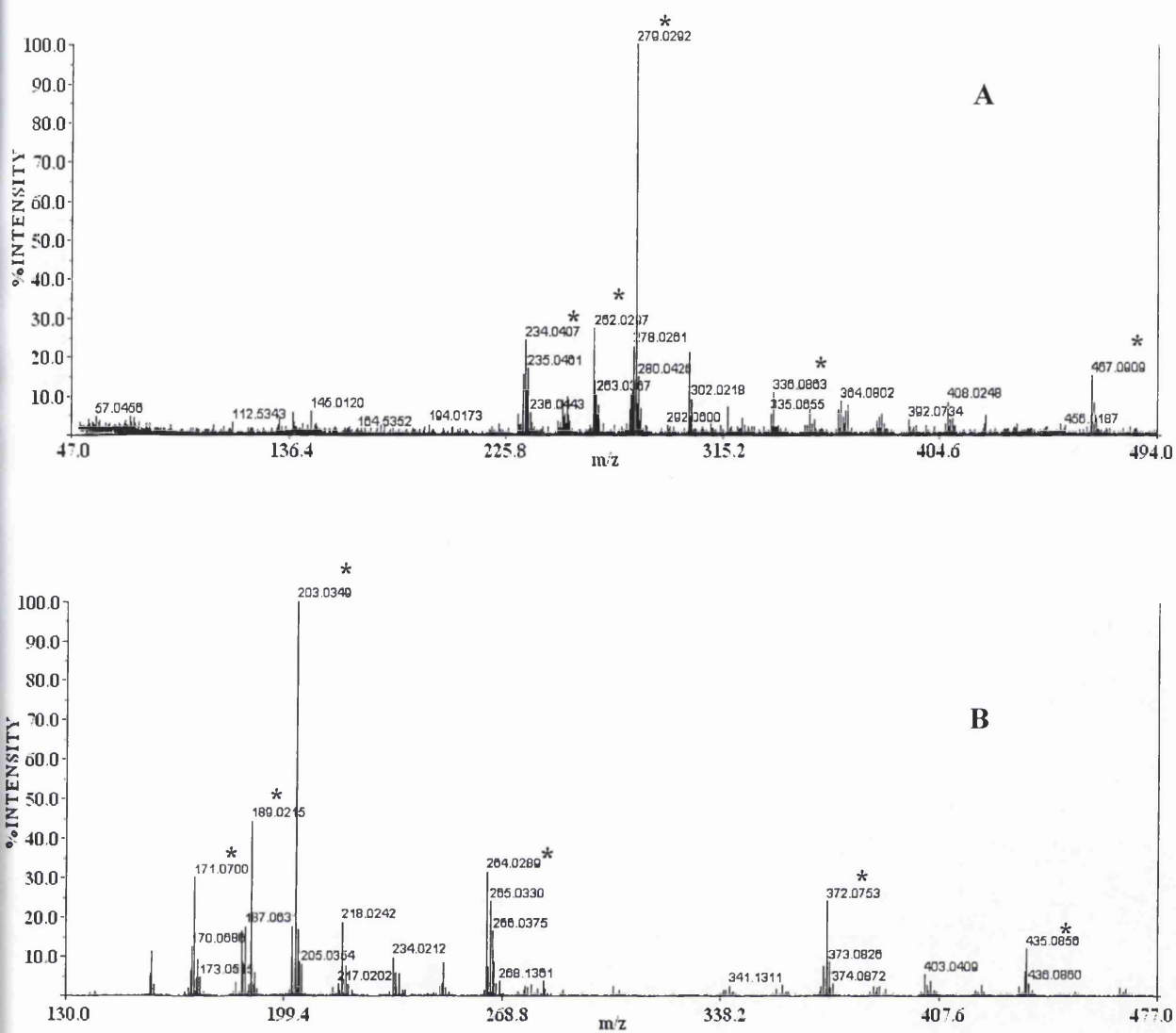


Figure 5.7: Mass spectra of the novel matrices; A = Fluram, molecular weight of 278, and B = Dansylhydrazine (DH), molecular weight of 265, indicating their associated ions (\*) as illustrated in table 5.2.

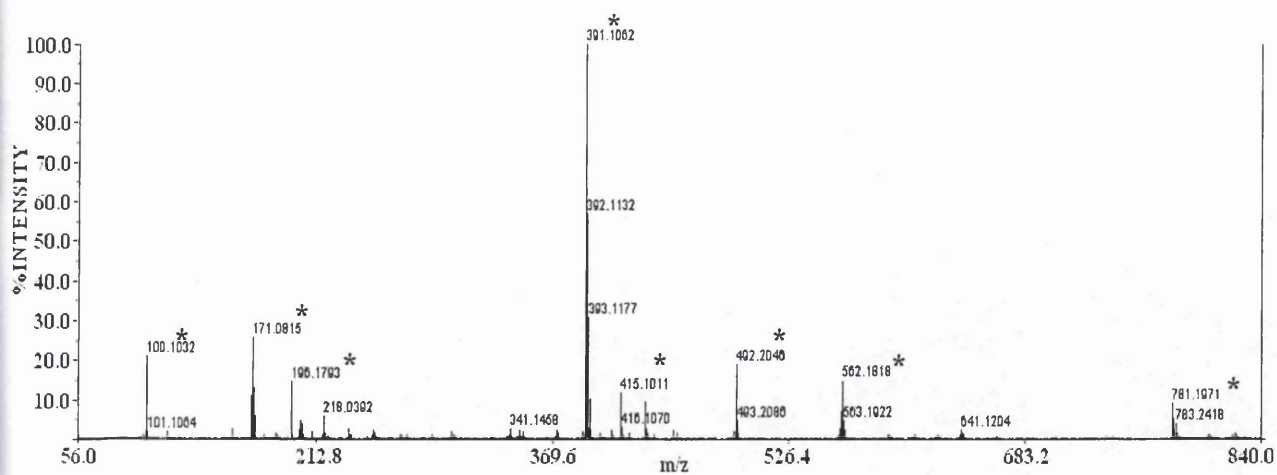


Figure 5.8: Mass spectrum of dansyl-DL- $\alpha$ -aminocaprylic acid (DCA), molecular weight of 492, showing the ions common to this novel matrix (\*) as displayed in table 5.2.

Matrix	Nominal Molecular Weight (Da)	Nominal Molecular Ion $[M+H]^+$ ( $m/z$ )	Matrix Ions (Da)
$\alpha$ -cyano-4-hydroxycinnamic acid (CHCA)	189	190	172, 190, 212, 228, 234, 250, 294, 379, 417
2,5-dihydroxybenzoic acid (DHB)	138	139	154, 178, 192, 211, 268, 304, 333, 367
Sinapinic acid (SA)	224	225	207, 225, 263, 387, 471
Dansylhydrazine (DH)	265	266	435, 372, 266, 265, 203, 189, 171, 157
Dansylcadaverine (DC)	335	336	336, 335, 271, 203, 171, 157
11-(dansylamino)-undecanoic acid (DUA)	435	435	435, 434, 372, 203, 189, 171, 157
Dansyl-DL- $\alpha$ -aminocaproic acid (DCA)	492	492	781, 562, 492, 415, 391, 196, 171, 100
Fluram	278	279	467, 363, 316, 279, 262, 234

*Table 5.2: Neutral and charged molecular species of novel and conventional matrices and their commonly observed ions. These ions must be taken into consideration when inspecting mass spectra of tissue at a lower mass range.*

Some ions are observed specifically with the dansylated matrices, for example, the ion  $m/z$  171 is consistent with the loss of the tertiary amino group plus the two phenyl rings from the typical dansyl functional group. The identity of these matrix peaks is essential for low mass (below  $m/z$  500) analyses in order to avoid mis-identification of biomarkers during the eventual application.

Standard	Components	Charge State	Average $m/z$ of Ion	Monoisotopic $m/z$ of Ion	Final Concentration (with Matrix) /pmol/ $\mu$ L
CAL MIX 1	Bradykinin	+1	905.05	904.4681	1.0
	Angiotensin 1	+1	1,297.51	1,296.6853	1.3
	Glu-fibrinopeptide B	+1	1,571.61	1,570.6774	1.3
	Neurotensin	+1	1,673.96	1,672.9175	0.05
CAL MIX 2	Angiotensin 1	+1	1,297.51	1,296.6853	2.0
	ACTH (CLIP 1-17)	+1	2,094.46	2,093.0867	2.0
	ACTH (CLIP 18-39)	+1	2,466.72	2,465.1989	1.5
	ACTH (CLIP 7-38)	+1	3,660.19	3,657.9294	3.0
	Insulin	+2	2,867.80	2,865.8083	3.5
	Insulin	+1	5,734.59	5,730.6087	
CAL MIX 3	Insulin	+2	2,867.80	-	0.5
	Insulin	+1	5,734.59	-	
	Thioredoxin	+2	5,837.74	-	2.75
	Apomyoglobin	+2	8,476.78	-	4.0
	Thioredoxin	+1	11,674.48	-	2.75
	Apomyoglobin	+1	16,952.56	-	4.0
BSA	BSA	+2	33,216	-	
		+1	66,431	-	

*Table 5.3: Peptide and protein composition of the Sequazyme calibration mixes and BSA solution (Applied Biosystems) used for evaluating the novel matrices. The singly and doubly charged ions of those shown in the table may also be present in the experimental spectra.*

#### 5.2.2.2.2. Solvent composition

There was no significant trend regarding the preferred solvent mixtures of 30 and 50% organic (methanol for the dansylated derivatives, with ethanol and acetone for Fluram) for the analysis of proteins and peptides in both positive and negative modes. This characteristic appears to be matrix dependent and does not seem to have a major impact on ionisation at these organic solvent levels. Additives were also incorporated into solution to study the affect on protein and peptide ionisation in both positive and



negative modes. Conventional MALDI matrix protocols for these biomolecules involve CHCA (10mg/mL in 50:50 acetonitrile/0.1% TFA) for masses less than 10kDa, and sinapinic acid (10mg/mL in 30:70 acetonitrile/0.3% TFA) for masses greater than 10kDa. These percentages of TFA are used for operating in both positive and negative mode and were tested with the solvent systems chosen for the novel matrices. An improvement in ionisation, measured by the number of calibration constituents detected, was observed in positive mode for larger proteins, such as those in calibration mix 3 and BSA. Generally, for these particular proteins ionisation in negative mode proved to be quite poor. Positive mode experiments also highlighted the presence of adducts between matrix and some of the calibration mix. For example, Fluram was observed to form adducts with Angiotensin I with the loss of radical hydroxyl group from the adduct when no TFA was present in a solution of 50% acetone. However, upon addition of TFA the adduct formation was removed, suggesting that acidity has an impact on the positive mode ionisation processes involved in the MALDI plume for this particular novel matrix.

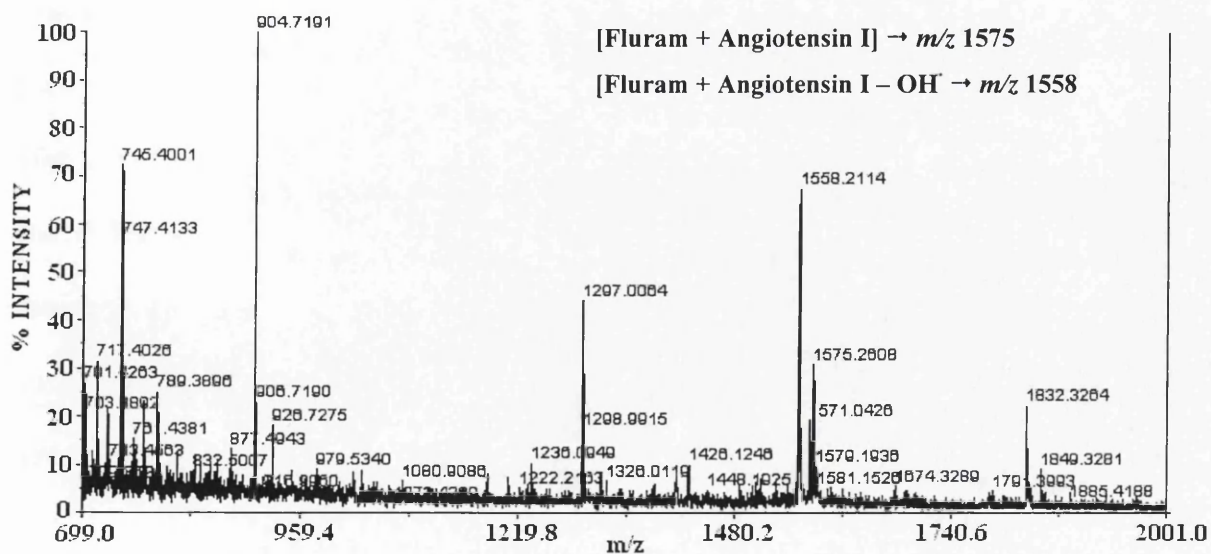


Figure 5.9: MALDI mass spectrum of calibration standard mix 1 (see table 5.3 for contents) in 50% acetone Fluram solution. Observed within this mass spectrum are

*Fluram adducts with Angiotensin I (m/z 1297 no adduct, 1575 with adduct), and a possible corresponding loss of radical hydroxyl group (m/z 1558).*

Conventional negative mode analyses of proteins and peptides were carried out in the same conditions as positive mode, using CHCA and sinapinic acid as described previously. Unlike, positive mode, no enhancement of ionisation was observed upon addition of TFA irrespective of concentration, and no adduct formation was observed. This supports the acid-base principle that for improving ionisation in positive mode a proton-donating acid is used as the additive. This would also suggest that for amplifying ionisation in negative mode, a proton-accepting base should be included in the matrix-sample mix.

#### 5.2.2.2.3. Matrix concentration

This parameter was investigated by testing matrix solutions with varying matrix compositions of 10, 25, and 50mg/mL with the calibration standard mixes and BSA. Throughout these experiments no specific trend was observed regarding matrix performance and its abundance in the sample spot. However, it was apparent that the 10mg/mL concentration consistently performed the most poorly for all the matrices tested. There was also no clear trend regarding the mass range of protein and peptide analysed and the concentration of matrix solution. Thus, for these matrices, concentration appears to have little effect on the ionisation of particular size proteins and peptides both in positive and negative mode.

#### 5.2.2.2.4. Ionisation mode for novel matrices

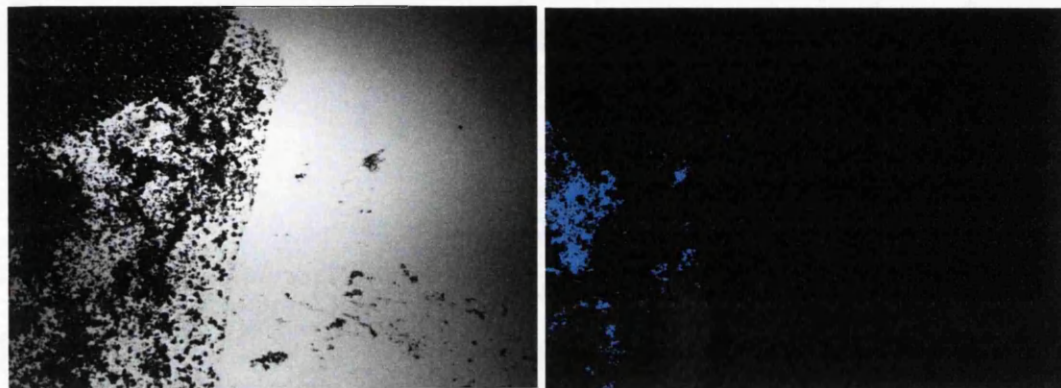
As stated earlier in this section, all of the diagnostic work, regarding optimization of the novel matrices for protein and peptide analysis was carried out in both positive and negative ionisation modes. Proteins and peptides were observed in both modes, although more of the calibration standards were detected in positive mode. This could be the result of two possibilities; the matrices may be more acidic (better suited to proton donating) and/or the proteins and peptides in the calibration mixes are relatively basic, more suited to accepting protons and therefore, function better in positive mode. However, results obtained for negative ionisation mode were acquired using a TFA additive frequently associated with positive mode ionisation.

#### 5.2.2.2.5. Investigation of optimum sample preparation conditions for novel matrices

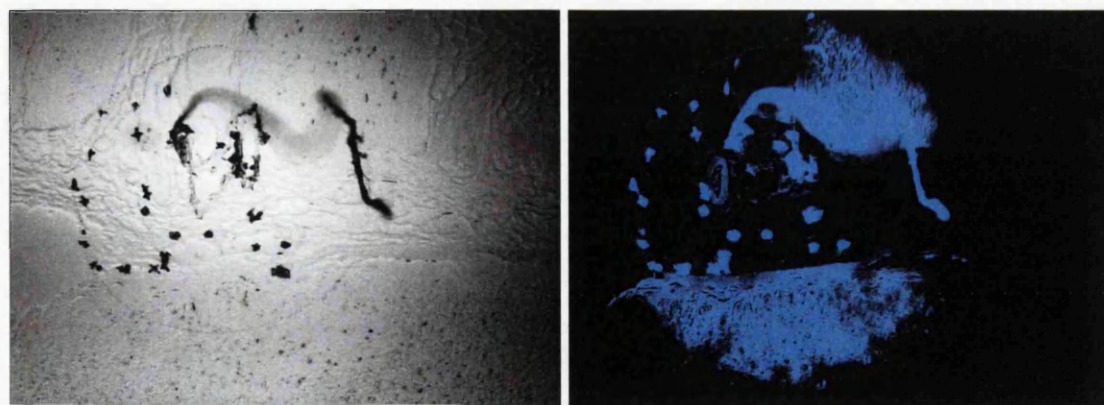
##### 5.2.2.2.5.1. Fluorescent microscopic results

The fluorescent capability of these novel matrices on tissue was investigated and compared to the chosen matrices of  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA), 2,5-dihydroxybenzoic acid (DHB) and sinapinic acid (SA). The chosen matrices showed a varied potential for use as a fluorescent 'tag' to highlight areas prepared with matrix. For example, CHCA showed a good fluorescent signal and good distribution for paraffinised tissue, but distribution is affected, becoming more isolated when deparaffinised. This distribution affect was also observed for DHB although this matrix generally showed a poor fluorescence signal. Sinapinic acid matrix generally performed poorly on both paraffin-embedded and deparaffinised tissue sections. For dansylated matrices, dansylhydrazine showed greater fluorescent signal and distribution over the whole spot for both paraffinised and deparaffinised tissue. These novel compounds, unlike the selected common matrices, have the potential of

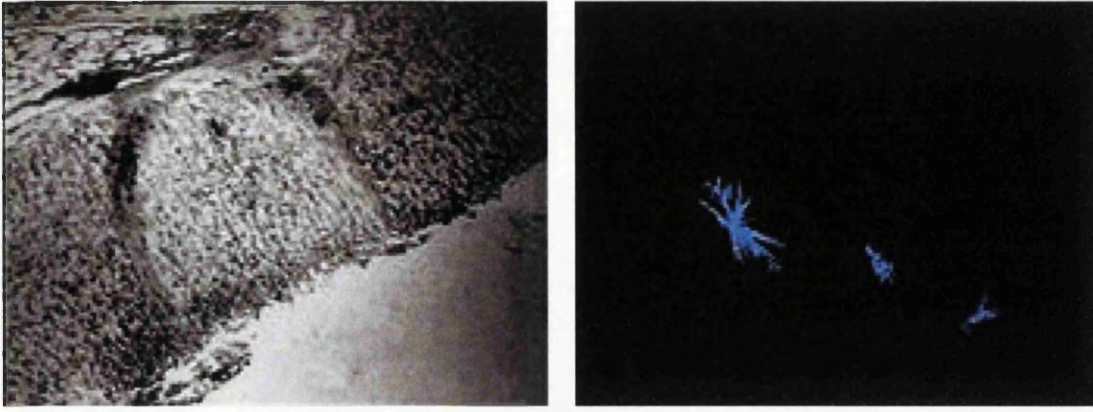
dispersing deeply into the tissue and is an encouraging result for achieving penetration and ionisation of membrane bound proteins within tissue sections.



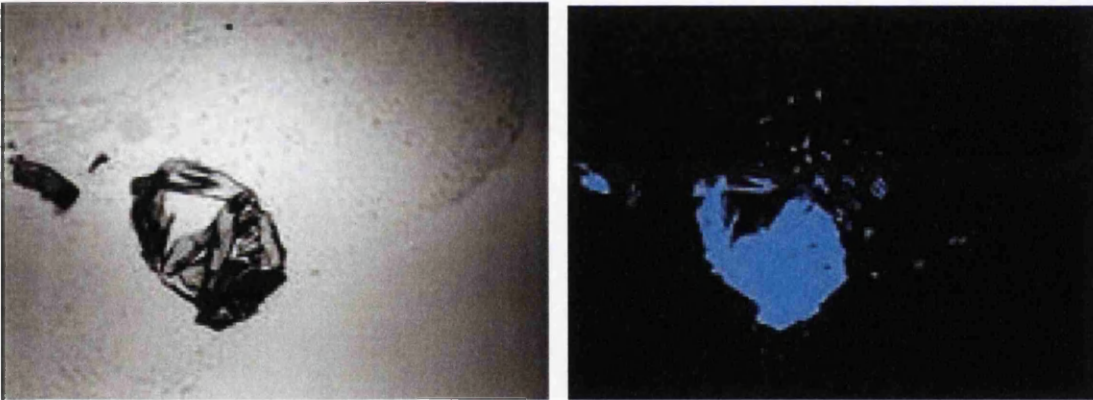
*Figure 5.10: Spot of  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) on deparaffinised tissue observed at visible (left) and fluorescent wavelengths (right).*



*Figure 5.11: Spot of  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) on paraffin-embedded tissue at visible (left) and fluorescent wavelengths (right).*

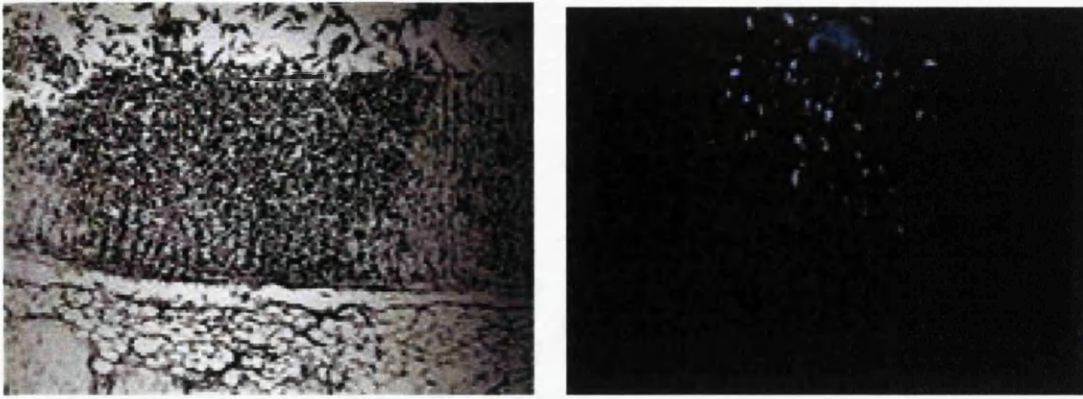


*Figure 5.12: Spot of 2,5-dihydroxybenzoic acid (DHB) on deparaffinised tissue at visible (left) and fluorescent (right) wavelengths.*

