The Development of Mass Spectrometry Based Approaches to Monitor Protease Activity in Biological Fluids

A thesis submitted to The University of Manchester of the degree of Doctor of Philosophy in the Faculty of Medical Sciences 2012

David N. Potier
School of Medicine
Contents

Contents ........................................................................................................................................... 2

List of Figures .................................................................................................................................... 9

List of Tables ..................................................................................................................................... 13

Abstract ............................................................................................................................................ 15

Declaration ....................................................................................................................................... 16

Copyright Statement .................................................................................................................... 16

List of Abbreviations .................................................................................................................... 17

Acknowledgements ....................................................................................................................... 21

1 Chapter 1: Introduction ................................................................................................................ 22

1.1 Project Background .................................................................................................................. 22

1.2 Cancer Classification ............................................................................................................... 22

1.3 Stages of Cancer .................................................................................................................... 23

1.4 Current Diagnosis ................................................................................................................... 24

1.5 Risk Factors Associated with Cancer ....................................................................................... 24

1.6 Molecular Causes of Cancer .................................................................................................. 25

1.7 Biomarkers ............................................................................................................................. 27

1.8 Personalised Medicine ............................................................................................................. 27

1.9 Enzymes .................................................................................................................................. 28

1.10 Proteases .................................................................................................................................. 29

1.10.1 Coagulation ....................................................................................................................... 31

1.10.2 WNT Signalling .................................................................................................................. 35

1.11 Enzyme Kinetics ..................................................................................................................... 39

1.12 How do Proteases Enter the Bloodstream? ........................................................................... 44

1.12.1 Apoptosis .......................................................................................................................... 45

1.12.2 Necrosis ............................................................................................................................ 48

1.12.3 Secondary Necrosis ........................................................................................................... 49
1.13 Mass Spectrometry ........................................................................................................49

1.13.1 Electrospray Ionisation ..........................................................................................51

1.13.2 Quadrupole Mass Analysers .................................................................................52

1.13.3 Ion Trap Mass Analysers ......................................................................................53

1.13.4 High Performance Liquid Chromatography (HPLC) .........................................54

1.13.5 Liquid Chromatography-Mass Spectrometry (LC-MS) ....................................55

1.13.6 Ion terminology ....................................................................................................56

1.13.7 4000 Q-Trap MS Instrument ..............................................................................58

1.14 Quantitative Mass Spectrometry ..............................................................................60

1.14.1 Selected Reaction Monitoring ............................................................................60

1.14.2 Chemical Labelling .............................................................................................62

1.14.3 Stable Isotope Labelling with Amino Acids in Cell Culture (SILAC) .................62

1.15 Project Outline ..........................................................................................................63

2 Chapter 2: A Comparative Study between Selected Reaction Monitoring and a Novel
Signal Boosting with Ion Trap Scanning (SBITS) Method to Detect Low Abundance
Analytes on a 4000 Q-TRAP Mass Spectrometer ..........................................................65

2.1 Introduction ..................................................................................................................65

2.1.1 Signal Boosting and Ion Trap Scanning (SBITS) .................................................65

2.1.2 MRM-Initiated Detection and Sequencing (MIDAS) .........................................69

2.1.3 Precursor Ion Scanning .......................................................................................69

2.1.4 Enhanced Resolution (ER) Scanning .................................................................70

2.1.5 Enhanced Product Ion (EPI) Scanning ..............................................................70

2.1.6 isobaric Tagging for Relative and Absolute Quantitation (iTRAQ) .................71

2.1.7 Project Description ..............................................................................................72

2.2 Materials & Methods ...............................................................................................74