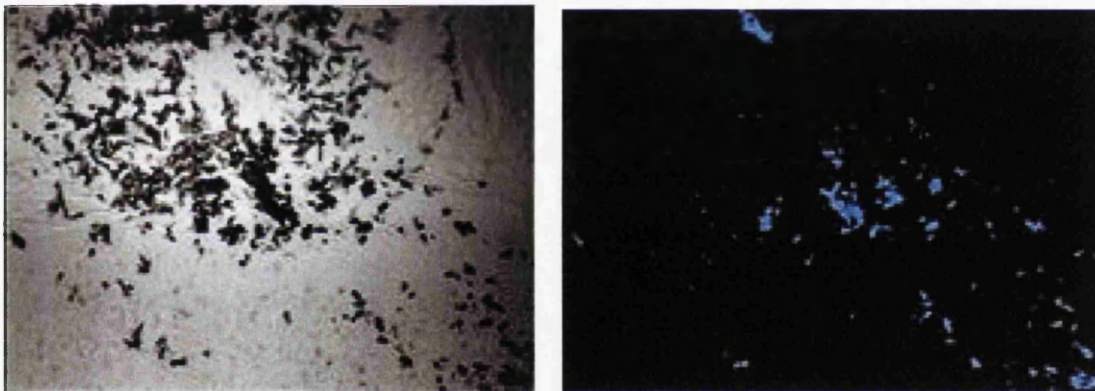


*Figure 5.13: Spot of 2,5-dihydroxybenzoic acid (DHB) on paraffin-embedded tissue at visible (left) and fluorescent (right) wavelengths.*

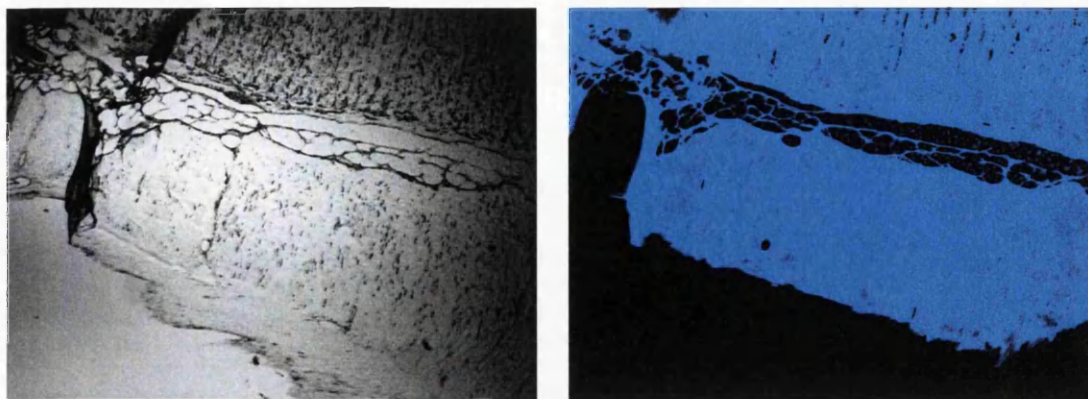




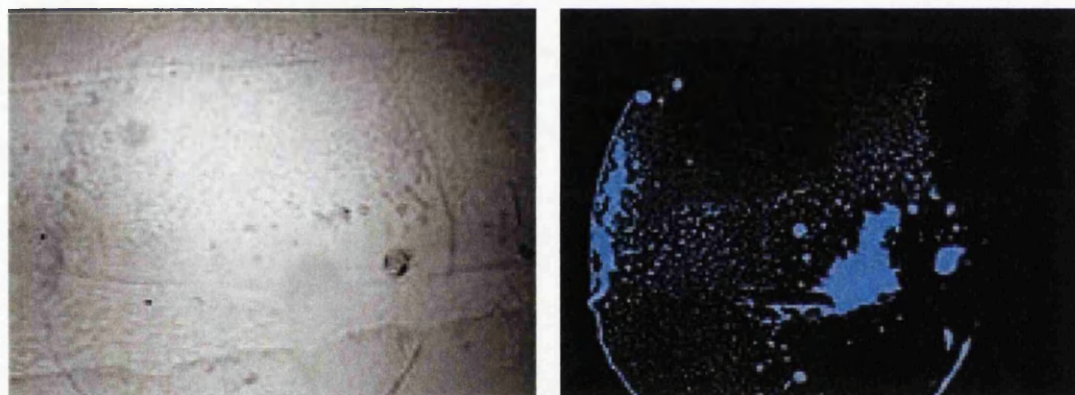
*Figure 5.14: Spot of sinapinic acid matrix on deparaffinised tissue at visible (left) and fluorescent (right) wavelengths.*



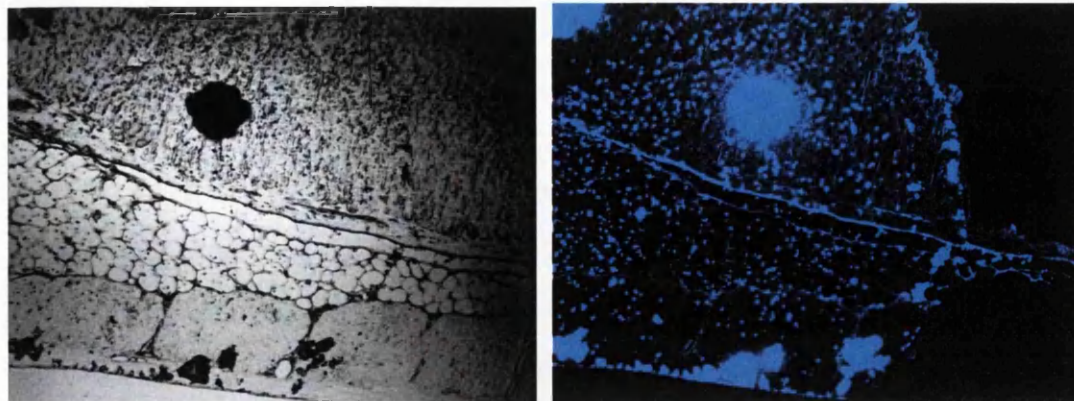
*Figure 5.15: Spot of sinapinic acid matrix on paraffin-embedded tissue observed at visible (left) and fluorescent (right) wavelengths.*



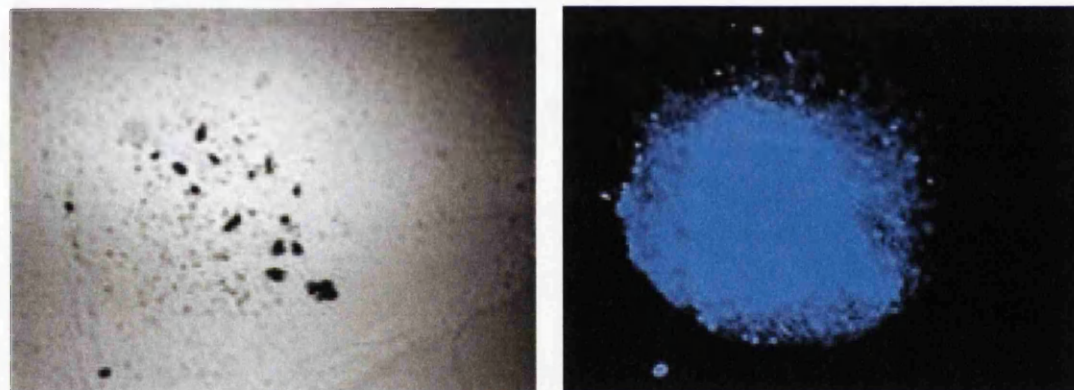
*Figure 5.16: Spot of novel matrix dansyl-DL- $\alpha$ -aminocaprylic acid on a deparaffinised tissue section at visible (left) and fluorescent (right) wavelengths.*



*Figure 5.17: Spot of novel matrix dansyl-DL- $\alpha$ -aminocaprylic acid on paraffin-embedded tissue section at visible (left) and fluorescent (right) wavelengths.*

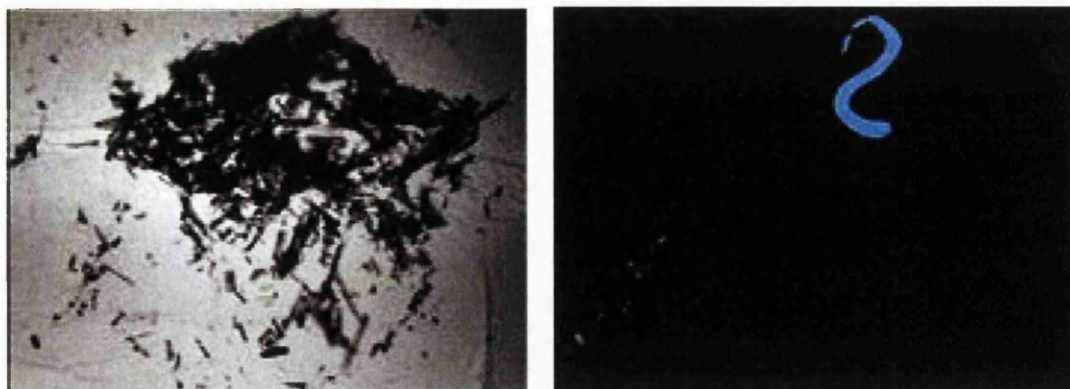


*Figure 5.18: Spot of dansylhydrazine novel matrix on deparaffinised tissue pictured at visible (left) and fluorescent (right) wavelengths.*



*Figure 5.19: Spot of novel matrix dansylhydrazine on paraffin-embedded tissue at visible (left) and fluorescent (right) wavelengths.*





*Figure 5.20: Spot of novel matrix Fluram on deparaffinised tissue at visible (left) and fluorescent (right) wavelengths.*



*Figure 5.21: Spot of novel matrix Fluram on paraffin-embedded tissue at visible (left) and fluorescent (right) wavelengths.*

#### 5.2.2.2.5.2. Mass spectrometric suitability of novel matrices

The suitability of both matrix and their preparative conditions were measured according to:

- a. the number of protein and peptide peaks successfully observed,
- b. minimal laser intensity used (ideally 100 units from threshold i.e. the minimum laser intensity required in achieving a signal) and
- c. the highest signal obtained for proteins and peptides.

Therefore the best performing matrices for specific mass ranges of protein and peptides in solution using positive mode are as shown in table 5.4. It is apparent that a relatively high concentration of acid is required to achieve good ionisation of proteins and peptides in positive mode. This is consistent with the effect observed with some of the conventional matrices such as CHCA and in particular, sinapinic acid. The data also suggests that a larger complement of organic solvent is required for a more effective crystallization and eventual ionisation process of these biomolecules. This is not however, consistent with the conventional matrices named above, as for larger proteins sinapinic acid is typically used with 30% organic solvent mix. Thus, it appears as though this effect may just apply to the novel matrices suggested in table 5.4. One relatively clear finding from these experiments is that dansylhydrazine has the potential to be used for the analysis of a wide range of proteins and peptides unlike the current matrices CHCA and sinapinic acid (molecular weights below and above 10kDa, respectively).

Calibration Mix	Mass Range ( $m/z$ )	Novel Matrix	Concentration (mg/mL)	Solvent Composition
1	900-2000	Fluram	50	30%ACETONE/0.1%TFA
		Fluram	25	30%ACETONE/0.1%TFA
		DCA	25	30%MEOH/0.1%TFA
2	1200-6000	DH	25	30%MEOH
		Fluram	50	30%ACETONE/0.1%TFA
		DH	50	50%MEOH/0.1%TFA
3	5000-20000	DH	50	30%MEOH/0.3%TFA
		DH	50	50%MEOH/0.3%TFA
		DH	25	50%MEOH
4	33000-66000	DH	50	50%MEOH/0.3%TFA
		DCA	25	50%MEOH/0.3%TFA
		DH	50	50%MEOH

Table 5.4: Optimum conditions found for analyzing the proteins and peptides in the Sequazyme calibration standards (Applied Biosystems), where DH = dansylhydrazine, DCA = dansyl-DL- $\alpha$ -aminocaprylic acid. Overall dansylhydrazine appears as a potential 'universal' matrix that may be suitable for a wide mass range.

Current literature suggests that cation/proton adduct-forming matrices and alkaline protein or peptide signals will be prone to a peak broadening effect, thought to be due to the desorption of salt molecules in the MALDI plume<sup>[33]</sup>. Thus, the mass shift associated with the salt adduct appears as part of the signal of the parent ion of the peptide or protein, resulting in an apparent broadening of the peak. Dansyl-DL- $\alpha$ -aminocaprylic acid (DCA) is known to exist as a salt in solution and should therefore have poorer peak resolution than the remaining novel matrices. This mass resolution effect for DCA is particularly apparent when mixed with standard calibration mix 1, showing a mass resolution for ions  $m/z$  905, 1297, 1571 and 1674 of 4408, 4032, 5363 and 4421, respectively. These values are considerably lower than the other dansylated matrices, and dansylhydrazine (DH) has approximately double the mass resolution for

these peaks of 9508, 11622, 9599 and 14505, respectively. This effect is not however, found for the matrix Fluram and could suggest that the ionisation process for this matrix with proteins and peptides is quite different to that of the other novel and conventional matrices. Unfortunately, the ionisation behaviour of Fluram has not been previously been investigated and hence, this trend will remain a hypothesis until further work is completed.

### **5.3. Comparative Studies of Novel and Current Matrices for Protein and Peptide Analysis**

One of the common strategies used in the identification of proteins by MALDI is a tryptic digestion followed by a database search of the peptidic fragments. CHCA, DHB and sinapinic acid are commonly used for these experiments, and the novel matrices that performed best for proteins and peptides less than 6kDa were run in a comparative study to assess their capability in identifying proteins. A range of proteins,  $\alpha$ -casein,  $\beta$ -casein, bovine serum albumin (BSA), cytochrome C, and carbonic anhydrase were digested overnight using the endoprotease trypsin in ammonium carbonate solution (50mM, pH8). The digestion was stopped by adding 0.1%TFA solution and 1 $\mu$ L of each digested protein solution was individually mixed with 7 $\mu$ L of the chosen novel (see table 5.4) and conventional matrices. MALDI mass spectra were recorded and the resulting peptide ions imported into the MS-Fit database (UniProt Knowledgebase, ExPASy, Geneva, Switzerland) where they were searched against a wide range of proteins. The protein identities generated by this software were ranked according to the number of peptides matched for a particular

protein. Of the conventional matrices chosen CHCA proved most successful in identifying the protein correctly. However, both sinapinic acid and DHB generated some false identities for the proteins and were not considered sufficiently reliable for this type of analysis<sup>[34]</sup>. The novel matrices in comparison to CHCA proved unsuccessful in identifying the chosen proteins and were unable to generate more than 25% of the expected peptide ions. Some of the unidentified peaks within the peptide spectra obtained with the novel matrices can be accounted for by adduct formation with matrix. Peaks are observed with the appropriate mass shift making them undetectable by the database since this peptide mass is not present in the protein database. For example, dansylhydrazine can cause the 1695, 1823 and 2181 peaks of carbonic anhydrase to be mass shifted by +263 leading to observed peaks of 1958, 2086 and 2444, respectively. This adduct-type reaction has been stated in the literature, indicating that dansylhydrazine reacts with ketone and aldehyde functional groups causing a mass of +263.07 for every ketone/aldehyde group<sup>[27, 28]</sup>. There is no apparent trend in site specific adduct formation of these tryptic peptides, apart from the c-terminus lysine or arginine residues common to all peptides generated with this enzyme.

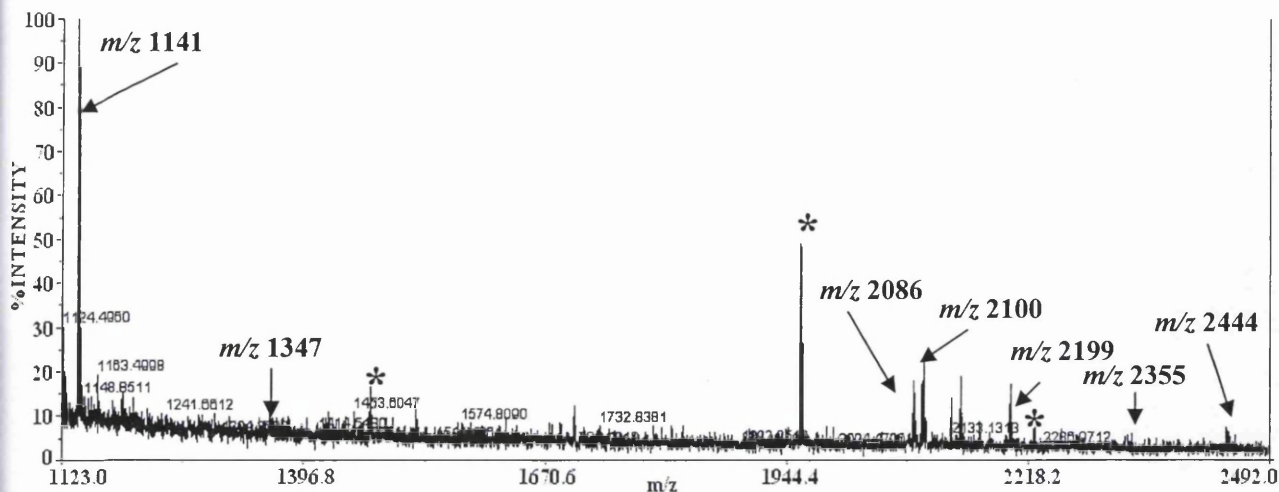


Figure 5.22: Spectrum of carbonic anhydrase digested using trypsin and ionised using the novel matrix dansylhydrazine. The peaks labelled by \* are those peptides that show a mass shift of +263 due to the adduct formation with the matrix at  $m/z$  2231, 1973 and 1466. The original peptide peak of these adducts were at  $m/z$  of 1968 (one missed cleavage), 1710 (one missed cleavage) and 1203 (one missed cleavage) respectively. Also shown in this spectrum are the expected tryptic peptide peaks at  $m/z$  2355, 2199, 2100, 1347 and 1142 for amino acid positions 37-58 (one missed cleavage), 37-57 (no missed cleavages), 59-76 (no missed cleavages), 156-169 (no missed cleavages) and 10-18 (no missed cleavages), respectively.

A similar mass shift however, has previously been observed when using dansylchloride to derivatise tryptic peptides to improve sequence coverage<sup>[29]</sup>. However, this mass shift was of a single dansyl group causing an increase in mass of +233. This disparity is due to the inclusion of two amino groups attached to the sulphur of the dansyl moiety and accounts for the remaining 30Da. If these mass shifts are included with the identified peaks, the novel matrices performed considerably better in identifying proteins. Dansylhydrazine proved most promising

since it showed comparative performance to the conventional DHB matrix (see table 5.5).

Matrix	Number of Matching Peaks	Number of Mass Shifted Peaks
$\alpha$ -cyano-4-hydroxycinnamic acid (CHCA)	11/20	-
2,5-dihydroxybenzoic acid (DHB)	5/20	-
Dansylhydrazine (DH)	5/20	3

*Table 5.5: Number of matching peptide peaks for CHCA, DHB, and the novel matrix DH, of ionised tryptic peptides for carbonic anhydrase II identification. From the data it is clear that dansylhydrazine has the potential to perform better than DHB for identification providing the mass shifted peaks are included in the mass peptide list of the protein database.*

#### 5.4. Application of Novel Matrices onto Tissue Surfaces

The viability of tissue imaging using the three selected novel matrices (see table 5.4), CHCA, sinapinic acid and DHB were tested with paraffin-embedded, de-paraffinised, and frozen tissue sections. The paraffin and de-paraffinised sections originated from rat liver, and the frozen sections were obtained from porcine kidney. The sections were cut on a microtome and OTF-5000 cryostat (Bright Instrument Co. Ltd., Cambridge, UK) respectively and were 10 $\mu$ m in thickness.

##### 5.4.1. Preparation Protocol for Tissue Sections

###### 5.4.1.1. Paraffin sections

A minimum of nine 10 $\mu$ m sections were generated using a microtome and placed directly from the blade to the stainless steel MALDI target plate. Both novel (as

shown in table 5.4) and conventional matrices were applied using the ‘dried droplet’ procedure ready for subsequent analyses.

#### 5.4.1.2. De-paraffinised sections

The paraffin-embedded tissue sections were prepared using the following protocol<sup>[35]</sup>:

1. incubate sections for 5 minutes in two washes of xylene solvent,
2. apply three ethanol washes of 100, 95, 80, 50% for three minutes each and
3. rinse twice with distilled water for three minutes.

After this is completed matrix is applied at a particular region of the tissue using the ‘dried droplet’ method and each matrix having a separate but subsequent tissue section. Once the matrices have dried the plate is then submitted for analysis using the Voyager DE-STR MALDI-ToF (Applied Biosystems) mass spectrometer.

#### 5.4.1.3. Frozen sections

Fresh porcine kidney was frozen using a 2-methylbutane dry ice mix and stored in a -80°C freezer. Sections were obtained using an OTF-5000 cryotome set at -27°C from tissue embedded in Shandon cryomatrix (Thermo Fisher Scientific, Waltham, MA, USA) and carefully placed onto a stainless steel MALDI target plate. Initially a matrix solution was applied using the ‘dried droplet’ method to a particular region of the tissue and each matrix spotted on separate but subsequent sections. Mass spectral data was obtained per sample spot of the target plate and, CHCA and sinapinic acid were capable of successfully ionizing, in positive mode, biomolecules believed to be protein or peptide, unlike the matrix DHB.



## 5.4.2. Mass spectrometric results

### 5.4.2.1. Frozen tissue sections

#### 5.4.2.1.1. Dried droplet method

CHCA and sinapinic acid were capable of successfully ionising, in positive mode, protein or peptide within this section, unlike the matrix DHB. CHCA was observed to ionise a greater number of biomolecules showing stronger MALDI spectra, although sinapinic acid could obtain ion signals with much improved observed peak resolution. This was later refuted on closer inspection, as the calculated resolution of peaks using the sinapinic acid matrix was poorer than CHCA due to a peak splitting effect. Sinapinic acid does have an added advantage of functioning well in negative mode with comparable results. Analyses of the section carried out in positive mode with CHCA and sinapinic acid matrices showed a number of ions and their corresponding mass-to-charge are displayed in table 5.6.

Several parameters were tested with the aim of improving ionisation of the frozen tissue peptides and proteins, including layering novel matrices with and without conventional matrices, varying drying temperature (hence crystallization time), and using antigen-retrieval buffers (Novagen, Merck KGaA, Darmstadt, Germany) for archived tissue sections.

<i>m/z</i> of Unknown Ion	Matrix		Possible Identity
	CHCA	Sinapinic Acid	
3798	✓		Unknown
4160	✓		Unknown
4958/4966	✓		Unknown
5463		✓	Unknown
5614/5622	✓		Unknown
6003/5988	✓		Tubulin-specific chaperone d (fragment)*
6291/6227	✓		Unknown
7519/7522	✓		Unknown
8021		✓	Unknown
8300		✓	Ferritin light chain (fragment)
8447/8455	✓		Unknown
9048	✓		Unknown
9713/9728	✓		Unknown
10068		✓	Unknown
10087		✓	Unknown
11702		✓	Unknown
12448		✓	Unknown
13802	✓		Unknown
14270	✓		Fatty acid binding protein (liver)
15052	✓	✓	Cytochrome b5,
15804/16044	✓	✓	Cellular retinol binding protein 1 / Superoxide dismutase [Cu-Zn]
29801		✓	Unknown

Table 5.6: List of typical ions detected using the conventional matrices CHCA and sinapinic acid on frozen tissue section of porcine kidney. Identities were assigned using the SwissProt protein database, with known proteins and peptides retrieved from characterization of pig kidney tissue<sup>[36]</sup>. \* peptides of corresponding mass-to-charge and of similar origin to those published<sup>[36]</sup>.

#### 5.4.2.1.2. Matrix layering

This was initially investigated to improve matrix coverage and peptide/protein ionisation by applying multiple layers of novel matrix on the tissue surface after each

has dried. The results obtained using the 'dried droplet' method for just one layer of novel matrix were relatively poor in ionising tissue protein or peptides. Hence, it was considered that the matrix was becoming too dilute as a result of dispersion within the tissue section. It is believed that ionisation is best achieved in a molar excess of matrix, and this effective dilution would therefore hinder ionisation. If this dispersion does occur it is likely that the efficiency of the laser imparting energy to the matrix will be affected. The layering of several applications for the novel matrices was investigated using the 'dried droplet' method. The use of layers did improve the overall ionisation of the tissue and lowered the laser intensity required for analysis in positive mode. However, as the number of layers increases the number of peaks observed in negative mode surpasses those found in positive mode. Exact reasons for this remain unknown although, it is proposed that the density of matrix crystals, the functional groups of the matrices and the unknown aspects of the ionisation mechanisms involved in the MALDI process possibly all play a part.

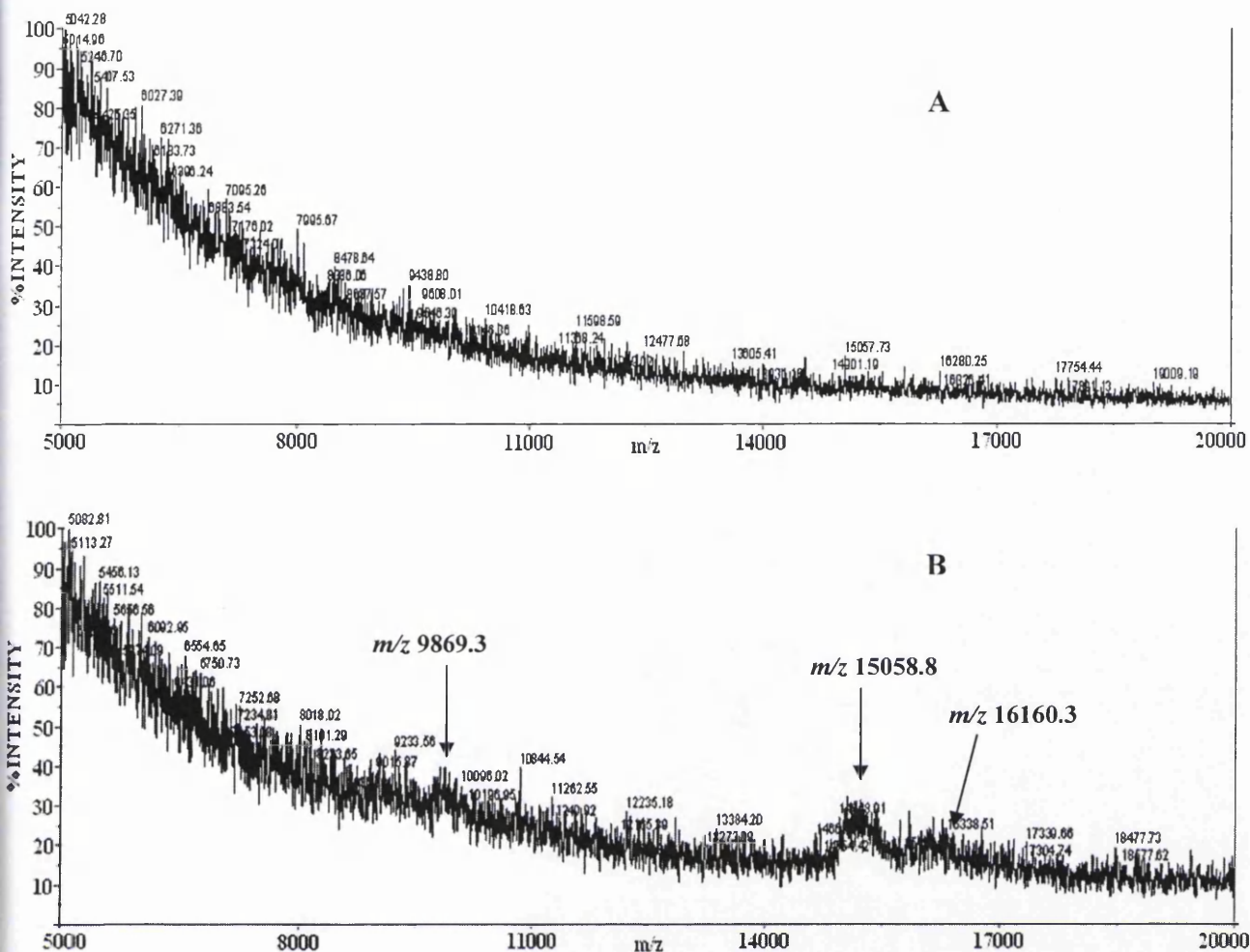


Figure 5.23: Effect of layering novel matrix dansylhydrazine using the 'dried droplet' method on fresh porcine kidney tissue. A - positive mode mass spectrum using one layer of dansylhydrazine matrix, B - positive mode mass spectrum involving triple layers of dansylhydrazine matrix on fresh tissue section. Ionisation of the tissue section seems to improve when using several (three) layers of matrix deposited upon each other. This effect is also observed with the other chosen novel matrices when run in negative mode.

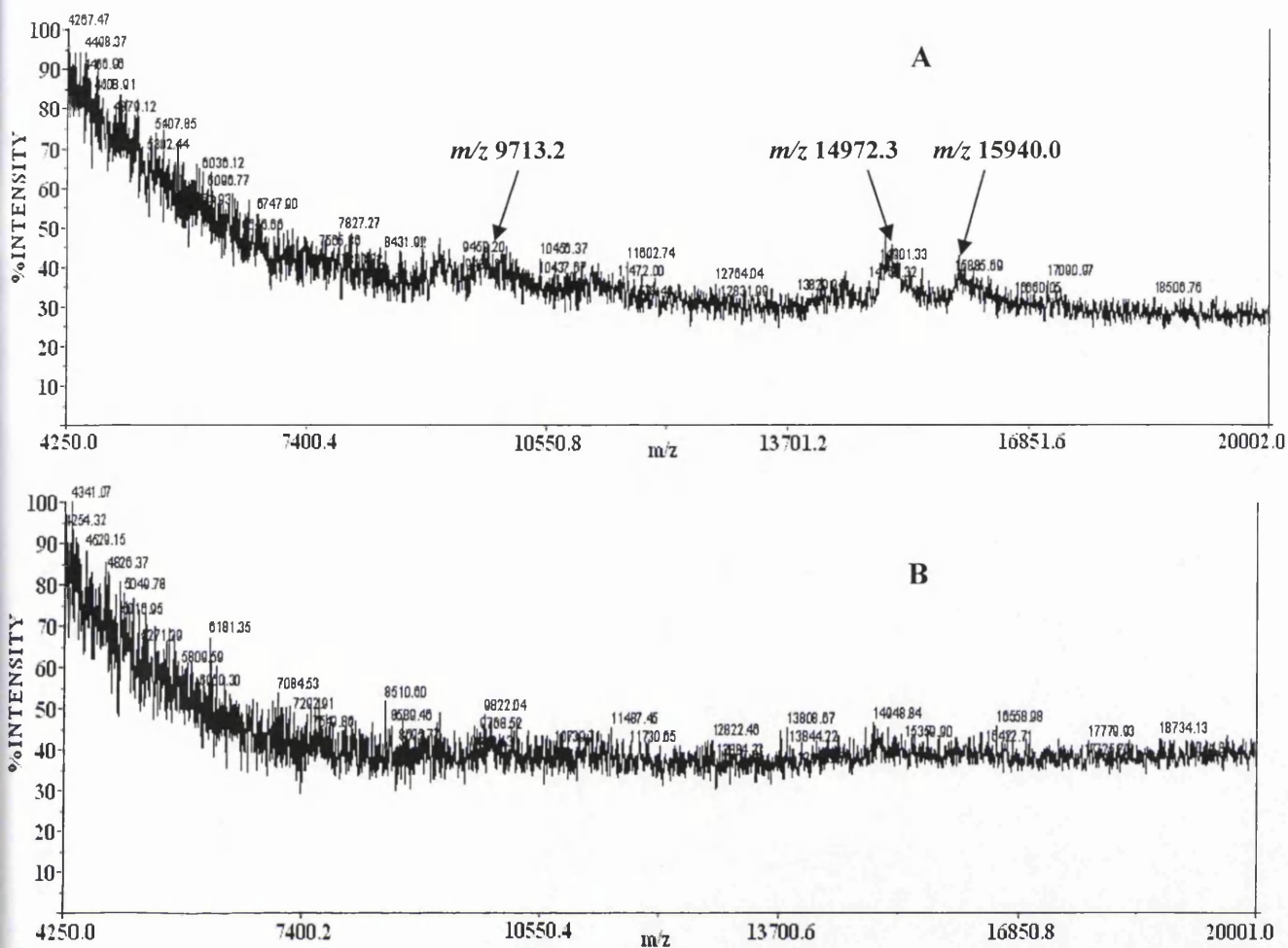


Figure 5.24: Mass spectra of "triple-layered" Dansylhydrazine on fresh tissue section of porcine kidney analysed in negative mode (A) and positive mode (B). It is apparent that more components of the tissue can be analysed in negative mode. The mechanism responsible for this is likely to be related to the ionisation process of the novel matrix, dansylhydrazine.

The conventional matrices, CHCA and sinapinic acid, appeared to ionise fresh tissue well, and more than 15 discernable peaks were observed at relatively low laser intensities. Applying several layers of matrix had the benefit of lowering the background signal observed with high laser intensities, thus improving the sensitivity. This did, however, reduce the peak resolution and would therefore result in an increase in the mass error of the measured ion. A single layer of these conventional

matrices did show a higher number of peaks, although this may be due to the higher laser power required to obtain a spectrum of sufficient intensity.

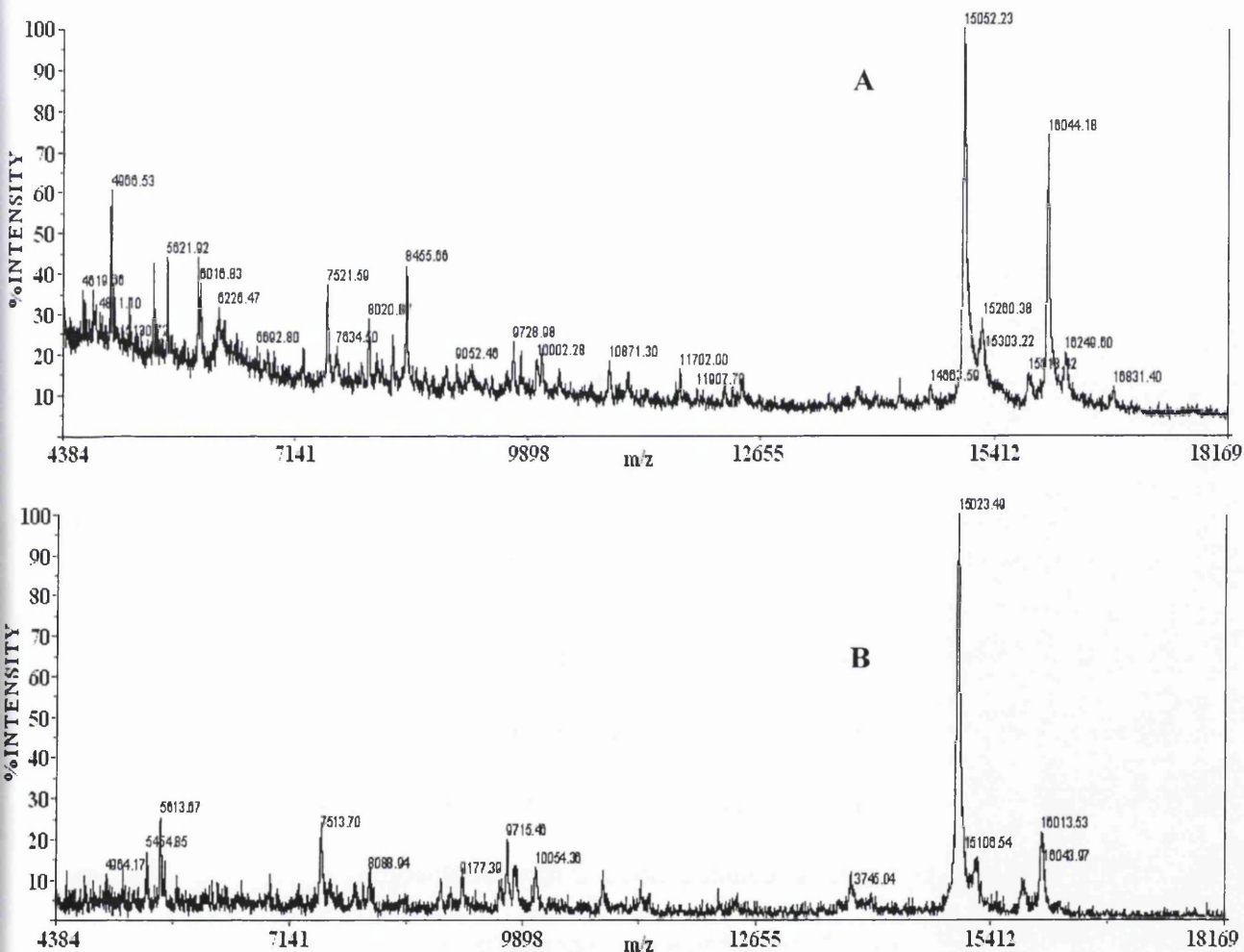


Figure 5.25: Mass spectra showing the effect of multiple surface layers of conventional matrix sinapinic acid on the ionisation of fresh tissue. A - tissue with one layer of matrix with many more peaks observed but with a higher background signal using a greater laser intensity. B - tissue with three layers of sinapinic acid matrix showing less peaks but with a lower background signal.

#### 5.4.2.1.3. Drying temperature effects on fresh tissue sections

This was investigated as it is likely to contribute to the crystallization process of the MALDI matrices. The effect on crystallization was tested using room temperature (25°C), and heating the MALDI target plate (with spotted tissue section) to 50°C using a heating mantle. When applied to one layer of matrix there was no improvement in ionisation observed for the novel matrices, for analysis in both positive and negative modes. An increase in temperature had an adverse effect on the conventional matrix sinapinic acid. At room temperature 14 peaks were present in the positive ionisation mode MALDI spectrum, and this decreased to 9 when heated to 50°C. This indicates the importance of sample preparation on MIMS tissue imaging, and for future experiments all subsequent preparations were carried out at room temperature. Analyses in negative ionisation mode showed no improvement in ionisation, when compared to results in positive mode. We have shown previously that multiple layers of matrix increased the ionisation of tissue in positive mode and tested this in combination with heating for any further improvement. At 50°C ionisation in positive mode with the novel matrices was poor when compared to room temperature. This is comparative with the conventional matrices perhaps indicating that a slower rate of matrix crystallisation is more preferable for tissue ionisation.

#### 5.4.2.1.4. 'Sandwich' preparative method using conventional matrices

This methodology was investigated as multiple layers of matrix were found to significantly improve ionisation and it could combine good ionisation and tissue penetration. It involved depositing matrix, either CHCA or sinapinic acid on the target MALDI plate prior to tissue application. Once the frozen tissue section is

placed on the plate CHCA, sinapinic acid, DHB, or the chosen novel matrices were applied to the tissue using the 'dried droplet' protocol.

The combination of the conventional and novel matrices prepared according to the 'sandwich' protocol resulted in an improvement in the ionisation of the tissue and mass resolution when compared to using the 'dried droplet' protocol and the novel or chosen matrices alone. The mass spectra obtained with matrix dansylhydrazine and CHCA consisted of a greater number of peaks than using CHCA alone. This is possibly due to adduct formation of the tissue biomolecules with dansylhydrazine, as this effect has previously been observed during the tryptic peptide experiments (section 5.3). Adduct formation may also explain the disparity of peak masses observed with experiments carried out with sinapinic acid alone and with dansylhydrazine or Fluram. However, this apparent effect seems to occur with specific biomolecules as there are only a few additional peaks present in the mass spectra involving CHCA and sinapinic acid. For example, within the mass spectrum obtained using Fluram and sinapinic acid a peak is observed at  $m/z$  10085 and is absent when using sinapinic acid only. Peaks of similar mass-to-charge are present in both mass spectra at an approximate  $m/z$  of 9600 and 15000. This could imply that this unique peak of  $m/z$  10085 is an adduct of the ion at  $m/z$  9600, possibly involving the addition of two Fluram moieties.

According to the literature the novel matrix DCA should result in ion peaks with a lower resolution value than the remaining matrices as it is a salt<sup>[33]</sup>. However, unlike the results described in section 5.2.2.2.5, preparations involving DCA actually showed better peak resolution than the other novel and conventional matrices (figures 5.26-5.28). This was not the only resolution effect observed. Peaks obtained with sinapinic acid in both positive and negative mode appear to have significantly better



resolution when compared to those obtained with CHCA (figures 5.29-5.30). However, when calculated at 50% peak height, resolution is actually very poor for the sinapinic acid preparations. This may be due to a peak splitting affect which is not observed with the other conventional matrices or with the novel matrices when used alone. This is apparent when comparing the peaks present with all matrix preparations, at approximately  $m/z$  15000 and 16000. Poor peak resolution is a fundamental problem for determining the identity of biomolecules for use as biomarkers. In order for accurate mass or tryptic peptide database searching to be successful good resolution is required to obtain the true mass-to-charge of the ion. Therefore, if sinapinic acid results in poor peak resolution of the biomolecular ions, this matrix can not be used for either of the functions stated above for tissue analysis.

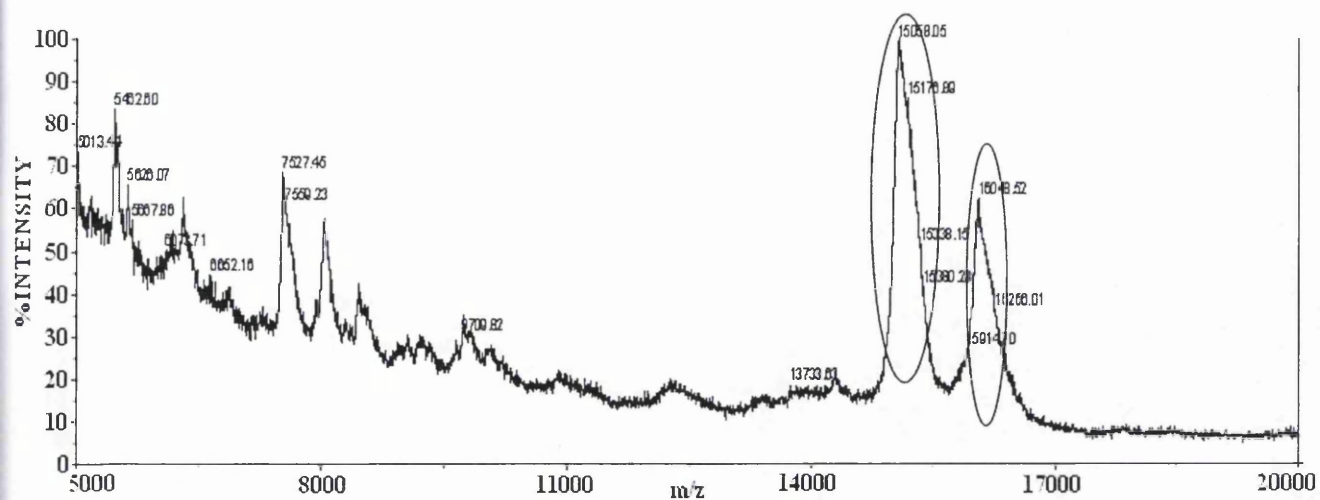


Figure 5.26: Mass spectrum of frozen tissue section using DCA in 30:70 methanol/0.1% TFA and a layer of CHCA showing a number of peaks including those at  $m/z$  15000 and 16000 (circled).

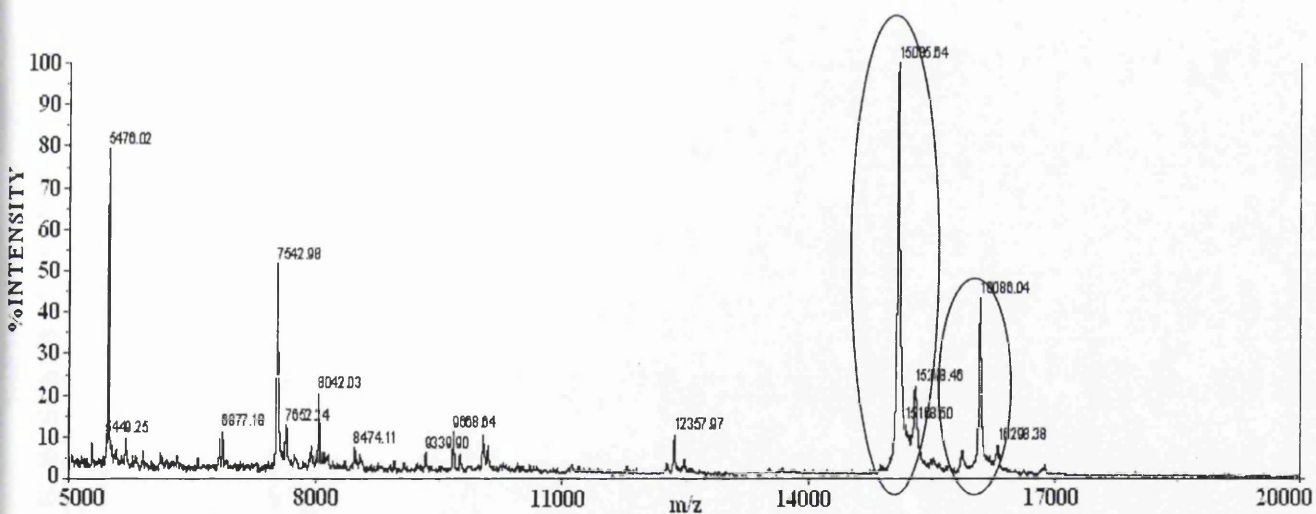


Figure 5.27: Mass spectrum of frozen tissue section using DCA in 30:70 methanol/0.1% TFA and a layer of sinapinic acid showing the peaks at  $m/z$  15000 and 16000 (circled).

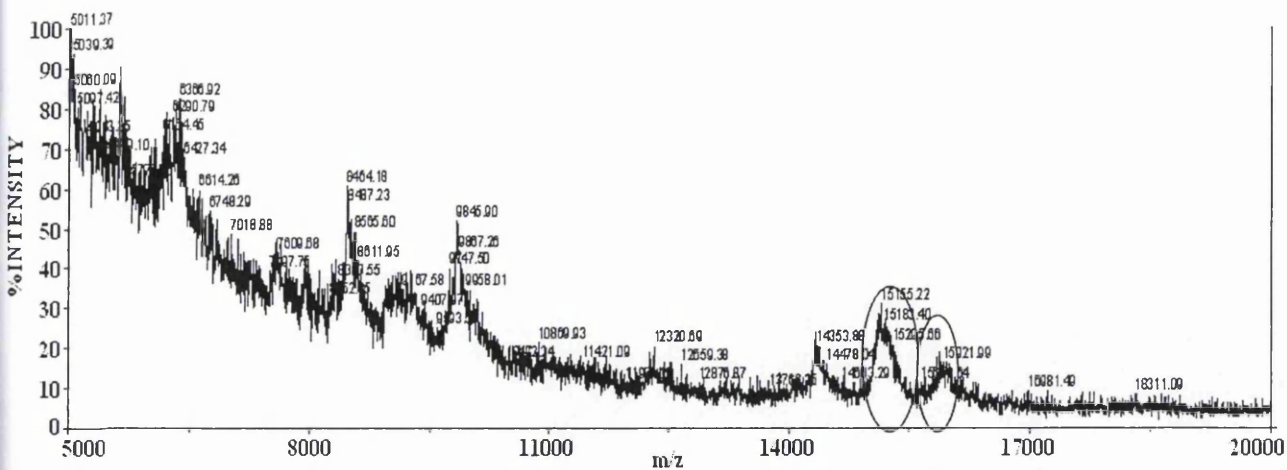


Figure 5.28: Mass spectrum of frozen tissue section using DH in 30:70 methanol/0.1%TFA and a layer of CHCA. This matrix unexpectedly showed much poorer peak resolution illustrated by peaks at  $m/z$  15000 and 16000 (circled) when compared to DCA.

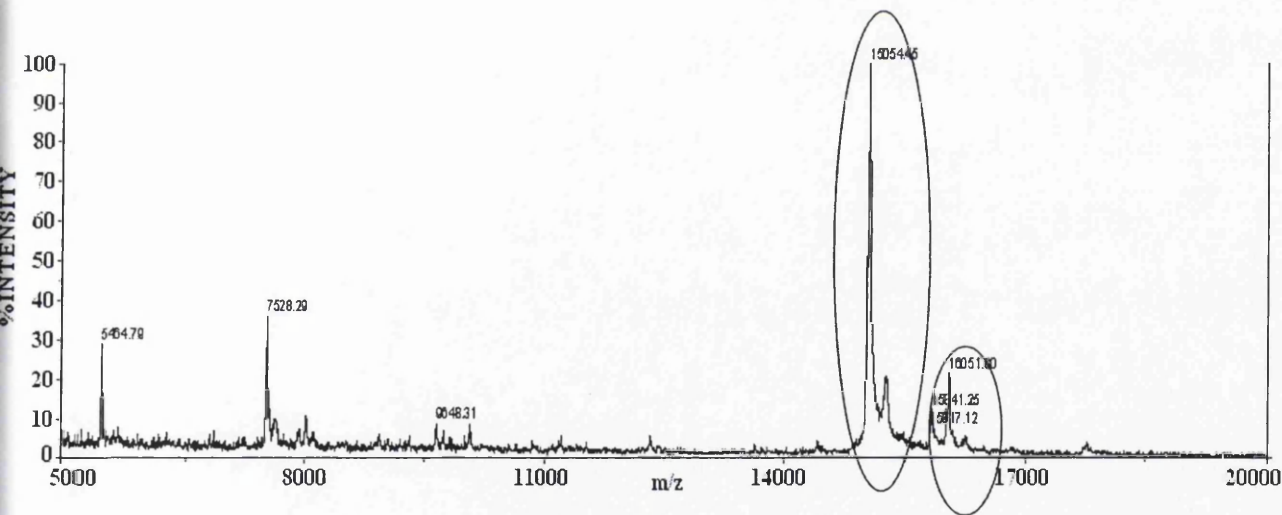


Figure 5.29: Mass spectrum of a frozen tissue section prepared using the 'sandwich' preparative method with sinapinic acid matrix. This has good observed resolution but the calculated resolution is in fact very poor (see circled peaks).

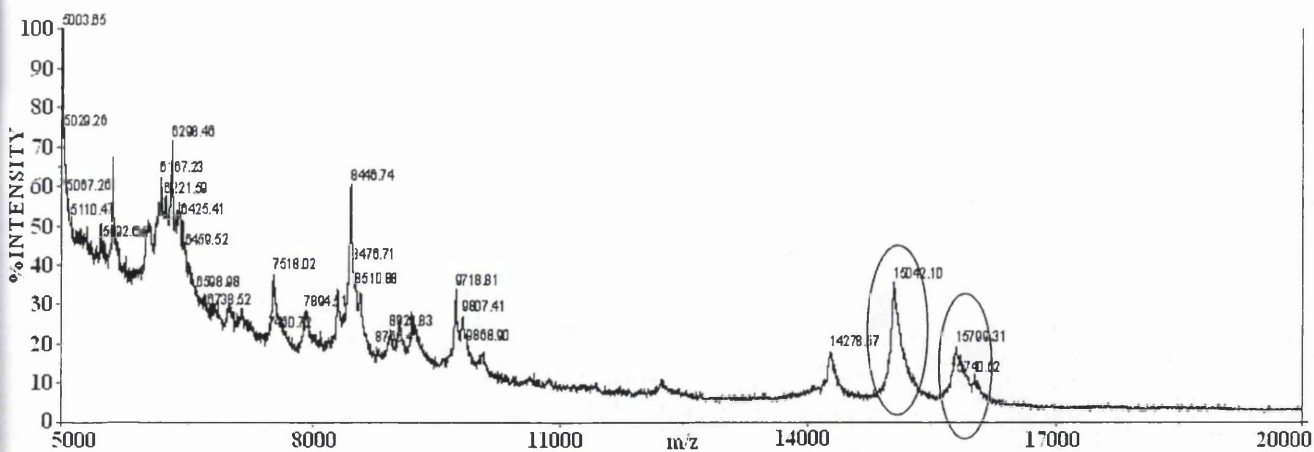


Figure 5.30: Mass spectrum of frozen tissue section prepared using the 'sandwich' method with the conventional matrix CHCA.

#### 5.4.2.2. Paraffinised and de-paraffinised tissue sections

Archived tissue sections that have been in contact with paraffin generally give poor MALDI mass spectra when prepared with the conventional matrices. Almost all standard tissue biopsies are contained within paraffin due to their ease of storage and high degree of stability over long periods of time. Hence, a procedure capable of screening and imaging these sections would be of great use in tracking disease states. A possible effect of paraffin on tissue ionisation is the reduction of the number of potential ionic groups that can result in a mass-to-charge signal. Buffer solutions have been developed to 'activate' a greater number of reactive groups of proteins and peptides within deparaffinised tissue during immunohistochemical staining to enhance their signal. The capabilities of this procedure were therefore investigated for improving the ionisation of these tissue sections with MALDI mass spectrometry.

#### 5.4.2.2.1. “5X” AntigenPlus retrieval buffers

A selection of buffers, obtained from Novagen (Merck KGaA, Darmstadt, Germany), designed for antigen retrieval of tissue for were chosen for investigating additional preparative methods for MIMS. These buffers are intended to improve the sensitivity of histochemical staining and subsequent immunological analysis, by ‘activating’ functionalities of the *in-situ* proteins and peptides. The buffers are available at three different pH values of 6, 7.4 and 10. The neutral buffer is capable of functioning at room temperature, and is designed for more fragile tissue sections. The preparation involving the remaining two buffers include a period of heating to enhance the number of functionalities ‘activated’ and the intensity of the staining process.

##### 5.4.2.2.1.1. Preparation of pH 6 and 10 antigen retrieval buffer

The archive tissue section is placed on the MALDI target plate and deparaffinised as in section 5.4.1.2. The antigen retrieval buffer is diluted 1:5 with deionised water and heated until boiling. The MALDI plate and deparaffinised section is placed in this buffer and heated for 20 minutes to activate the functional groups of tissue proteins and peptides. Following this preparation the plate is removed from the buffer solution and left to dry. The novel and conventional matrices are then applied to the tissue using the ‘dried droplet’ method ready for analysis.

##### 5.4.2.2.1.2. Preparation of pH 7.4 antigen retrieval buffer

A deparaffinised tissue section (see 5.4.1.2.) is placed on a MALDI plate and soaked in the antigen retrieval buffer at room temperature for 10 minutes. The plate is retrieved, dried and matrix applied as in the ‘dried droplet’ method for subsequent analysis.

#### 5.4.2.2.1.3. Results

The neutral pH 7.4 buffer offered little improvement in the ionisation of the deparaffinised tissue regardless of matrix used. It is understood that acid-base chemistry has a fundamental role in ionisation within the MALDI plume. The pH 6 and 10 buffers should promote conditions favourable for ionisation through the dissociation of ionic bonds. Therefore we can hypothesise that these buffers would produce a greater number of peaks within the MALDI mass spectra when compared to those at neutral pH. If this acid-base hypothesis is correct the pH 6 and 10 buffers could have the potential to specifically enhance the ionisation of tissue in positive and negative modes respectively. An additional benefit is that the preparation for these buffers involves a considerable period of heating prior to MALDI analysis. This can denature proteins or peptides present in the tissue section and prime or 'activate' them for ionisation as a greater number of potential ionic sites will be exposed. However, in practice the inclusion of the pH 10 buffer showed little improvement of the ionisation of tissue when analysed in both positive and negative modes. Hence, there must be an additional factor other than the acid-base chemistry within the MALDI plume that is essential for successful ionisation biomolecules within a deparaffinised tissue section.

#### 5.4.2.2.2. Comparison of immunohistochemical testing and mass spectrometric data

Current techniques used in the screening of fresh or archived tissue for disease states and biologically relevant molecules, such as over/under-expressed proteins, involve immunohistochemical staining. This technique comprises of testing for chosen 'antigens' or proteins using appropriate antibodies and is followed by the application of a staining solution which will preferentially bind to the antigen-antibody complex.

Deparaffinised tissue sections from various human organs were tested by immunohistochemical staining for particular protein(s) and then compared to a subsequently cut section analysed by MALDI mass spectrometry. Matrices used here included sinapinic acid, CHCA and dansylhydrazine in conditions shown for calibration mix 3 and BSA standard (see table 5.4). These are seen to perform better than other conventional matrices with proteins and peptides in solution and on tissue. Tissue sections analysed and the relevant 'antigen' protein are shown in the table 5.7.

Tissue Section	Proteins Tested for with Immunohistochemistry	Molecular Weight (Da)
Appendix	CD-45 (precursor)	100k
Tonsil	CD-45 (precursor)	100k
Breast	Progesterone receptor (precursor)	Membrane associated component 1
		Membrane associated component 2
Thyroid	Calcitonin (precursor)	15.4k
Melanoma	S100 (precursor)	β-chain
		γ-chain
		p-chain
		z-chain

Table 5.7: Proteins that tested positive for the tissue sections analysed using the chosen novel and current imaging matrices, and their approximate molecular weight.

5.4.2.2.2.1. Results: testing for CD-45 protein

The regions of appendix and tonsil tissue that stained positive for the protein CD-45 (precursor) did not present ions in the MALDI mass spectrum at the expected mass of 100kDa whether run in positive or negative modes. It is widely accepted that there

are difficulties in obtaining mass spectra large biomolecules and in particular proteins. Exact reasons for this are unclear however a possibility could be the degradation of the protein into smaller subunits during the laser ablation or ionisation process. This is consistent with a greater number of peaks observed at a mass range of 2-20kDa in positive ionisation mode for the tissue that produced a positive immunohistochemical stain for CD-45. To test this hypothesis a predicted fragmentation pattern was generated using BioLynx software (Waters, Milford, MA, USA) and compared to the relevant mass spectra. Unfortunately, of the peaks present none matched the expected fragment peptides perhaps indicating a rearrangement of fragments or an unexpected fragmentation process. However, as expected, the use of the pH 6 buffer resulted in improved ionisation of the tissue section in positive mode when compared to that in negative mode.

#### 5.4.2.2.2. Results: Progesterone Receptor (precursor)

This is actually present as two subunits of masses 21.7 and 23.8kDa. Unlike other regions of tissue, areas that tested positive immunohistochemically for this protein showed a peak present at a mass of 28.7kDa and several other masses below 15kDa. These results were acquired using CHCA or sinapinic acid matrices in positive and negative ionisation modes whilst, no discernable peaks were observed with dansylhydrazine. A closer inspection of some related proteins indicate the high mass peak at 28.7kDa could correspond to a 255 amino acid fragment of the intact progesterone receptor protein and is believed to have a role in the regulation of DNA transcription. It is possible that the smaller masses are degradative products of the membrane associated components 1 and 2 or perhaps the total protein, progesterone receptor. Expected fragmentation patterns of these proteins were generated using



BioLynx software (Waters) and compared against the mass spectra of the breast tissue. This indicated that the peaks previously unassigned could be fragments of membrane component 1 ( $m/z$  6317, 11347, 14009) and the whole progesterone receptor protein ( $m/z$  6980) as shown in table 5.8.

<b>Ion (<math>m/z</math>)</b>	<b>Possible Origin of Fragment</b>	<b>Predicted Fragment (<math>m/z</math>)</b>	<b>Amino Acid Residue Fragment</b>
6320, 6316	Membrane component 1	6317	1-69
6981, 6978	Progesterone receptor	6980	1-64
11347, 11340	Membrane component 1	11347	1-114
14012, 14003	Membrane component 1	14009	1-138

*Table 5.8: Putative origin of peaks present in spectra of breast tissue obtained using CHCA and sinapinic acid matrices.*

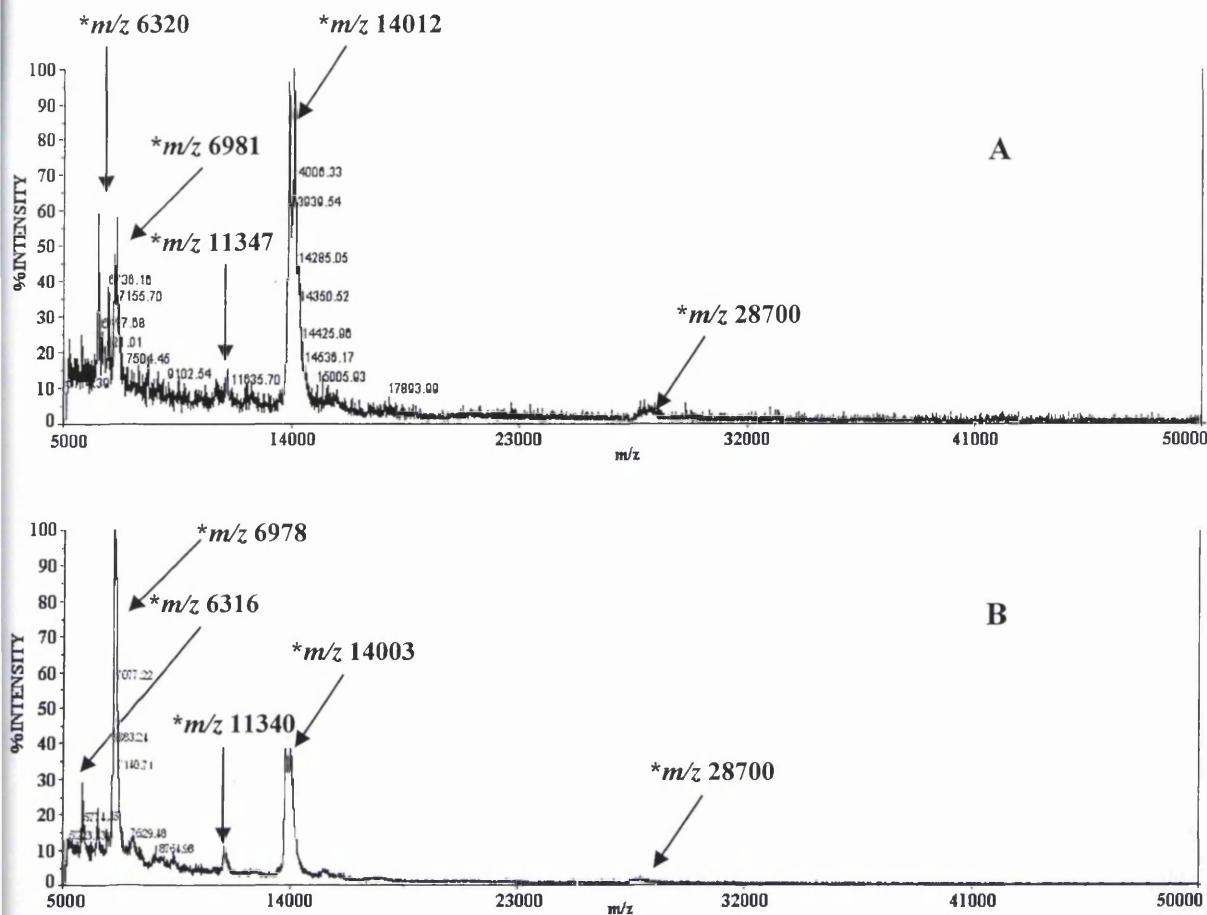


Figure 5.31: MALDI mass spectra of a region of human breast tissue that tested positive for the protein progesterone receptor (precursor) obtained using matrices CHCA (A) and sinapinic acid (B). An ion specific to these relevant regions of tissue is at approximately 28.7kDa. This is not consistent with the values in table 5.7 but could be due to the presence of intact progesterone receptor protein. Also observed are several peaks below the mass of this protein and the membrane components that correspond to the fragmentation of progesterone membrane component 1 and progesterone receptor (\*) at m/z 6317, 11347, 14009 and m/z 6980, respectively.

#### 5.4.2.2.2.3. Results: Calcitonin (precursor)

This 15.4kDa protein was identified in specific regions of thyroid tissue by immunohistochemical means. A subsequent tissue section was prepared using the antigen retrieval buffers with CHCA, sinapinic acid and dansylhydrazine matrices, then analysed using MALDI mass spectrometry in both positive and negative modes.

Unlike CHCA, sinapinic acid and dansylhydrazine were unsuccessful in generating a possible molecular ion of this protein at  $m/z$  15,400 in both positive and negative modes. In negative mode however, this ion was absent from the CHCA mass spectrum and replaced by a lower mass ion at  $m/z$  15,100. This could imply that a loss of an amino acid(s), possibly basic in nature, to expose a readily ionisable carboxyl group has occurred. Again, dansylhydrazine was unable to generate any peaks within the MALDI mass spectrum when operated in both positive and negative modes. However, a peak consistent with both CHCA and sinapinic acid matrices is at  $m/z$  13,909 and is indicative of Calcitonin gene-related peptide 1 (precursor). This is capable of inducing vasodilation in several different vessels, contribute to neurosignalling within the central nervous system and elevate levels of platelet cAMP<sup>[37,38]</sup>.

Sinapinic acid mass spectra were associated with a greater number of ions in both positive and negative modes and some were consistent with CHCA prepared tissue. The remaining peaks could be due to the degradation of this protein into fragment peptide ions when ionised with the 'hotter' sinapinic acid matrix. To test this theory the expected fragmentation pattern of both Calcitonin proteins were generated using BioLynx software (Waters) and compared to the relevant MALDI mass spectrum for a corresponding peak. Using this method we can suggest several low intensity peaks that are present in both the CHCA and sinapinic acid mass spectra could be as a result of Calcitonin (precursor) fragmentation.

<i>m/z</i> of Peak	<i>m/z</i> of Predicted Fragment	Amino Acid Residue Fragment
9320	9319	1-88
7705	7702	1-71
5670	5666	1-54
4713	4715	1-45
3720	3718	1-35

Table 5.9: Masses present in the MALDI mass spectrum of prepared thyroid tissue that correspond to specific fragmentation ions of the protein Calcitonin (precursor).

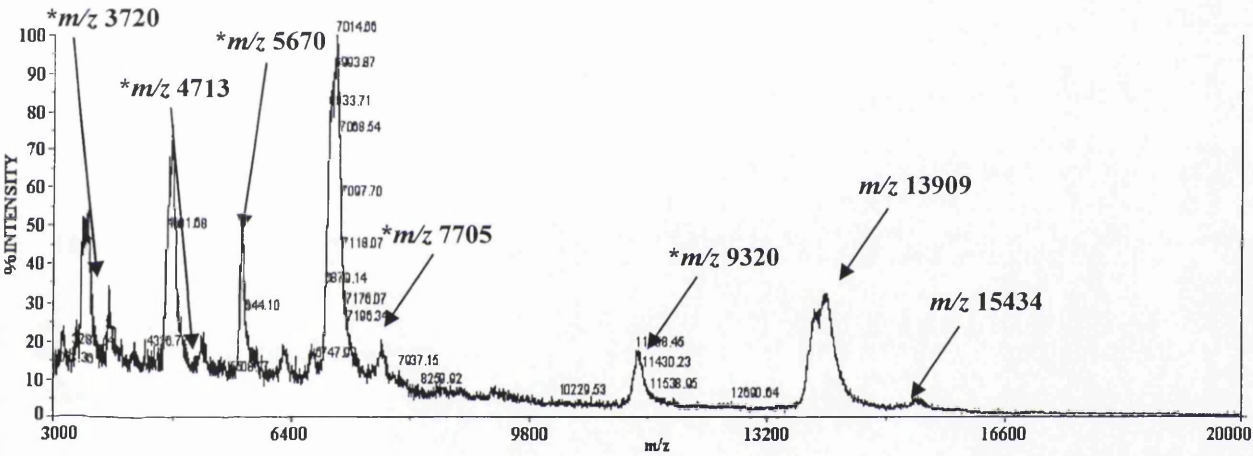


Figure 5.32: Positive ion MALDI mass spectrum of a region of thyroid tissue prepared using pH 6 antigen retrieval buffer with CHCA and tested immunohistochemically positive for the 15.4kDa protein Calcitonin (precursor). Also present is a peak at *m/z* 13909 and could correspond to Calcitonin gene-related peptide 1 (precursor). Peaks labelled with \* are those that could be generated by the fragmentation of 15.4kDa Calcitonin (precursor).

5.4.2.2.2.4. Results: S100 (precursor)

This protein can exist as four different chains,  $\beta$ ,  $\gamma$ , p, and z of masses 10.7, 9, 10.4 and 11.5kDa, respectively. This protein is common to cancerous tissue and gave a positive result within melanoma using an immunohistochemical test. A subsequent

tissue section was prepared using pH 6 antigen retrieval buffer and analysed by MALDI mass spectrometry in both positive and negative modes using CHCA, sinapinic acid and dansylhydrazine. Both the conventional matrices appeared successful in ionising the z-chain in positive mode with a peak observed at  $m/z$  11500, and displayed several other peaks not consistent with the molecular species of the S100 proteins. However, in positive mode sinapinic acid was capable of ionising more S100 proteins than CHCA, showing both the z- and  $\gamma$ -chain and the latter present at  $m/z$  9000.

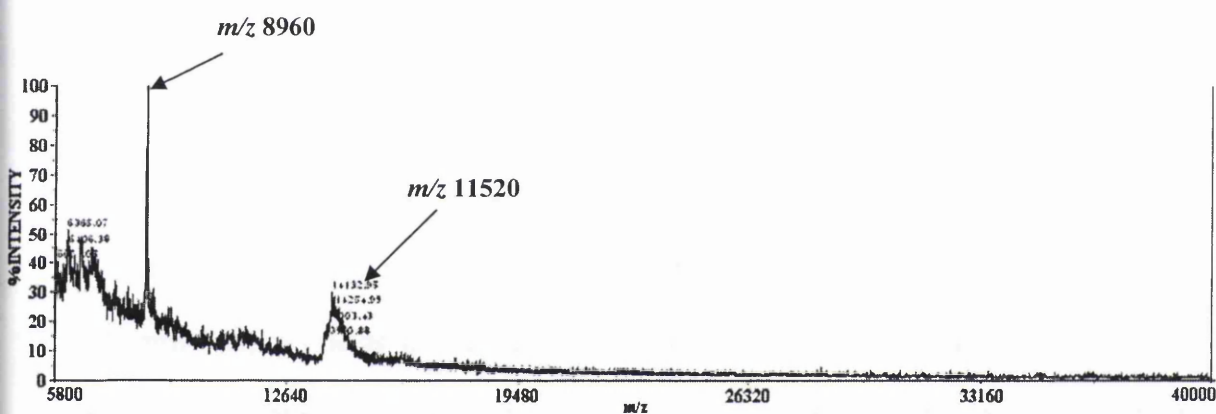


Figure 5.33: Positive mode mass spectrum of melanoma tissue section prepared using the pH6 antigen retrieval buffer and sinapinic acid containing two of the relevant proteins for S100 family;  $\gamma$ -chain at  $m/z$  9000, z-chain at  $m/z$  11500.

These peaks are absent in the negative ion mode mass spectrum, and peaks are observed at a reduced mass-to-charge of 11200 and 8700. This set of proteins may require the loss of a moiety prior to exposing a functional group capable of maintaining a negative charge. Dansylhydrazine however, proved unsuccessful in ionising the tissue section and no peaks were observed within the resulting mass spectrum.

#### 5.4.2.2.2.5. Discussion: Immunohistochemical vs Mass Spectral Imaging

This data has shown the capability of an antigen retrieval buffer in priming paraffin embedded archived tissue sections for analysis by MALDI mass spectrometry. One major concern with this type of antigen retrieval is the persistent heating of tissue within the buffer solution. This has the potential to cause fragmentation and diffusion of some protein or peptide material within the section, thus affecting the accuracy of its native location. However, in spite of this we have demonstrated that CHCA and sinapinic acid with this procedure can generate comparative data regarding the presence of known proteins in areas initially confirmed by immunohistochemical testing. In addition, several other peaks have been observed and are consistent with related proteins and peptides to those subject to the immunohistochemical protocol. Some of these peaks could be indicative of protein fragmentation and in particular the 'b-series' of fragment ions. A commonality with all these sites of cleavage is that the c-terminal amino acids contain neutral or basic side chains. This is contrary to published literature regarding site specific fragmentation. In-source fragmentation has been associated with a greater number of 'a-' and 'c-series' ions rather than the 'b-series', and fragmentation of singly charged peptides ionised by MALDI will preferably cleave at the C-terminus of acidic amino acid residues<sup>[39, 40]</sup>. Hence, it is probable that the fragmentation observed within the MALDI mass spectra of this tissue was generated prior to analyses and during the fixation process. This fragmentation may prove beneficial in assigning a particular disease state to tissue. Current immunohistochemical tests are reliant on antibodies generated in bulk that are specific to particular sites on the antigen protein. If this target area of the protein is damaged or fragmented during the fixation process it may appear unrecognisable to the relevant antibody. MALDI mass spectrometry is not affected by this type of

discrimination, and therefore has a greater potential to be used as a general screen of the tissue section. Currently, this interrogation of the section is reliant on the laborious expert visual examination by a pathologist, and hence, this novel MALDI approach could also improve the efficiency of analyses.

## 5.5. Conclusions

The aims of this body of work were to:

- identify a novel hydrophobic matrix capable of penetrating internal proteins and peptides within a tissue section that allows analysis by both mass spectrometry and fluorescence detection;
- assess performance each matrix in ionising peptides and proteins and their application to frozen, deparaffinised, and paraffin-embedded tissue sections and,
- improve the preparation of archived paraffin-embedded tissue sections for MALDI mass spectrometry.

A selection of fluorescent compounds known to interact with biomolecules as potential novel MALDI matrices have been tested. Each dansylated matrix showed better observed penetration into the tissue, yet maintaining fluorescence detection, than the standard matrices CHCA, sinapinic acid and DHB. Of these novel matrices dansylhydrazine proved most successful in ionising proteins and peptides by forming both a protonated molecule and reacting to form an adduct. These additional mass shifted peaks, when included in a tryptic peptide database search, can improve the probability of the original protein/peptide identification. In comparison to CHCA,

dansylhydrazine did not perform as well for both intact proteins or peptides and those subject to tryptic digestion. Interestingly, CHCA did show a potential to be identified on tissue using fluorescence detection. This matrix does appear the most suitable for ionising these biomolecules regardless of tissue preparation, but is limited to analysing those that are associated with the surface of the section. This work suggests that we have the potential to obtain a total image of frozen tissue using CHCA and dansylhydrazine in combination to ionise proteins or peptides at the surface or at depth, respectively. Unfortunately, it appears that the antigen retrieval protocol is insufficient preparation for dansylhydrazine to ionise de-paraffinised tissue sections. Hence, future work will be required for this total imaging principle using CHCA and dansylhydrazine to be applied for archived sections.

#### 5.5.1. Future Work

Developments in sample preparation regarding matrix deposition have recently shown that electrospray application is necessary in achieving sufficient signal intensity and reproducibility across the tissue section. Hence, this technique should be investigated using dansylhydrazine as an individual matrix and as a mix with CHCA, to assess its suitability as a fluorescent matrix capable of analysing internal tissue proteins and peptides. This matrix then has the potential to be used in tandem with CHCA and current protocols to obtain a full 'image' of the section. Additional work would include improving the compatibility of dansylhydrazine with deparaffinised sections and assessing the degree to which the antigen retrieval buffer protocol would displace proteins or peptides from within the tissue. This may be carried out by spiking the tissue with a known protein not expected to be present, such as a bovine protein within a human section, and measuring the degree of movement. If this buffer



protocol is deemed unsuitable for detecting these biomolecules and their location, then perhaps other antigen retrieval methods could be investigated.

## References

1. Todd, PJ., Schaaff, TG., Chaurand, P., and Caprioli, RM., *Journal of Mass Spectrometry*, 2001. **36**: p. 355-369.
2. Chaurand, P., Fouchecourt, S., DaGue, BB., Xu, BJ., Reyzer, ML., Orgebin-Crist, M-C., and Caprioli, RM., *Proteomics*, 2003. **3**: p. 2221-2239.
3. Pierson, J., Norris, JL., Aerni, H-R., Svenningsson, P., Caprioli, RM., and Andren, PE., *Journal of Proteome Research*, 2004. **3**: p. 289-295.
4. Xu, BJ., Caprioli, RM., Sanders, ME., and Jensen, RA., *Journal of American Society for Mass Spectrometry*, 2002. **13**: p.1292-1297.
5. Maddalo, G., Petrucci, F., Iezzi, M., Pannellini, T., Del Boccio, P., Ciavardelli, D., Biroccio, A., Forli, F., Di Ilio, C., Ballone, E., Urbani, A., and Federici, G., *Clinica Chimica Acta*, 2005. **357**: p. 210-218.
6. Yew, JY., Dikler, S., and Stretton, AO., *Rapid Communications in Mass Spectrometry*, 2003. **17**: p. 2693-2698.
7. Karas, M., Bachmann, D., and Hillenkamp, F., *Analytical Chemistry*, 1985. **57**: p. 2935-2939.
8. Karas, M., and Hillenkamp, F., *Analytical Chemistry*, 1988. **60**: p. 2299-2301.
9. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., and Yoshida, T., *Rapid Communications in Mass Spectrometry*, 1988. **2**(8): p. 151-153.
10. Ehring, H., Karas, M., and Hillenkamp, F., *Organic Mass Spectrometry*, 1992. **27**: p. 472-480.
11. Zenobi, R., and Knochenmuss, R., *Mass Spectrometry Reviews*, 1998. **17**: p. 337-366.
12. Krause, J., Stoeckli, M., and Schlunegger, UP., *Rapid Communications in Mass Spectrometry*, 1996. **10**: p. 1927-1933.
13. Beavis RC., and Chait, BT., *Rapid Communications in Mass Spectrometry*, 1989. **3**: p. 233-237.
14. Karas, M., Bahr, U., Strupat, K., Hillenkamp, F., Tsarbopoulos, A., and Pramanik, BN., *Analytical Chemistry*, 1995. **67**: p. 675-679.
15. Spengler, B., Kirsch, D., and Kaufmann, R., *Journal of Physical Chemistry*, 1992. **96**: p. 9678-9684.

16. Zhu, L., Parr, GR., Fitzgerald, MC., Nelson, CM., and Smith, LM., *Journal of American Chemical Society*, 1995. **117**: p. 6048-6056.
17. Thierolf, M., Bahr, U., and Karas, M., *Proceedings of the 45th ASMS Conference on Mass Spectrometry and Allied Topics*. Palm Springs: FL, 1997. p. 856.
18. Dreisewerd, K., *Chemical Reviews*, 2003. **103**: p. 395-425.
19. Vertes, A., Irinyi, G., and Gijbels, R., *Analytical Chemistry*, 1993. **65**: p. 2389-2393.
20. Muddiman, DC., Gusev, AI., and Hercules, DM, *Mass Spectrometry Reviews*, 1995. **14**: p. 383-429.
21. Schwartz, SA, Reyzer, ML, and Caprioli, RM., *Journal of Mass Spectrometry*, 2003. **38**: p. 699-708.
22. Desnovers, L., Therien, I., and Manjunath, P., *Molecular Reproductive Developments*, 1994. **37(4)**: p. 425-435.
23. Visser, SA., Smulders, CJ., Gladdines, WW., Irth, H., van der Graaf, PH., and Danhof, M., *Journal Chromatography B, Biomedical Scientific Applications*, 2000. **745(2)**: p. 357-363.
24. Yamada, K., Onodera, M., and Aizawa, Y., *Journal Pharmacological Methods*, 1983. **9(2)**: p. 93-100.
25. Hoffmann, K., Resch-Genger, U., Mix, R., and Friedrich, JF., *Journal of Fluorescence*, 2006. **16**: p. 441-448.
26. Seiler, N., and Weichmann, M., *Journal of Chromatography*, 1967. **28**: p. 351-362.
27. Chayen, R., Dvir, R., Gould, S., and Harell, A., *Analytical Biochemistry*, 1971. **42(1)**: p. 283-286.
28. Anderson, JM., *Analytical Biochemistry*, 1986. **152(1)**: p. 146-153.
29. Park, S-J., Song, J-S., and Kim, H-J., *Rapid Communications in Mass Spectrometry*, 2005. **19**: p. 3089-3096.
30. Cohen, SL., and Chait, BT., *Analytical Chemistry*, 1996. **68**: p. 31-37.
31. Chan, T-WD., and Colburn, AW., *Organic Mass Spectrometry*, 1992. **27**: p. 53-56.
32. Klein, B., Sheehan, JE., and Grunberg, E., *Clinical Chemistry*, 1974. **20(2)**: p. 272-274.
33. Strupart, K., Hillenkamp, F., and Karas, M., *International Journal of Mass Spectrometry and Ion Processes*, 1991. **111**: p. 89-102.

34. Beavis, RC, Chaudhary, T., and Chait, BT. *Organic Mass Spectrometry*, 1992. **27**: p. 156-158.
35. Crockett, DK., Zhaosheng, L., Vaughn, CP., Lim, MS., and Elenitoba-Johnson, KSJ., *Laboratory Investigation*, 2005. **85**: p. 1405-1415.
36. Arthur, JM., Thongboonkerd, V., Scherzer, JA., Cai, J., Pierce, WM., and Klein, JB., *Kidney International*, 2002. **62**: p. 1314-1321.
37. Petermann, JB., Born, W., Chang, JY., and Fischer, JA., *Journal of Biological Chemistry*, 1987. **262**: p. 542-545.
38. Kitamura, K., Kangawa, K., Kawamoto, M., Ichiki, Y., Matsuo, H., and Eto, T., *Biochemical, Biophysical Research Communication*, 1992. **185**: p. 134-141.
39. Lennon, JJ., and Walsh, KA., *Protein Science*, 1997. **6**: p. 2446-2453.
40. Wattenberg, A., Organ, AJ., Schneider, K., Tyldesley, R., Bordoli, R., and Bateman, RH., *Journal of American Society for Mass Spectrometry*, 2002. **13**: p. 772-783.

## CHAPTER 6:

# Identification of a Novel Haemoglobin Variant by Ion Mobility Spectrometry Coupled to Time-Of-Flight Mass Spectrometry

### 6.1. Introduction

Human haemoglobin (Hb) is an iron containing metalloprotein located within red blood cells and is necessary for oxygen transport within the blood. It exists naturally as a tetrameric aggregate consisting of two  $\alpha$  and two  $\beta$  peptide chains of 141 and 146 amino acids in length at masses (in adults) of 15126.4 and 15867.2Da, respectively. Contained within this tetrameric structure are four heme groups of mass 616.2Da and having an elemental formula of  $C_{34}H_{32}N_4O_4Fe$ . Variations or mutations within the amino acid sequence of the haemoglobin protein can occur by a process of substitution and there are more than a thousand known variants carried by approximately 1 in 800 people. Common variants consist of a single amino acid substitution within  $\alpha$  or  $\beta$  peptide chains and are most frequently detected during antenatal or neonatal screening. Additional sources of diagnosis have included preanaesthetic tests prior to an operation and indirectly through diabetic testing, by measuring the proportion of glycated haemoglobin. These variants are linked to many clinical conditions ranging from Sickle cell anaemia, and haemolytic anaemia to haemoglobin thalassaemia.

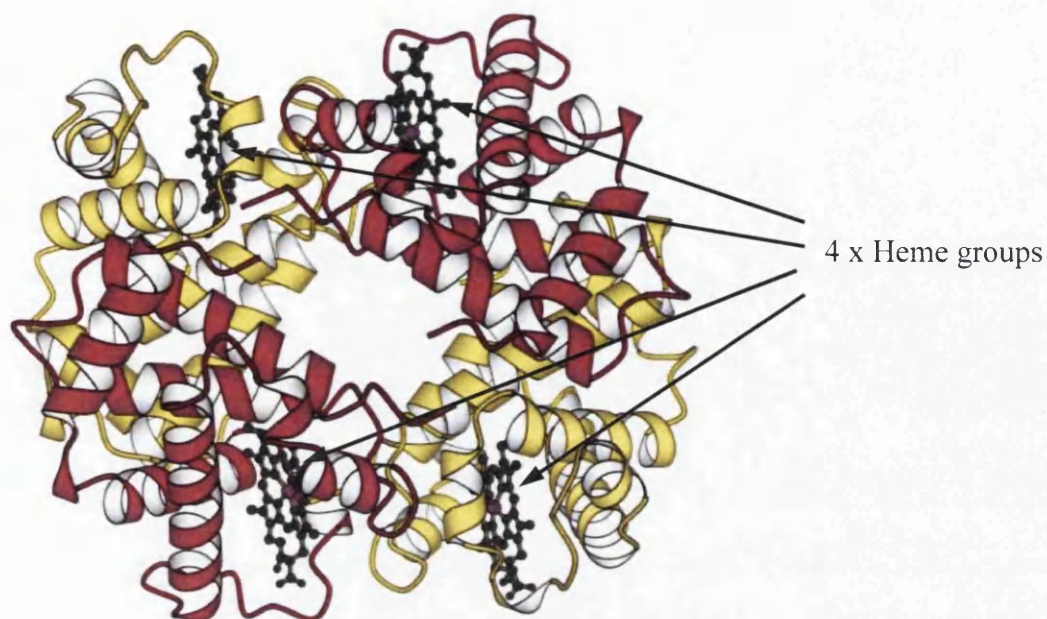


Figure 6.1: 3-Dimensional structure of human haemoglobin showing two  $\alpha$  and  $\beta$  peptide chains (in magenta and yellow) and the heme groups located within the each chain<sup>[1]</sup>.

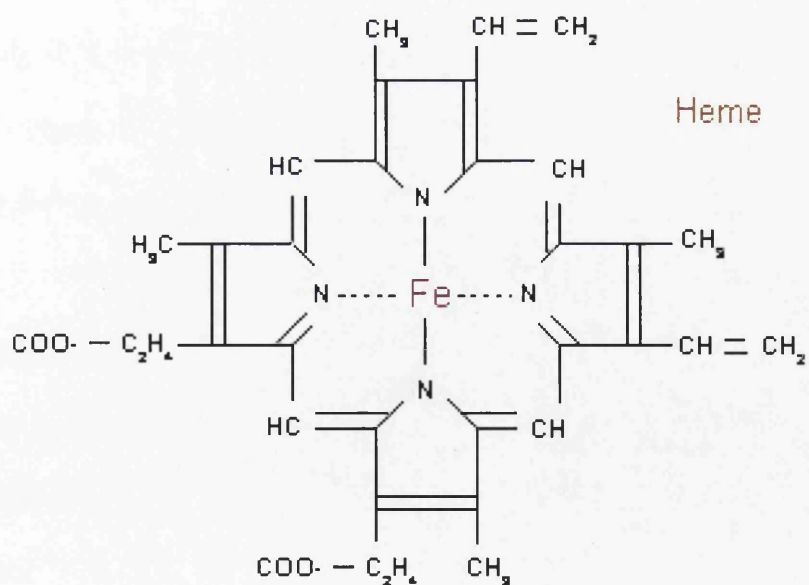


Figure 6.2: Cross-sectional view of a heme moiety indicating the interaction of the amino groups of the four peptide chains with the iron element.

### 6.1.1. Current preparative methods for identifying variants

Definitive characterisation of the haemoglobin mutation requires structural elucidation and is often carried out using mass spectrometry with sample preparation involving endoproteases such as trypsin and chymotrypsin. These techniques can be relatively costly and require a high degree of scientific expertise. Therefore, initial screening protocols can employ the following techniques;

#### 6.1.1.1. Electrophoresis

This technique under basic pH conditions has been the conventional screening protocol applied to blood samples for variant detection since the 1960's. It was developed over many years initially from a paper medium, to starch, and then cellulose acetate to optimise separation. These advances enabled the detection of all common clinically relevant variants, such as haemoglobin-S (Hb-S), Hb-C, Hb-D-Punjab, Hb-E, Hb-O-Arab, and Hb-Lepore by 1973<sup>[2]</sup>. This protocol was well suited for routine screening as it was capable of separating Hb-A and Hb-F in less than 30 minutes analysis time. However, more recently due to advanced techniques a greater number of variants of over 900 have been discovered and resulting in an increased requirement for a more accurate identification technique.

#### 6.1.1.2. High-performance liquid chromatography

HPLC is now commonly used in clinical laboratories for this application as it is easily automated and requires a smaller blood sample volume (approximately 5 $\mu$ L) than the electrophoretic technique. Existing protocols are reliant on the separation of the blood sample using a cation exchange column on which the retention time will differ if a variant is present. Amino acids other than arginine, lysine, histamine, asparagine and

glycine are neutral when placed in solution. Arginine, lysine and histamine can form positive charges while asparagine and glycine become negatively charged when in solution. Hence, if an amino acid substitution of a variant causes a deviation with the overall charge of the protein then a retention time shift will be observed on the cation exchange column. These principles are used to estimate the amino acid change from the normal haemoglobin molecule, often designated Hb-A0. A decrease in retention time from Hb-A0 is associated with an increase in the number of negative charges (J-like = 1 charge, I-like = 2 charges) as the protein is less likely to interact with the negatively charged cation exchange resin. An increase in retention time therefore corresponds to an increase in the number of positive charges (D-like = 1 charge, C-like = 2 charges) as the protein is much more likely to bind to the oppositely charged column resin. However, if a variant co-elutes at the retention time of Hb-A0 it is assumed that the substitution involves no change in the overall charge state of the protein. It is this type of mutation that is most problematic to identify and often require an additional technique such as DNA sequencing or the more rapid mass spectrometry as shown in latter section of this chapter.

#### 6.1.1.3. DNA sequencing

DNA sequencing is a definitive technique regarding the acquisition of structural information and the eventual variant identification, but unfortunately, this technique is not readily available within the vast majority of clinical laboratories due to a shortage of technical expertise and resources. For example, the reagent solutions used for the protocols incur a greater cost than the HPLC or electrophoretic methods. This procedure does however, have excellent sensitivity and would be better suited to confirming the identity of the variant rather than a general screening procedure.



#### 6.1.1.4. Liquid chromatography-mass spectrometry

Mass spectrometry has been used for analysing tryptic peptides of haemoglobin chains since the early 1980's in the form of field desorption MS<sup>[3]</sup>. Several different modes of ionisation have been examined to improve the reproducibility and sensitivity for these biomolecules, although neither of these techniques were particularly easy to use. The development of electrospray ionisation (ESI) and high quality deconvolution software has enabled the molecular weights of the intact haemoglobin chains, including the variants, to be determined<sup>[4]</sup>. It has been suggested that most clinical laboratories looking for these variants have previously avoided mass spectrometry as a conventional screening tool due to lengthy sample preparation procedures formerly associated with this technique<sup>[5]</sup>. A great advancement in minimising blood sample preparation for haemoglobin studies was achieved by Nakanishi and co-workers. This group removed the HPLC separation step and analysed complex mixtures of tryptic peptides directly by infusion electrospray ionisation-mass spectrometry (ESI-MS). However, despite achieving relatively good sequence coverage along the whole of  $\alpha$  and  $\beta$ -chains, the peptides containing cysteine residues were poorly observed. This, in addition to a high level of expertise required for data analysis and interpretation, has resulted in clinical laboratories persevering with HPLC for conventional screening.

## 6.2. The Analytical Problem

Current conventional screening protocols rely on either an electrophoretic or chromatographic separation. The specificity of separation is not sufficient for categorically identifying a wide range of haemoglobin variants that exist in the current patient population. A common scenario encountered is the presence of an additional

peak co-eluting with the normal haemoglobin chains or their glycosylated moieties, giving the appearance of an abnormally shaped chromatographic peak. Glycosylated haemoglobin is often used as a biomarker for charting diabetes progression and interference with this peak can result in an incorrect glycation or blood glucose level. Using a modified methodology of that fashioned by Green and co-workers<sup>[6]</sup> blood samples with suspicious HPLC chromatograms were interrogated using ESI-MS with the added dimension of ion mobility separation. Unlike cation exchange HPLC, mass spectrometry enables a more accurate determination of haemoglobin glycation as it involves the individual monomer chains and not their corresponding dimers. Hence, this procedure should provide both an accurate variant identification and glycation level<sup>[7]</sup>. To ensure sufficient reliability of results this is usually carried out in conjunction with a phenotypic (DNA sequencing) technique to provide additional evidence regarding the 'type' of variant present. There are four mass spectrometry experiments required to determine the identity of the variant and to characterise the relevant mutation (see figure 6.3).

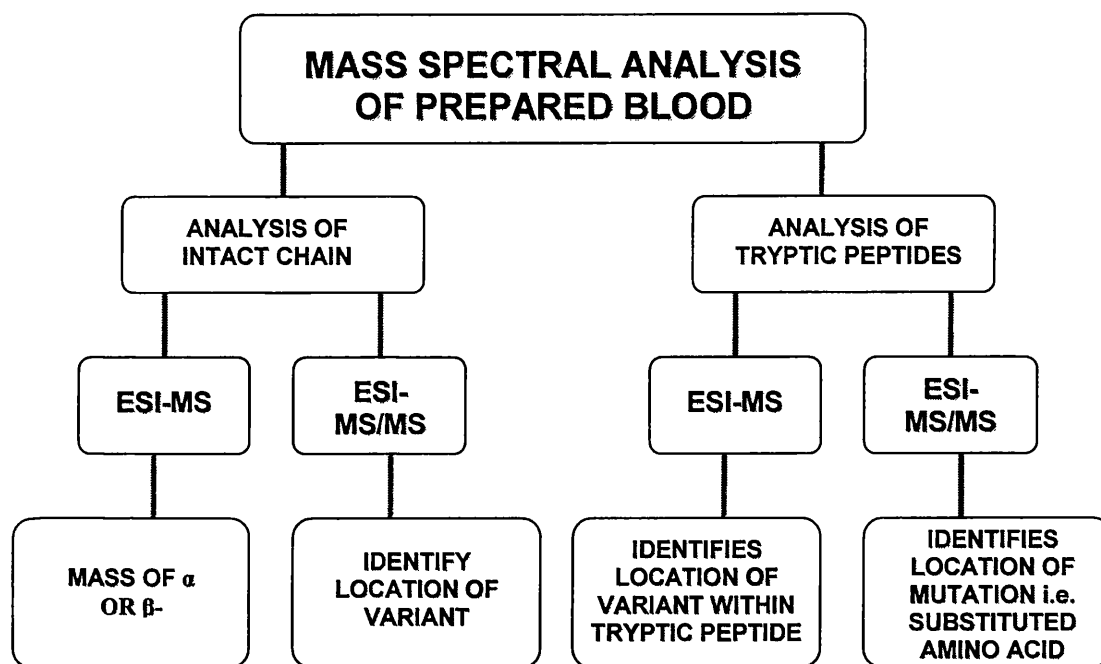


Figure 6.3: Flow chart illustrating the stages of variant identification by mass spectrometry.

### 6.3. Application of Novel Analytical Mass Spectrometric Methodology for Identifying the Novel Variant

#### 6.3.1. Sample preparation

##### 6.3.1.1. Intact haemoglobin chains

For this experiment 10 $\mu$ L of a whole blood sample was diluted with 490 $\mu$ L of deionised water. A 20 $\mu$ L aliquot of this stock sample solution was further diluted with 180 $\mu$ L of 50:50 acetonitrile:0.2% formic acid. This dilution was then desalted with prepared BIORAD (Hercules, CA, USA) AG50WX8 ion exchange beads (100-200 mesh hydrogen form), and infused into the Synapt HDMS (Waters, Milford, MA, USA) and the Quattro II (Waters) mass spectrometers at a flow rate of 5 $\mu$ L/minute.

#### 6.3.1.2. Tryptic digestion

Whole blood samples were diluted 1:50 with deionised water, and 20 $\mu$ L of 50:50 acetonitrile:0.5% formic acid was added to 100 $\mu$ L of this stock solution. This solution was vortex mixed and left to stand for 5 minutes at room temperature. To this solution 5 $\mu$ L of 5mg/mL trypsin (TPCK treated solution, Sigma Aldrich, Munich, Germany) or  $\alpha$ -chymotrypsin (TLCK treated, Sigma Aldrich) were added following 6 $\mu$ L of 1M ammonium bicarbonate solution (pH 7.8). The resulting mixture was again vortexed, pulse centrifuged for approximately 10 seconds once the solution became clear, and then incubated for 30 minutes at 37°C. Prior to infusing into the mass spectrometers, this digested solution was diluted 1:4 by adding 40 $\mu$ L to 360 $\mu$ L of 50:50 acetonitrile:0.2% formic acid.

#### 6.3.2. Ion mobility mass spectrometry analysis

##### 6.3.2.1. Ion mobility mass spectrometry

This type of mass spectrometry has developed over the past 25 years from an existing technique known as plasma or ion chromatography<sup>[8, 9]</sup>. Prior to its current use with large biomolecules it was mainly used to analyse volatile organic compounds<sup>[10]</sup> and understand the electronic states of ions<sup>[11]</sup>. Modern ion mobility techniques combine electrospray ionisation (ESI) or matrix-assisted laser/desorption ionisation (MALDI) mass spectrometry with an ion mobility cell to further characterise large biomolecules such as peptides and proteins. Its fundamental principle is the separation of gas-phase ions based on their mobility and is dictated by their collision cross-section through a neutral target gas, such as helium. Ions enter the collision cell containing a buffer gas at a specific pressure (p) and migrate through the cell with the application of a linear electric field (E). The frequency of collision with the inert buffer gas determines the

ion separation observed and is related to the 3-dimensional physical volume of each ion or collision cross-section. Hence, this technique is not only capable of separating ions according to mass-to-charge but also conformation or shape. This offers the advantage, when compared to liquid chromatography-mass spectrometry (LC-MS), of further isolating of large biomolecules, such as proteins and peptides, of the same mass-to-charge and polarity but with a different overall conformation. This enables the detection of components present at very low concentrations and isobaric species previously undiscovered by LC-MS alone. In addition, ion mobility has several added advantages; separation is carried out over a microsecond timescale unlike LC methods that often require hour(s) of sample preparation, and unlike some gel methods, ion mobility separation is highly reproducible<sup>[12]</sup>.

#### 6.3.2.1.1. Fundamentals of ion mobility

Separation is based variations in the drift velocity ( $v_d$ ) of an ion through a buffer gas with an applied electric field (E). An ion's drift time is defined by the mobility constant (K) and is related to drift velocity by the following equation<sup>[13]</sup>:

$$(v_d) = KE \qquad \text{Equation 6.1}$$

This relationship therefore contains information regarding the degree of interaction of the ion and buffer gas. For atomic ions this expression implies that the mobility of these ions is primarily dependent on their electronic state<sup>[14]</sup>. However, polyatomic ions are more complex and require a more complex expression to describe their mobility. This is known as the average collision cross-section and as most biomolecules are polyatomic in nature we will use this term to characterise ion

mobility throughout this chapter. The mobility constant (K), shown in equation 6.1, is often given as a value at standard temperature and pressure known as the reduced mobility constant (K<sub>0</sub>):

$$K_0 = \frac{p}{760} \times \frac{273.15}{T} \times K \quad \text{Equation 6.2}$$

The interaction of an ion with the buffer gas and hence its mobility constant is dependent on several parameters; the charge of the ion (q), the density of the buffer gas (N), the gas temperature (T), the ion's collision cross-section (Ω<sub>0</sub>) and the ion-neutral (buffer gas) complex reduced mass (μ). These factors form the fundamental equation used for determining ion mobility (shown below) and it is apparent that mobility is inversely proportional to the square root of the reduced mass and the collision cross-section.

$$K = \frac{3}{16} \times \frac{q}{N} \times \left( \frac{1}{\mu} \times \frac{2\pi}{k_b} \right)^{\frac{1}{2}} \times \frac{1}{\Omega_0} \quad \text{Equation 6.3}$$

The ion-neutral complex mass (μ) formed between the buffer gas (m<sub>1</sub>) and the analyte ion (m<sub>2</sub>) is considered a constant during protein and peptide analyses. This is because the value of the ion-neutral complex approaches the mass of the buffer gas when measuring ions of mass greater than 500Da. Hence, a more concise expression can be used to describe ion mobility under these conditions indicating that it is inversely dependent on the collision cross-section alone (see equation 6.4)<sup>[15]</sup>.

$$K \propto \frac{1}{\Omega_0}$$

*Equation 6.4*

where,

$$\Omega_0 = \mu b v_{rel}$$

and,

$$\mu = \frac{m_1 \times m_2}{m_1 + m_2}$$

The collision cross-section itself is dependent on several parameters, and for ions above 500Da it will account for the scattering angle between the ion and buffer gas, with their relative velocities ( $v_{rel}$ ) and mean geometry of the analyte ion ( $b$ )<sup>[16]</sup>. From the latter part of the expression shown in equation 6.4 the collision cross-section now has physical parameters which can be controlled or varied to achieve the appropriate separation. Experiments involving biological macromolecules the collision cross-section becomes of fundamental importance for successful ion separation. For example, large peptide or protein ions of similar mass-to-charge can be separated and distinguished as those with a more open conformation (a larger cross-sectional volume) will have a greater propensity to interact with the buffer gas and have a longer drift time. It is this principle that is so effective for the separation of proteins and peptides and, providing sufficient resolution, enables the isolation and detection of relatively large isomeric species.

#### 6.3.2.1.2. Degree of ion mobility separation or resolution

Unlike most mass spectrometric techniques the parameter of peak resolution in ion mobility is more complex and is often considered to have greater similarities with chromatographic resolution. This concept, originated by Hill and co-workers<sup>[17, 18]</sup>, relates the drift time of the ion and the width of the ion peak:

$$R = \frac{t}{\Delta t} \times \frac{t}{W} \quad \text{Equation 6.5}$$

where,

R = Resolution

W = Width at peak base

= 4.7 x standard deviation( $\sigma$ ) of arrival time.

Peak width is dependent on four factors; initial pulse width, space-charge effects, diffusive behaviour of ions through the buffer gas, and interaction of the analyte ion with the buffer gas. As with chromatographic separations, a high level of resolution observed as a narrow peak width is required to give an accurate drift time and hence, these four parameters should be minimised to obtain sufficient resolution. The diffusive behaviour of the ion through the buffer gas can also be quantified using the expression shown in equation 6.6.



$$\sigma = (2Dt_d)^{\frac{1}{2}} \quad \text{Equation 6.6}$$

where,

$\sigma$  = spatial standard deviation of the diffusion of the ion  
through the buffer gas,

$D$  = diffusion coefficient of the ion and

$t_d$  = time of ion within drift cell.

Resolution can also be expressed as a function of the applied electric field, drift time and temperature by substituting the Nernst-Einstein relationship into equation 6.6:

$$D = K \times \frac{k_b T}{q} \quad \text{Equation 6.7}$$

where,

$K$  = ion mobility at temperature,  $T = \frac{v_d}{E}$

$v_d$  = drift velocity

therefore,

$$W_d = 4.7\sigma = \left( \frac{44.2k_b T v_d t_d}{qE} \right)^{\frac{1}{2}} \quad \text{Equation 6.8}$$

As the diffusion width greatly exceeds the remaining peak broadening factors the peak width ( $W$ ) approximates to the diffusion width and peak resolution can be described by equation 6.9.

$$R = \left( \frac{qEL}{44.2k_bT} \right)^{\frac{1}{2}}$$

*Equation 6.9*

Using this expression, peak resolution can now be described in terms of drift time and electric field, with the ion mobility resolution inversely proportional to the temperature of the analyses.

Hill and co-workers also identified that different buffer gases interact differently with the analyte ions, and can affect analyte separation and elution order<sup>[19]</sup>. Using chromatographic principles, a term describing separation was developed by comparing the ratio of the mobility constants of two individual ions travelling through a drift tube. This ion mobility separation factor ( $\alpha$ ) is defined by the following expression:

$$\alpha = \frac{K_1}{K_2}$$

*Equation 6.10*

where,

$K_1$  = mobility constant of the ion with the faster drift time

and,

$K_2$  = mobility constant of the ion with the slower drift time.

This expression is particularly useful for describing the capability of an ion mobility cell in achieving adequate separation of ions. For example, a separation factor that is close to or is equal to 1 indicates that the cell does not have the sufficient selectivity to separate the sample, and hence other parameters known to affect separation must be employed. The main outcome of this previous research implied that smaller drift

gases such as helium and nitrogen are better suited to separating large polyatomic ions as they have shorter overall drift times within these gases.

Recent advances in ion mobility have involved the separation of enantiomers<sup>[20]</sup>. Additives have been included within the buffer gas such as a chiral vapour, and have improved both the resolution and overall separation of these ions. The chiral modifier can interact differently with each enantiomer and alters the mobility drift time enough to allow these individual ions to be distinguished. The exact mechanism of this interaction is not yet fully understood and further research is required prior to conventional use.

#### 6.3.2.1.3. Sensitivity in ion mobility experiments

This is the major drawback encountered with ion mobility as significant quantities of ions can be lost due to diffusion within the ion mobility cell. This is apparent by the following equation describing the passage of ions through the cell:

$$A = 1 - e^{\left(\frac{-r_c^2}{4Dt}\right)} \quad \text{Equation 6.11}$$

where,

A = attenuation of ions through cell,

$r_c^2$  = radius of mobility cell exit,

D = degree of diffusion of ion perpendicular to the applied electric field and

t = drift time through mobility cell

To achieve sufficient sensitivity and levels of detection it is essential that the appropriate parameters are optimised for the analyses. Fortunately diffusive effects can be relatively easily controlled at low electric fields as the Nernst-Einstein expression can be applied. However, experiments at high applied electric fields, such as those requiring a high level of peak resolution, can encounter problems with ion transfer. For example, as the field increases the analyte ion transfers less translational energy into the buffer gas, resulting in a decrease in ion transmission and incorrect drift times<sup>[21]</sup>. To account for this field strength effect McDaniel and Moseley<sup>[21, 22]</sup> applied the Wannier expression to determine the ion diffusion coefficient maximising sensitivity under these conditions (see equation 6.12).

$$D = K \times \frac{kT}{q} + \frac{1}{3} m \times \frac{m + M}{m + 1.908M} \left( \frac{E^2 K^3}{q} \right) \quad \text{Equation 6.12}$$

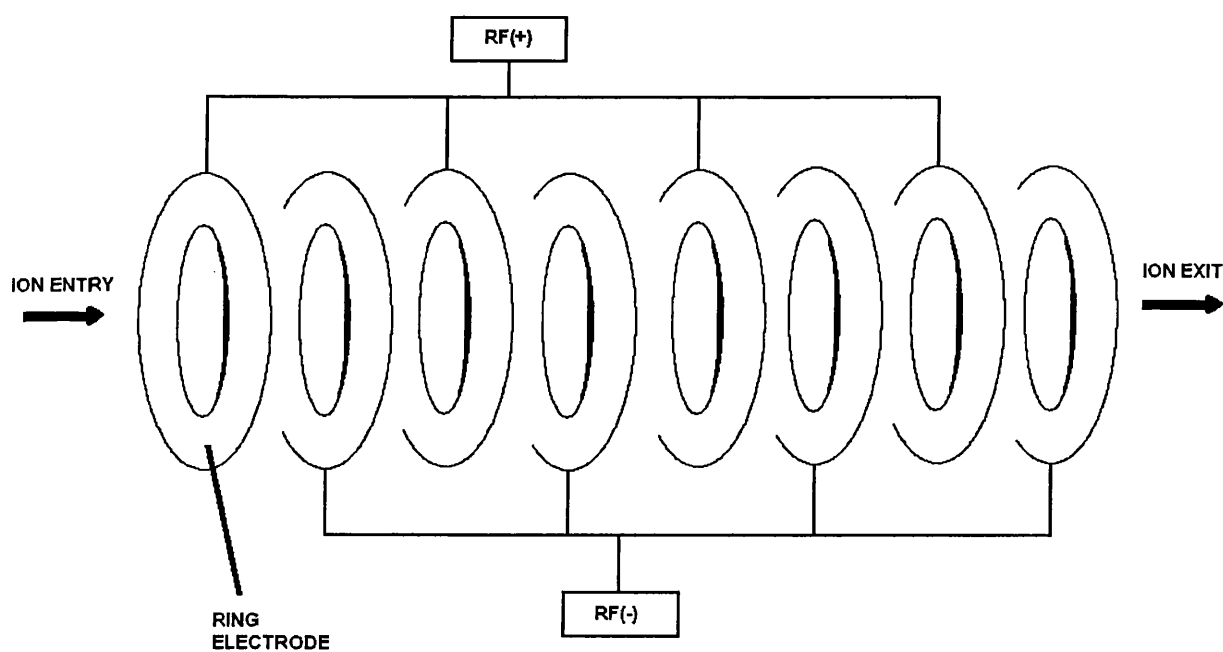
This expression (equation 6.12) enables ion mobility analyses to be carried out with maximum peak resolution and therefore separation or selectivity, and yet maintain a good level of sensitivity for complex biological samples.

#### 6.3.2.1.4. Instrumental developments in ion mobility mass spectrometry

Historically ion mobility instruments contained a mobility cell that is essentially a drift tube filled with a low pressure gas such as helium. Recent developments in this technology have included the introduction of a stacked ring ion guide (SRIG) that employs a travelling radio frequency (rf) wave capable of transmitting ions orthogonally into a time of flight mass analyser<sup>[23]</sup>. The major advantages of this

modified mobility cell are its capability of functioning as a fragmentation cell and an improved duty cycle.

Radio-frequency (rf) only ion guides are commonly used in most mass spectrometers in the form of the quadrupole, hexapole or octapole. A less familiar rf ion guide is the SRIG and was first invented by Bahr and co-workers in 1969<sup>[24]</sup>, to be further developed by Gerlich's research group<sup>[25, 26]</sup> in the early 1990's. It is essentially a stack of ring electrodes each with opposing rf potentials in sequence. Application of the rf potential out-of-phase results in the SRIG becoming an effective radially restricting ion trap (see figure 6.4).



*Figure 6.4: Diagram of stacked ring ion guide (SRIG) capable of retaining ions by an applied RF voltage only. Ions enter the first ring electrode and oscillate by the application of an RF voltage applied out-of-phase throughout the SRIG.*

This group later modified the SRIG by surrounding it with electrodes that have a potential difference designed to provide an axially directed field and limit specific ion losses along the cell. The effective potential within the SRIG provides it with an 'ion trap or ion pipe' characteristic and is described by the following expression:

$$V_{\max} = \left[ \frac{I_1^2(r/\delta)\cos^2 + I_0^2(r/\delta)\sin^2(z/\delta)}{I_0^2(\rho/\delta)} \right] \quad \text{Equation 6.13}$$

$$V_{\max} = \frac{qV_{RF}^2}{4m\omega^2\delta^2}$$

where,

$r$  = radial coordinate

$z$  = axial coordinate

$\delta$  =  $d/\pi$

$d$  = centre-to-centre spacing of the electrodes

$\rho$  = aperture radius

$V_{RF}$  = half of the applied peak-to-peak voltage with angular frequency

$\omega$

$q$  = ionic charge

$m$  = mass of the ion

$I_0, I_1$  = zero and first-order modified Bessel functions, respectively

and

$V_{\max}$  = maximum effective potential at  $r = \rho, z = d(i + 1/2)$ .

The travelling wave ion guide (TWIG) was developed from this modified SRIG to project ions through the cell and achieve separation as a form of ion mobility. A SRIG, containing a low pressure ( $\sim 0.2$  mbar) inert gas, such as argon, has a travelling voltage wave applied to the existing radial rf voltage in the form of a direct current (dc) potential on the first electrode. This propels ions to the second electrode and the dc potential is then switched to the second electrode. This process is repeated throughout the SRIG providing the travelling wave on which the ions flow and the basis of movement in ion mobility. Studies involving this technology has shown that at high SRIG gas pressures the drift velocity of the ion is representative of the expression given as equation 6.1. Under these practical considerations the 'smaller' ions will have low ion mobility as the ion will roll over the travelling wave more frequently than a 'larger' ion resulting in a longer drift time, as shown in figure 6.5.

Mass spectra based on ion mobility separation are acquired by trapping ions within the source ion guide and transferring them at set intervals into the TWIG using an ion gate. This 'packet' of ions enters the TWIG on the continuous travelling wave and results in ion separation as they approach the time-of-flight (ToF) mass analyser of the instrument subject to this investigation, the Synapt High Definition Mass Spectrometry (HDMS) hybrid mass spectrometer. Prior to entering the ToF analyser an electronic 'pusher' aligns the ions so they commence flight within the analyser at the same time. This signal is seen as a mass spectrum for a particular drift time and is repeated periodically to observe the whole mobility range of ions exiting the TWIG. The Synapt HDMS system is essentially a hybrid instrument based on the previous Q-ToF Ultima (Waters) design where the collision cell is replaced by a TWIG capable of ion mobility separations and fragmentation.

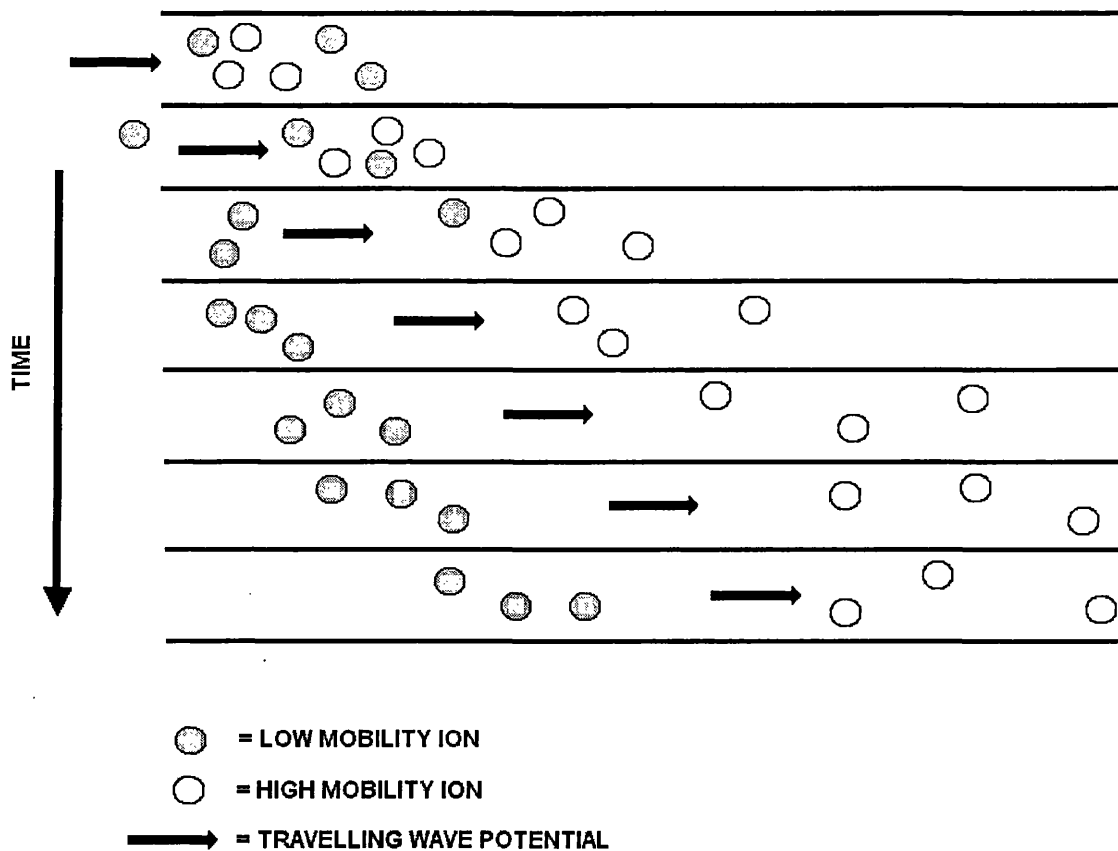


Figure 6.5: Illustration of mobility separation of ions within the travelling wave ion guide (TWIG). Separation is achieved by the number of times an ion moves over the travelling wave potential. Ions with high mobility experience a lower number of interactions with the travelling wave, and large ions with low mobility have a greater interaction resulting in larger drift times.



## Synapt HDMS System Hybrid Quadrupole / IMS / oa-ToF

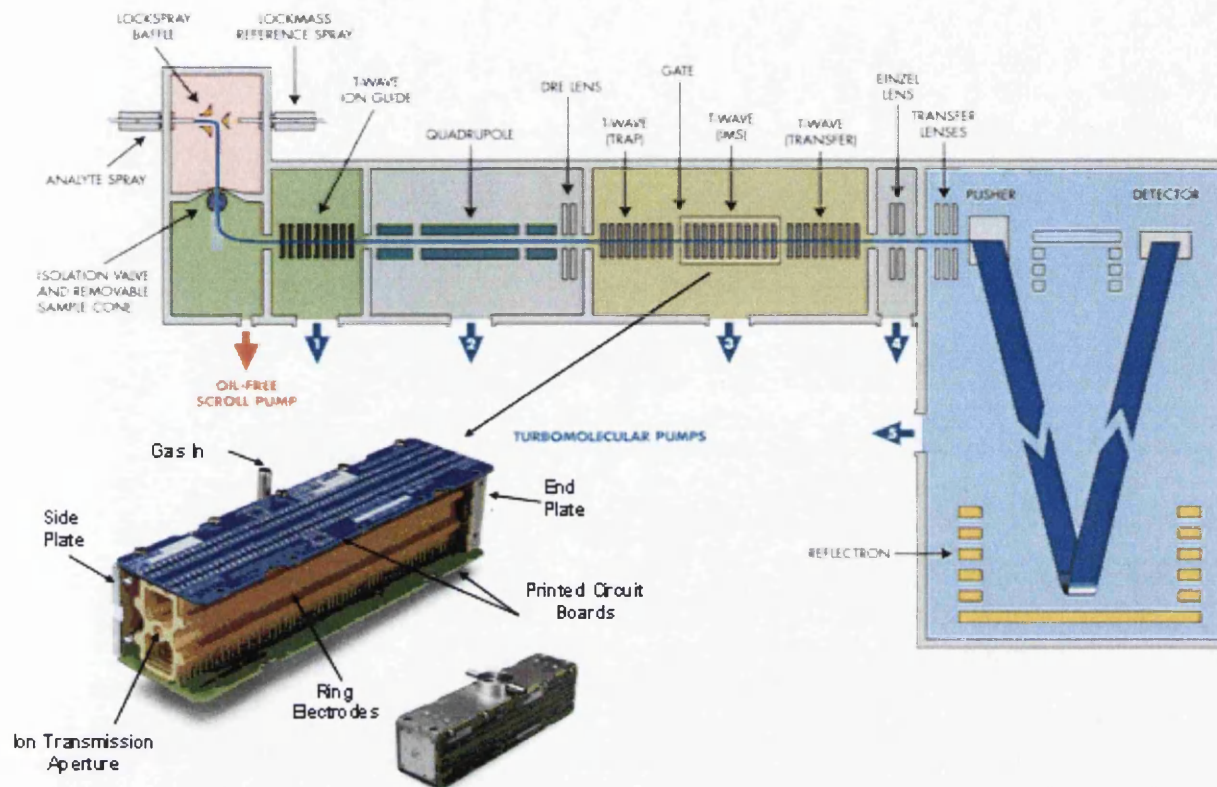


Figure 6.6: Schematic of Synapt HDMS ion mobility mass spectrometry system, showing an indepth view of an actual transmission wave ion guide (TWIG) device in place of the conventional collision cell. The TWIG illustration clearly displays the numerous stacked ring electrodes used in the ion mobility separation<sup>[23]</sup>.

The inclusion of a TWIG is advantageous as a major limitation of using ToF analysers is the length of the duty cycle. This is essentially the percentage of ions not detected per spectral acquisition or the time taken for the detector to scan the mass range associated with the ToF analyser. The 'pusher' time therefore can be optimised to improve the sensitivity of ions within a particular mass range. For example, if the mass window is narrowed, the detector requires less time to scan the given mass range, and allowing the accumulation of signal for these particular masses. The

TWIG apparatus enables this synchronisation of the ToF ‘pusher’ with each ion packet of differing mobilities and yet maintaining the enhanced mass-specific sensitivity. Primarily the advantage of using a ToF mass analyser is the large mass range and can enable the analysis of intact proteins with the use of multiple charging and deconvolution algorithms.

An additional advantage of using a TWIG is its capability of fragmenting ions. The propulsion of ions using a travelling wave results in the continuous application of kinetic energy within the TWIG. This energy wave can result in precursor ion fragmentation and the mobility separation of fragment ions<sup>[23]</sup>. This is advantageous when compared to conventional collision cells, as it exhibits less interference between several multiple reaction monitoring (MRM) experiments, yet maintaining sensitivity and resolution of both the precursor and fragment ions. This fragmentation and ion mobility separation was utilised for the work in this chapter to identify the location of the amino acid substitution associated with a mutation of haemoglobin chains.

### 6.3.2.2. Instrument parameters

#### 6.3.2.2.1. Synapt HDMS mass spectrometry system

<b>Instrument Parameter</b>	<b>Instrument Reading</b>
Capillary Voltage (V)	3.2
Cone Voltage (V)	30
Mobility Cell Gas	N <sub>2</sub>
Mobility Gas Pressure	0.5mbar
Trap/Transfer Lenses are Gas Pressurised with Argon	10 <sup>-2</sup> mbar
Velocity of Travelling Wave Constant (m/s)	300
Travelling Wave Pulse Height (dc amplitude – V)	7.0-17.0

*Table 6.1: Ion mobility instrument parameters for the full mass scan and fragmentation experiments.*

All fragmentation (MS/MS) experiments were carried out pre-ion mobility within the trap (first quadrupole, see figure 6.6) and at collision energy of 20eV.

### 6.3.2.3. Identification of a novel haemoglobin variant

Routine post natal screening of newborns had generated a suspicious liquid chromatogram for a patient. The sample of this subject (figure 6.7) showed an unexpected shoulder (designated P00) to a peak commonly used for measuring the glycation level of haemoglobin (SA1C). These appeared as one abnormally-shaped peak with some LC conditions and could provide an incorrect haemoglobin glycation percentage if unresolved.

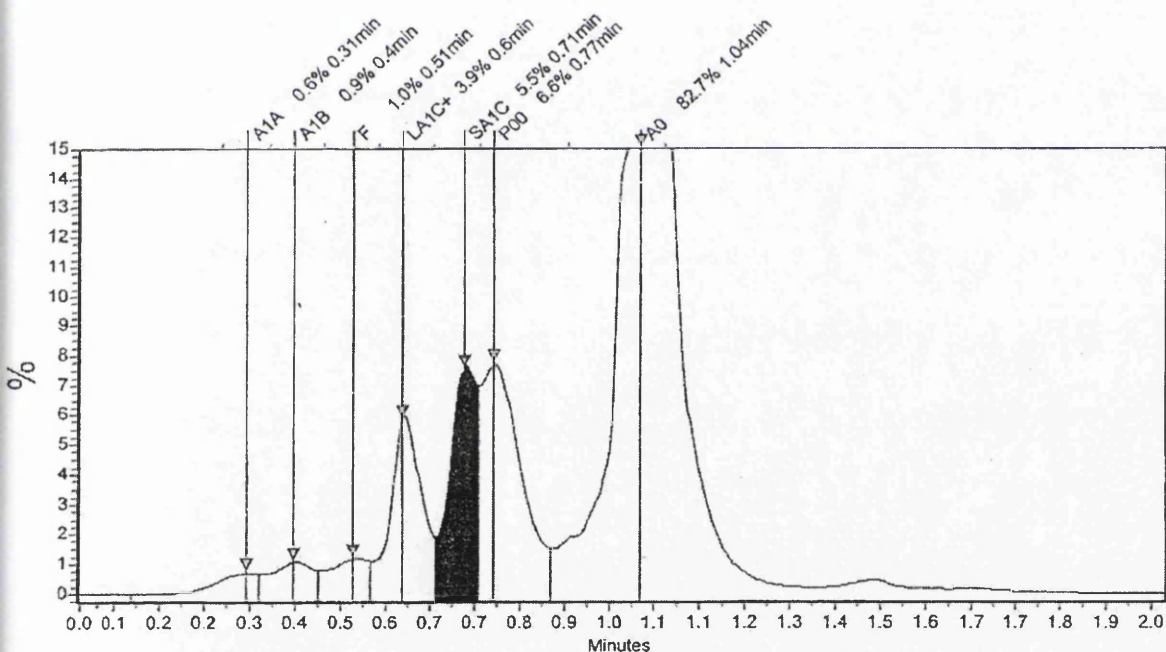


Figure 6.7: Initial chromatogram of blood sample subject to this investigation obtained during the routine haemoglobin screen. Highlighted in black background is the normal glycated haemoglobin peak labelled SA1C. Located to the right of this is the suspected variant, previously unresolved by the first chromatographic analysis and incorrectly included within the glycation calculation.

The glycation value is a fundamental indicator used in assessing the long-term development of diabetes and is frequently confirmed using mass spectrometry<sup>[7]</sup>. Hence, it is of great importance to identify any variant that co-elutes with glycosylated haemoglobin to ensure an accurate reading for diabetes diagnoses. These circumstances therefore, warranted further investigation to identify the exact composition and amino acid position of the mutation with the endeavour to improve speed of response for treatment of future sufferers.

#### 6.3.2.3.1. Analysis of intact haemoglobin chains – discovering a mutant

Infusion of blood samples containing denatured intact chains initially appear as highly populated spectra consisting of multiply charged ions of both the  $\alpha$  and  $\beta$  chains and their corresponding salts. Conventional deconvolution techniques, such as Maximum Entropy<sup>[27]</sup> (Waters) could provide an erroneous total ion mass-to-charge value from this data as it is unable to discriminate between multiple charged ions of more than one precursor. Desalting the blood sample (see section 6.3.1.1) has proven vital as it ensures good sensitivity of the multiply charged parent ions and minimises spectral complications due to adduct formation. The ion mobility dimension of the Synapt HDMS system enables the isolation of a particular charge state of a specific chain. Both haemoglobin chains differ in mass affecting the mass of peaks of the charge states, resulting in different conformations and hence, different ion mobilities<sup>[23, 28]</sup>. This is shown quite clearly by the ion chromatograms shown overleaf and for tryptic peptide mixtures in figure 6.11 and 6.12.

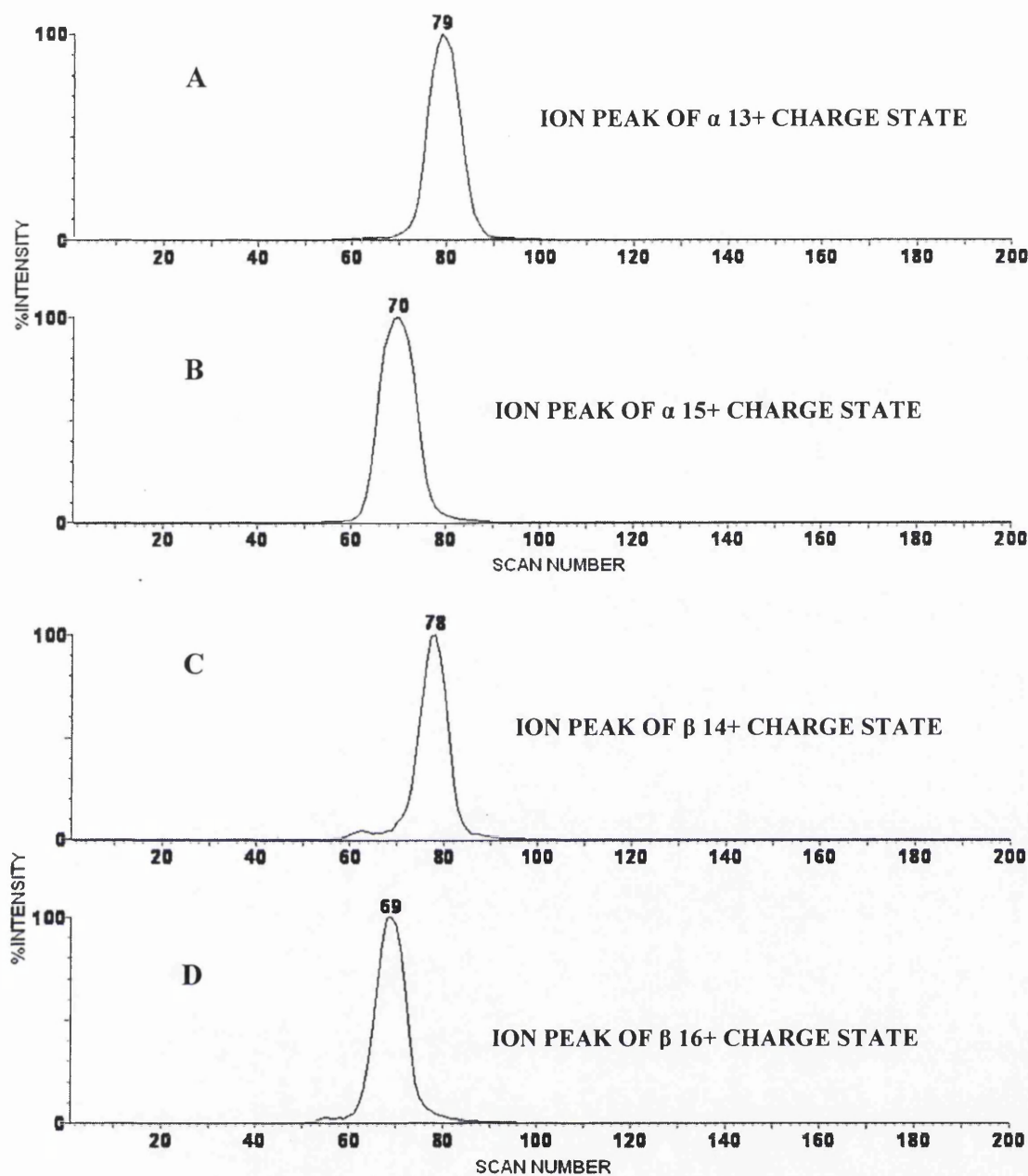


Figure 6.8: Ion chromatograms of the ions associated with the 13+ and the 15+ charge state of the  $\alpha$  haemoglobin chain (A and B) and the 14+ and 16+ charge state of the  $\beta$  haemoglobin chain (C and D). Each ion mobility scan has duration of  $65\mu\text{s}$ , therefore the ions involved in the 13+ charge state envelope have an ion mobility time which is the product of the scan duration and number, i.e.  $5.07\text{ms}$ . This indicates the speed of the experiment, and can be completed within  $13\text{ms}$ . The relatively high resolution of the ion mobility T-wave cell allows separation of these charge states of the respective chains, allowing deconvolution without interference.

Following mobility separation and deconvolution using Maximum Entropy software the normal  $\alpha$  and  $\beta$  haemoglobin chains were measured to have a mass of 15126.8 and 15867.5Da, respectively. These values are consistent with the literature<sup>[26]</sup> and when used as a control can indicate the mass difference associated with the haemoglobin variant. The samples subject to this investigation showed a mass shift of +30Da specific to the  $\beta$  chain only. This was deduced as mutations involving the  $\alpha$  haemoglobin chain did not correspond to the mass shift associated with variant peak. The +30Da mutation of the  $\beta$  chain shows a neutral change in the HPLC run and is consistent with several amino acid substitutions as indicated by table 6.2. However, the measured mass-to-charge of the variant parent ion can not confirm the actual substitution and is unable to identify relevant location. Hence, additional experiments involving endoproteases, such as trypsin, to obtain peptide ion information are required.

<b>Possible Amino Acid Substitutions for Variant due to Single Base Change in the Nucleotide Codon</b>	
<b>Mass Change</b>	<b>Amino Acid Change</b>
+30	Alanine (Ala) ↔ Threonine (Thr) Arginine (Arg) ↔ Tryptophan (Trp) Glycine (Gly) ↔ Serine (Ser) Threonine (Thr) ↔ Methionine (Met) Valine (Val) ↔ Glutamine (Glu)

*Table 6.2: Illustrating the possible amino acid substitution associated with the variant is consistent with a single nucleotide base change of +30Da.*

This variant sample also exhibited a greater degree of glycation illustrated by the clear presence of peaks +162Da above the normal haemoglobin chains. This glycation mass shift is consistent with the addition of a hexose based carbohydrate



such as fructose, galactose, glucose or mannose. Past research indicates that this is in fact due to the addition of a glucose moiety<sup>[29, 30]</sup> and this is a common biomarker in assessing the development of diabetes. The differences attributed to the variant chain are clearly visible in the spectra shown in figure 6.9.

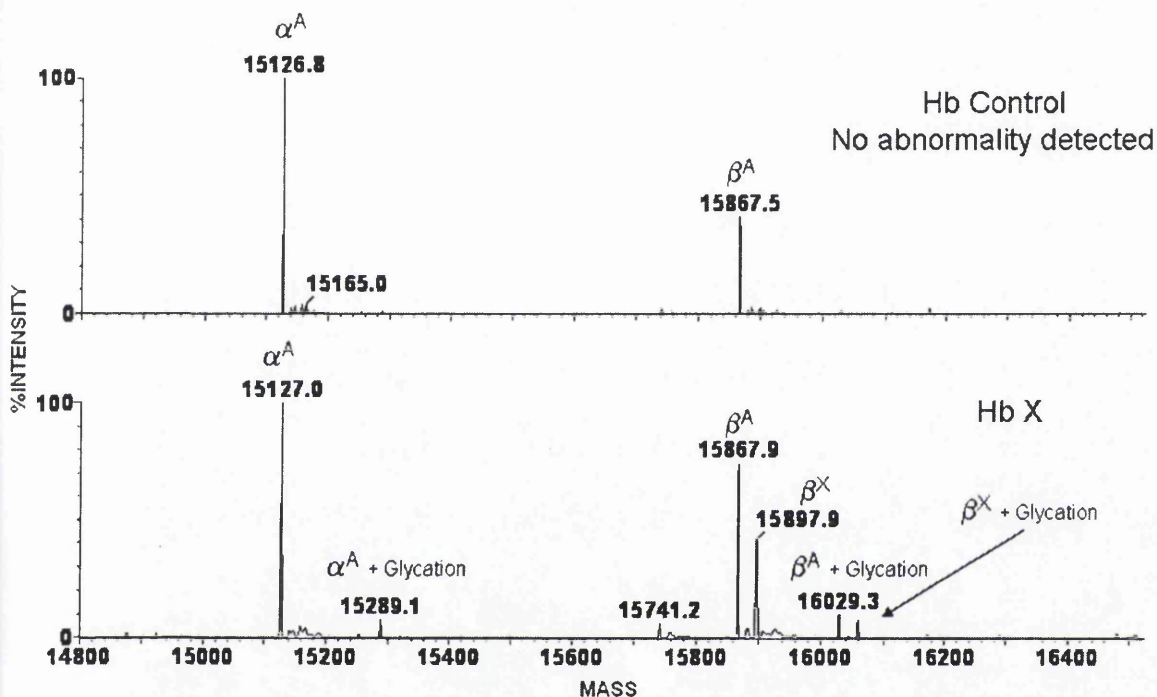
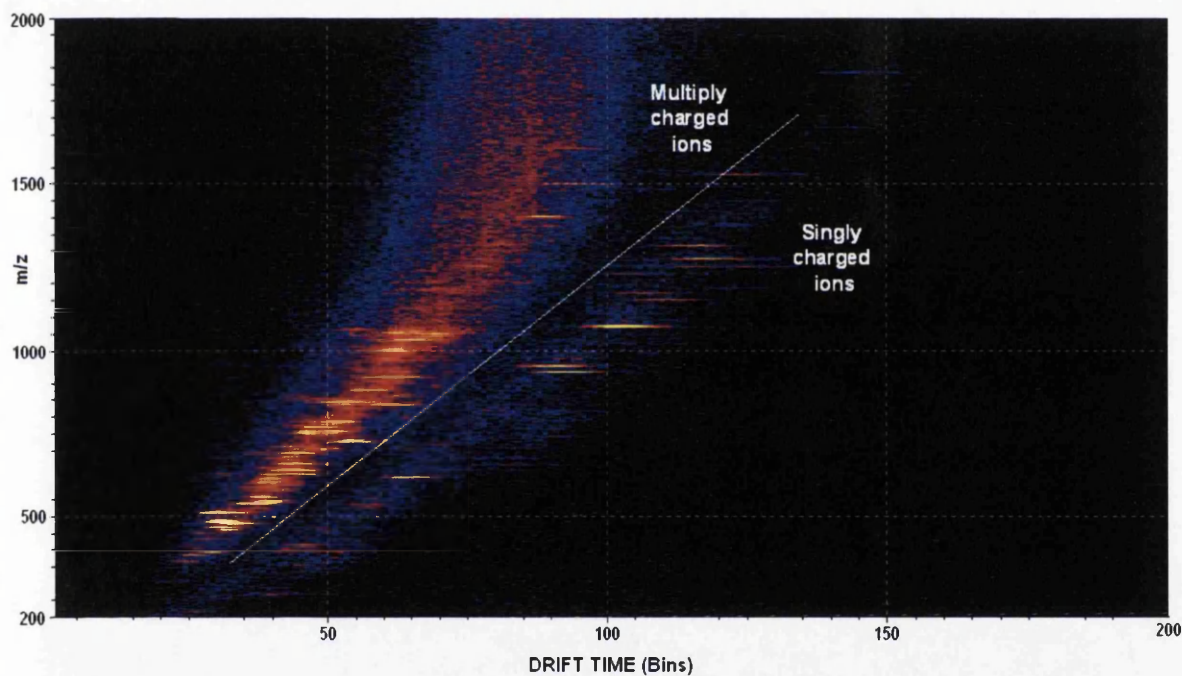


Figure 6.9: Deconvoluted mass spectra of the normal control blood sample (top) and that containing the haemoglobin variant, HbX (bottom). Normal  $\alpha$  and  $\beta$  haemoglobin chains are present in both spectra at 15126.8Da( $\alpha^A$ ) and 15867.5Da( $\beta^A$ ), and their glycated moieties at 15289.1Da and 16029.3Da, respectively. The variant blood sample contains a peak displayed at 15897.9Da and indicates the presence of a variant chain showing a +30Da shift from the normal  $\beta$  chain.

### 6.3.2.3.2. Analysis of tryptic haemoglobin peptides

The variant blood sample was digested with the endoprotease, trypsin, producing peptide fragments at known cleavage locations. This digested sample would normally appear as a complex mixture of relatively low mass ions, in addition to any multiple charged peptide ions present. The ion mobility capability of the Synapt HDMS system enables pre-analysis separation of multiply and singly charged ions, and ions of differing molecular weight.



*Figure 6.10: Mobility viewer chart illustrating the partition between multiply and singly charged ions, and the separation of ions according to mass-to-charge. The multiply charged ions have a more compact chemical structure and therefore have a smaller collision cross section, accounting for the lower mobility drift time.*

The data contained in the mobility viewer chart (see figure 6.10) can be individually selected to display the relevant mass spectrum. The extraction of the singly charged peptide ions considerably reduces the complexity of the sample spectrum allowing the



variant peptide ion to be easily identified when compared to peptides of the control haemoglobin.

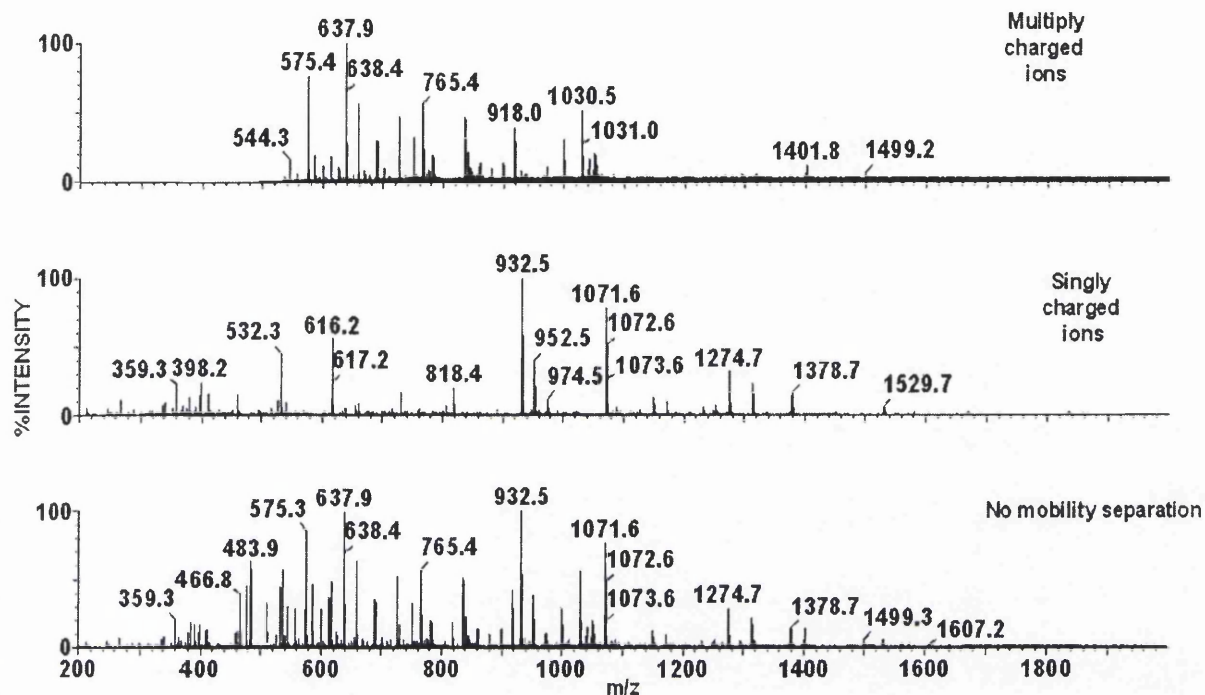


Figure 6.11: Mass spectra illustrating the clarity achieved through isolating known peptide ions of a particular charge state when ion mobility extraction is applied. This characteristic is of particular importance in identifying the modified peptide ion common to the haemoglobin variant under investigation.

To easily identify the variant tryptic peptide a predicted tryptic digestion of the normal haemoglobin chains was generated using BioLynx software (Waters). This can generate a report containing masses of the expected tryptic peptides and their multiply charged moieties, providing the amino acid sequence of the parent protein is known. This in conjunction with ion mobility extraction can identify the peptide suspected to contain the mutation. The +30Da mutation or substitution was found to

exist within the second tryptic peptide of the  $\beta$ -chain ( $\beta T2$ ) at  $m/z$  481.8 for the doubly charged ion (see figure 6.12).

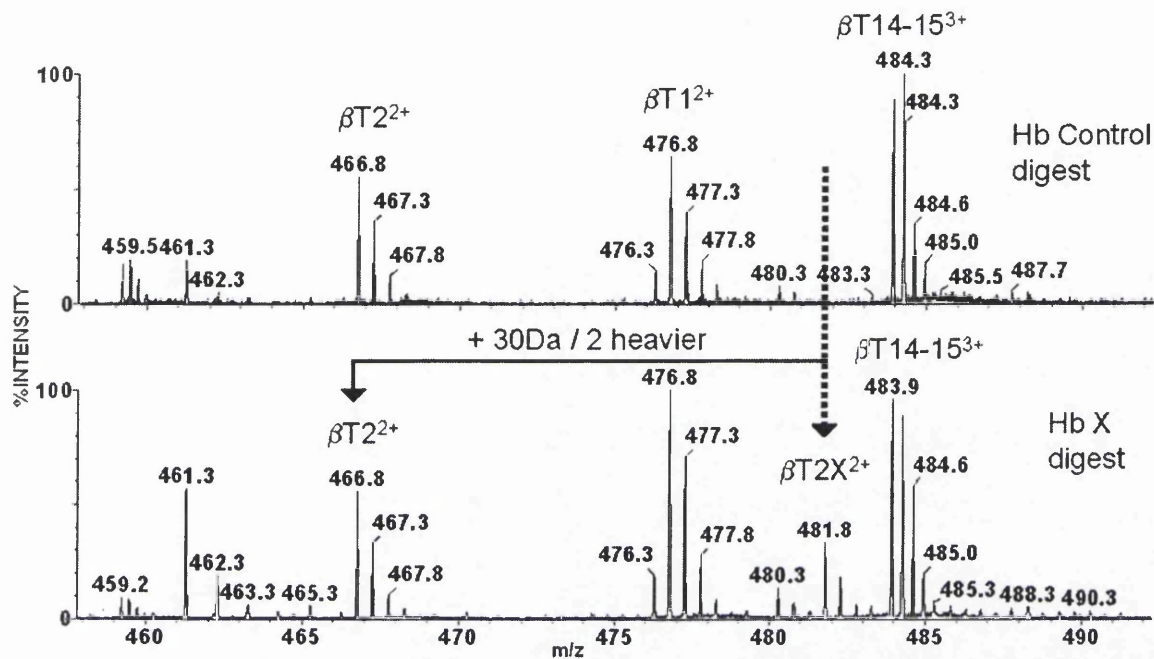


Figure 6.12: Mass spectra of tryptic peptide ions common to the normal haemoglobin digest and the variant (X) digest. The variant contains a doubly charged peptide peak at 481.8 that is not present within the normal control sample. This peak is consistent with a +30Da shift of the singly charged second tryptic peptide ion of the  $\beta$ -haemoglobin chain ( $\beta T2^{2+}$ ).

This tryptic peptide contains the following amino acids in sequence:

\*       \*       \*       \*       \*       \*       \*

(Lysine)–Serine–Alanine–Valine–Threonine–Alanine–Lysine–Tryptophan–Glycine–Lysine–(Valine)

(K) – S – A – V – T – A – L – W – G – K – (V)

where,

\* = possible site of mutation

There are a finite number of substitutions likely to cause the mutation, as shown in Table 6.2, and when compared to the above tryptic peptide sequence there are six possible locations for the mutation as indicated by the star symbol(\*). Confirmation of the location and the identification of the amino acid substitution will require further fragmentation of the tryptic peptide, in addition to the specificity of the cleavage associated with trypsin.

#### 6.3.2.3.3. Pre-mobility MS/MS investigations of tryptic peptides

These fragmentation experiments were carried out on the lowest but most intense charge state observed that illustrates the presence of the variant. This is to ensure sufficient sensitivity of product ions is obtained and enable the identification of the variant fragment ion. Again, the ion mobility facet of the Synapt HDMS system enables the isolation of the fragment product ions for the  $\beta$ -T2 tryptic peptides of interest ( $m/z$  466 and  $m/z$  481) from the multiply charged precursor ions. This extraction of data generates MS/MS mass spectra that offer enhanced clarity and ease of interpretation. It is quite clear from comparing the fragmentation MS/MS mass spectrum of the normal  $\beta$ -T2 tryptic peptide to the variant that the  $y''_8$  series ion ( $m/z$  845.50) is present in both mass spectra and shows a mass shift of +30Da within that of the variant only at  $m/z$  875.53. The  $y''_8$  ion of the normal tryptic peptide sequence involves the loss of the N-terminus serine residue with an alanine residue present at position two ( $y''_8$ ).

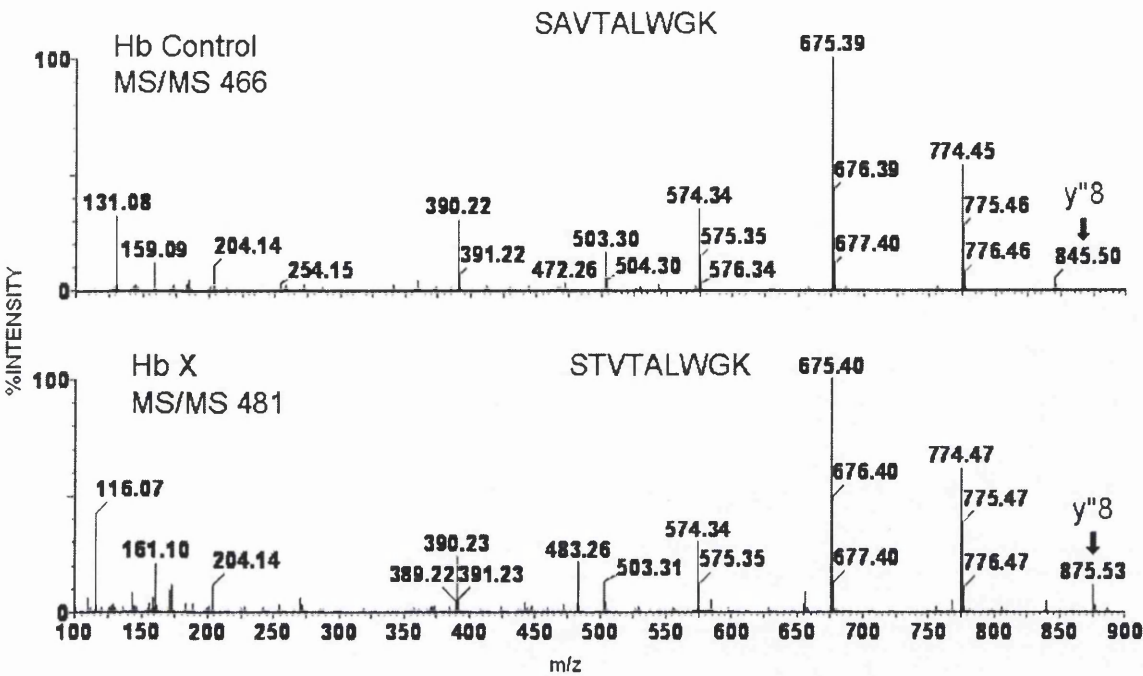


Figure 6.13: Fragmentation spectra of the tryptic peptide ion  $\beta$ -T2 within the normal control and the variant (HbX) haemoglobin chains. From the illustration it is clear that the ion at  $m/z$  875.53 is specific to the variant only and is consistent with a +30Da mass shift of the  $y^8$  ion and that observed with the intact  $\beta$  chain. A comparison of the amino acid sequences for this tryptic peptide ion indicate that the substitution is a replacement of the alanine residue at position two with a threonine residue.

According to the possible substitutions for this +30Da mass difference and a neutral charge change (table 6.2), this single base mutation must be the replacement of the alanine with a threonine residue. Further confirmation of this amino acid change is provided by inspecting the b series of fragment ions. The  $b^2$  ion of the normal control  $\beta$ T2 tryptic peptide should be observed at  $m/z$  159 accounting for serine and alanine, and the corresponding ion in the variant (serine and threonine) should be mass shifted by +30Da at  $m/z$  189. This is clearly shown in the mass spectra shown in figure 6.14.

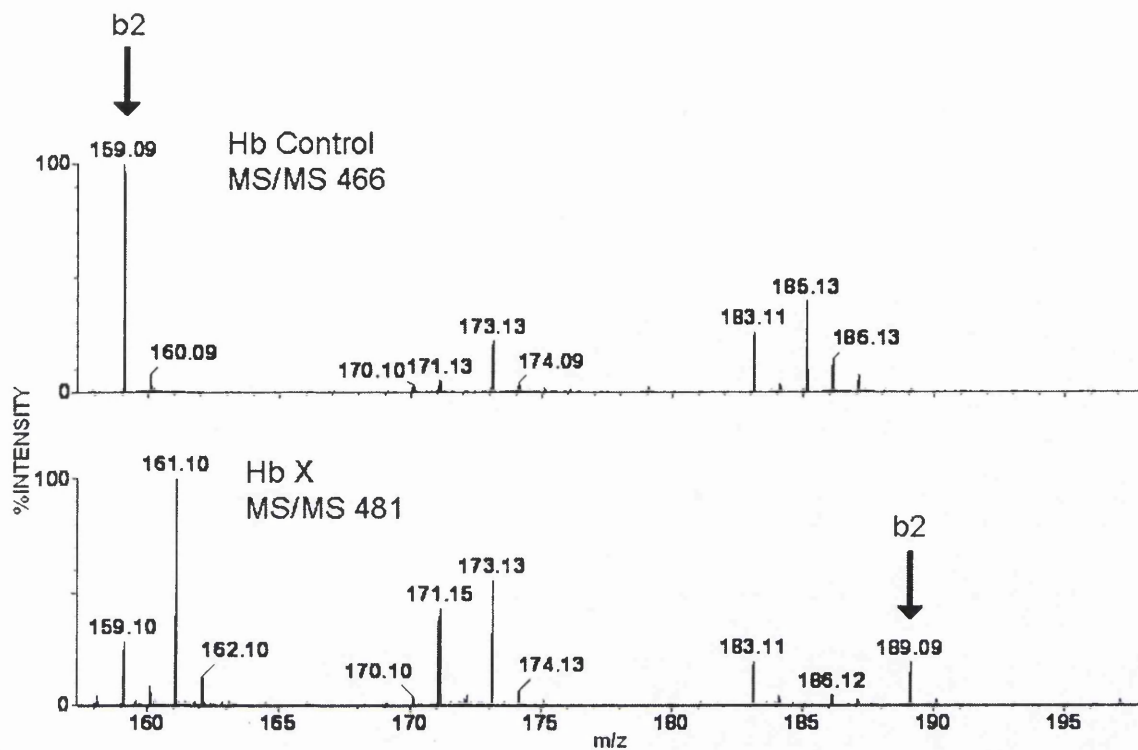


Figure 6.14: Mass spectra of the b-series of fragment ions observed for the normal and variant  $\beta$ -T2 tryptic peptide ions. The substitution of the alanine residue at position  $b_2$  ( $m/z$  159.1) with a threonine amino acid ( $m/z$  189.1) causing a +30Da mass shift is clearly shown.

#### 6.4. Conclusion

The application of the methodology developed by Green and co-workers<sup>[6]</sup> with ion mobility separation has enabled the rapid analysis and identification of a novel haemoglobin variant specific to the  $\beta$ -chain. Mass spectrometry has in this current study identified a new variant that co-elutes with glycated haemoglobin peaks present in chromatograms used for conventional post natal or general blood screening. In addition to difficult analyses such as these, mass spectrometry is capable of providing precise information regarding the characteristics of the variant unlike other

conventional screening techniques. The original mass spectrometric method has proven to be quite complicated for the type of variant highlighted within this chapter, and identification is often dependent on expert interpretation. However, through this work we have shown that the inclusion of ion mobility technology and data extraction enhances the clarity of results regarding multiple charging and differences associated with the variant. This modified protocol combining novel ion mobility and time-of-flight mass spectrometry enables the exact determination of the amino acid substitution or mutation, with its assignment to a haemoglobin chain and the specific location within the chain. In spite of the many advantages of this mass spectrometric technique the liquid chromatography screen is a more cost effective method for conventional analyses. Hence, it is likely that liquid chromatography will remain as the first step of the analytical procedure with rapid further elucidation of the blood sample carried out by mass spectrometry if required.

## References

1. Perutz, MH., *Scientific American*, 1978. 239(6): p. 68-77.
2. Schneider, RG., *Sickle cell disease*. In *Developmetns in Laboratory Diagnosis*. Mosby, St Louis, MO. 1973: p. 230-243.
3. Matsuo, T., Matsuda, H., Katakuse, I., Wada, Y., Fujita, T., and Hayashi, A., *Biomedical Mass Spectrometry*, 1981. 8: p. 25-30.
4. Green, BN., Oliver, RWA., Fallick, AM., Shackleton, CHL., Roitman, E., and Witkowska, HE., *Rapid Communications in Mass Spectrometry*, 1988. 2: p. 249-256.
5. Witkowska, HE., Bitsch, F., and Shackleton, CHL., *Haemoglobin*, 1993. 17(3): p. 227-242.
6. Wild, BJ., Green, BN., Cooper, EK., Lalloz, MRA., Erten, S., Stephens, AD., and Layton, DM., *Blood cells, molecules, and diseases*, 2001. 27(3): p. 691-704.
7. Roberts, NB., Green, BN., and Morris, M., *Clinical Chemistry*, 1997. 43(5): p. 771-778.
8. Karasek, FW., *Analytical Chemistry*, 1974. 46: p. 710A-720A.
9. Carr, TW., *Plasma chromatography*. New York: Plenum Press, 1984: p.1-39.
10. Eiceman, GA., and Karpas, Z., *Ion mobility spectrometry*. Boca Raton: CRC Press, 1994: p.1-15.
11. Kemper, PR., and Bowers, MT., *Journal of Physical Chemistry*, 1991. 95: p. 5134-5146.
12. Liu, X., Plasencia, M., Ragg, S., Valentine, SJ., and Clemmer, DE., *Briefings in functional genomics and proteomics*, 2004. 3(2): p.177-186.
13. Mason, EA., and McDaniel, EW., *Transport properties of ions in gases*. New York: John Wiley and Sons, 1988: p. 1-6.
14. Rowe, BR., Fahey, FC., Fehsenfeld, FC., and Albriton, DL., *Journal of Chemical Physics*, 1980. 73: p. 194-205.
15. Tammet, H., *Journal of Aerosol Science*, 1995. 26: p. 459-475.
16. Shvartsburg AA., and Jarrold, MF., *Chemical Physical Letters*, 1996. 261: p. 86-91.



17. Rokushika, S., Hatano, H., Baim, MA., and Hill, HH., *Analytical Chemistry*, 1985. **57**: p. 1902-1907.
18. Asbury, GR., and Hill, HH., *Journal of Microcolumn Separation*, 2000. **12**: p. 172-178.
19. Asbury, GR., and Hill, HH., *Analytical Chemistry*, 2000. **72**: p. 580-584.
20. Dwivedi, P., Wu, C., Matz, LM., Clowers, BH., Siems, WF., and Hill, HH., *Analytical Chemistry*, 2006. **78**: p. 8200-8206.
21. McDaniel, EW., and Moseley, JT., *Physical Reviews A*, 1971. **3**: p. 1040-1044.
22. Kemper, PR., and Bowers, MT., *Journal of American Society of Mass Spectrometry*, 1990. **1**: p. 197-207.
23. Giles, K., Pringle, SD., Worthington, KR., Little, D., Wildgoose, JL., and Bateman, RH., *Rapid Communications in Mass Spectrometry*, 2004. **18**: p. 2401-2414.
24. Bahr, R., Gerlich, D., Teloy, E., *Verhandl. DPG (VI)*, 1969. **4**: p.343-348.
25. Gerlich, D., *State-selected and state-to-state ion-molecule reaction dynamics, Part 1: Experiment*. Vol LXXXII, Ng CY, Baer M (eds). John Wiley and Sons: New York, 1992: p. 1-176.
26. Luca, A., Sclemmer, S., Cermak, I., and Gerlich, D., *Review of Scientific Instruments*, 2001. **72**: p.2900-2908.
27. Skilling, J., *Maximum Entropy and Bayesian Methods*. Kluwer Academic, 1989: p. 45.
28. Pringle, SD., Giles, K., Wildgoose, JL., Williams, JP., Slade, SE., Thalassinou, K., Bateman, RH., Bowers, MT., and Scrivens, JH., *International Journal of Mass Spectrometry*, 2007. **261**: p. 1-12.
29. Bookchin, RM., and Gallop PM., *Biochemical Biophysical Research Communications*, 1968. **32**: p. 86-93.
30. Bunn, HF., Haney DN., Gabbay KH., and Gallop PM., *Biochemical Biophysical Research Communications*, 1975. **67**: p. 103-109.



## CHAPTER 7:

### Conclusions

The primary objective of this work was to introduce and develop novel or existing techniques for the identification of new biomarkers within a range of biological matrices by mass spectrometry. Each biological sample was found to have inherent challenges for using modern mass spectrometric methods. These included accessing and ionising biomolecules within paraffin embedded tissue and identifying uremic analytes within a highly 'buffer-containing' dialysate sample. These difficulties were overcome by improvements in sample preparation and in some cases involved the use of existing protocols for completely novel applications.

The examination of haemodialysate solution has proved most fruitful in identifying new candidate biomarkers. Published research has included monitoring levels of known uremic solutes within other biological matrices such as blood, serum or urine and those involving dialysate have employed UV spectrophotometry with poor specificity<sup>[1]</sup>. We have shown that it is possible to detect both known and novel uremic solutes reproducibly within haemodialysate, a biological matrix previously deemed unsuitable for liquid chromatography-mass spectrometry (LC-MS). This developed methodology has been validated, which included stability and reproducibility investigations of the sample to test robustness. This highlighted a previously unrecorded thermally labile nature of some uremic solutes present in the dialysate solution. A total of 15 known uremic toxins and 6 thermally stable novel analytes have been detected and putative structural assignments made for 4 of the novel solutes. These have been named 5-(amino-1,2,-dihydroxy-ethyl)-3-nitrosooxy-[1,2,4]trioxine-3,6-diol, 2-(5,6-diamino-6-diazenyl-cyclohex-1-enyl)-2-hydroxy-

acetimidic acid, *N*-[2-(7-hydroxy-3-methyl-oxatahydro-imidazo[1,5- $\alpha$ ]pyridine-6-yl)-2-oxo-acetyl]-guanidine, and 3-(6-hydroxy-cyclohexa-1,3-dienyl)-2-imino-3-oxopropionaldehyde. Further confirmation of these identities should include the synthesis of standard reference materials of the novel solutes and extensive fragmentation studies. Once obtained these standards can be used to quantitate the amount excreted during haemodialysis and examine the true affect of exercise on haemodialysis adequacy. Overall, exercise appeared to alter the level of excretion and the chemical nature of individual uremic solutes dictated their removal during the dialysis treatment. We identified that the highly polar conventional biomarkers, urea and creatinine, are not representative of non-polar analyte excretion. Our suggestions to improve this technique include using a non-polar additive within the dialysate concentrate such as an inert surfactant, or attaching a non-polar polymer to the external side of a conventional dialysis membrane. However, this will require many years of clinical trials before it is considered as standard element of haemodialysis treatment. Hence, as a short term measure we believe that it is essential for adequacy to be monitored not only using polar analytes such as urea and creatinine, but solutes of a range of polarities, ensuring an unbiased measurement of dialysis performance.

The MALDI mass spectrometric analysis of whole tissue sections, in particular those that are paraffin embedded, posed a new range of challenges. Current MALDI matrices are unable to penetrate deep within the tissue limiting their use to imaging the surface only<sup>[2]</sup>. We have evaluated a range of novel dansylated MALDI matrices for this purpose which can be detected by both MALDI mass spectrometry and fluorescence spectroscopy. This was to aid in locating the MALDI matrix compound following application to the tissue section. Each dansylated MALDI matrix showed better penetration into the tissue sections, yet maintaining fluorescence detection, than

the standard MALDI matrices CHCA, sinapinic acid and DHB. Of these novel MALDI matrices dansylhydrazine proved most successful in ionising proteins and peptides by forming a protonated molecule and matrix-analyte adducts. These additional mass shifted peaks, when included in a tryptic peptide database search, can improve the probability of the original protein or peptide identification<sup>[3]</sup>. In comparison to CHCA, dansylhydrazine did not perform as well for both intact proteins or peptides and those subject to tryptic digestion. Interestingly, CHCA did show a potential to be identified on tissue using fluorescence detection. This work suggests that we have the potential to obtain a total image of frozen tissue by using CHCA and dansylhydrazine in combination, to ionise proteins or peptides at the tissue surface or at depth, respectively. Unfortunately, it appears that dansylhydrazine is unsuitable for the developed preparation protocols with paraffin embedded sections. Therefore further work is required for this total imaging principle to be applied to both frozen and archived paraffin embedded tissue.

Finally we have illustrated the compatibility of an existing protocol for identifying novel haemoglobin variants in blood samples<sup>[4]</sup> with a new ion mobility time-of-flight mass spectrometer, the Synapt HDMS system (Waters, MA, USA). We have identified a new variant that co-elutes with glycated haemoglobin peaks present in chromatograms used for conventional post natal or general blood screening. The original mass spectrometric method has proven to be quite complicated for the type of variant discovered, and identification is often reliant on expert interpretation. However, the inclusion of ion mobility technology and data extraction enhances the clarity of the results regarding multiple charging and variant characteristics. This enabled the exact determination of the amino acid substitution or mutation, with its assignment to a haemoglobin chain and the specific location within the chain.

The premise of this whole body of work was to identify novel solutes with the potential to act as a biomarker within a range of biological matrices. Techniques for novel applications have been introduced resulting in the examination of samples previously considered inaccessible for modern mass spectrometric investigations. We have identified candidate novel biomarkers within 2 out of 3 biological samples tested. Protocols developed for tissue imaging have proved promising in achieving complete coverage of the biological matrix but require additional work for archived sections.

## References

1. Senftlber, FC.; Halline, AG.; Veering, H.; Dayton, DA., *Clinical Chemistry*, 1976. **22**: p. 1522-1527.
2. Schwartz, SA., Reyzer, ML., Caprioli, RM., *Journal of Mass Spectrometry*, 2003. **38**: p. 699-708.
3. Park, S-J., Song, J-S., Kim, H-J., *Rapid Communications in Mass Spectrometry*, 2005. **19**: p. 3089-3096.
4. Wild, BJ., Green, BN., Cooper, EK., Lalloz, MRA., Erten, S., Stephens, AD., and Layton, D. M., *Blood cells, molecules, and diseases*, 2001. **27(3)**: p. 691-704.