2.2.1 Mass Spectrometry Equipment ...........................................................................74

2.2.2 isobaric Tagging for Relative and Absolute Quantitation (iTRAQ) Labelling ......74
2.2.3 StageTip Sample Purification ................................................................. 75
2.2.4 Signal Boosting & Ion Trap Scanning (SBITS) Sample Preparation for Method Development ................................................................. 75
2.2.5 SBITS Mass Spectrometry Scan Settings .................................................. 76
2.2.6 SBITS area of Fragmentation Investigation ............................................... 76
2.2.7 Q3 Acting as a Quadrupole or an Ion Trap in SBITS .................................... 77
2.2.8 Localised Scanning at Q3 in SBITS .......................................................... 77
2.2.9 Linearity Testing in SBITS ....................................................................... 77
2.2.10 Multiple Reaction Monitoring-Initiated Detection And Sequencing (MIDAS) Limit of Detection ........................................................................... 78
2.2.11 SBITS and SRM Selectivity Testing ......................................................... 79
2.3 Results ........................................................................................................ 80
2.3.1 SBITS Development and Optimisation ..................................................... 80
2.3.2 SBITS Linearity Testing and Comparison to SRM ...................................... 87
2.3.3 SBITS Selectivity Testing ....................................................................... 90
2.4 Discussion .................................................................................................... 93
2.5 Conclusion ................................................................................................... 96
3 Chapter 3: The Development and Optimisation of a Multi-Step Sample Preparation Method to Assay Protease Activity in Biological Fluids by SRM MS .................................................. 97
3.1 Introduction .................................................................................................. 97
3.1.1 Proteases ................................................................................................. 97
3.1.2 Project Description and Aim .................................................................... 98
3.1.3 Synthetic Peptides for Protease Activity Analysis .................................... 100
3.1.4 Target Peptide Recovery and Enrichment .............................................. 103
3.1.5 SRM Transitions .................................................................................... 109
3.2 Materials & Methods................................................................................... 111
3.2.1 MALDI-ToF/ToF Analysis ....................................................................... 111
3.2.2 LC-SRM Analysis .................................................................................. 111
3.2.3 Shotgun Peptidomics Analysis ................................................................. 111
3.2.4 Synthetic Peptide Design ................................................................. 112
3.2.5 Peptide Digestion ................................................................. 112
3.2.6 SRM Generation & Optimisation ....................................................... 113
3.2.7 SRM Linearity ................................................................. 113
3.2.8 Plasma Sample Preparation ............................................................... 114
3.2.9 Peptide Enrichment ................................................................. 114
3.2.10 Data Processing ................................................................. 117
3.2.11 Proof of Concept Testing ............................................................... 117
3.2.12 Sample Dilution Testing ............................................................... 119
3.2.13 HPLC Column Degradation ............................................................. 119
3.3 Results ............................................................................... 120
  3.3.1 Synthetic Peptide Design and Digestion .............................................. 120
  3.3.2 SRM Generation & Optimisation ....................................................... 123
  3.3.3 SRM Linearity ................................................................. 124
  3.3.4 Peptide Enrichment Investigation .................................................... 125
  3.3.5 Proof of Concept Testing ............................................................... 131
  3.3.6 HPLC column degradation ............................................................. 134
3.4 Discussion .............................................................................. 135
  3.4.1 Target Peptide Digestion ............................................................... 135
  3.4.2 Peptide Enrichment ................................................................. 135
  3.4.3 Proof of Concept Testing ............................................................... 135
3.5 Conclusion .............................................................................. 136

4 Chapter 4: Assaying the Activity of the Protease, Asparaginyl Endopeptidase (AEP) using SRM MS ................................................................. 138
  4.1 Introduction .............................................................................. 138
  4.1.1 Asparaginyl Endopeptidase (AEP)/Legumain .................................. 138
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.2 Aim of This Chapter</td>
<td>141</td>
</tr>
<tr>
<td>4.2 Methods and Materials</td>
<td>142</td>
</tr>
<tr>
<td>4.2.1 Chemicals</td>
<td>142</td>
</tr>
<tr>
<td>4.2.2 Mass Spectrometry Analysis</td>
<td>142</td>
</tr>
<tr>
<td>4.2.3 Fluorescence Analysis</td>
<td>142</td>
</tr>
<tr>
<td>4.2.4 AEP SRM Target Substrate Design</td>
<td>143</td>
</tr>
<tr>
<td>4.2.5 AEP SRM Target Substrate SRM Generation</td>
<td>143</td>
</tr>
<tr>
<td>4.2.6 Plasma Preparation for Pilot Experiments</td>
<td>146</td>
</tr>
<tr>
<td>4.2.7 SRM Background Levels in Plasma</td>
<td>146</td>
</tr>
<tr>
<td>4.2.8 Enzyme Activation</td>
<td>146</td>
</tr>
<tr>
<td>4.2.9 AEP SRM Target Substrate Digestion</td>
<td>146</td>
</tr>
<tr>
<td>4.2.10 Target Peptide Detection Investigation</td>
<td>147</td>
</tr>
<tr>
<td>4.2.11 Glufibrinopeptide Internal Standard SRM Generation and Optimisation</td>
<td>150</td>
</tr>
<tr>
<td>4.2.12 GluFib Internal Standard SRM Background Levels in Enriched Plasma</td>
<td>151</td>
</tr>
<tr>
<td>4.2.13 Cleaved SRM Substrate Linearity</td>
<td>151</td>
</tr>
<tr>
<td>4.2.14 Fluorescent Group Linearity in Plasma</td>
<td>152</td>
</tr>
<tr>
<td>4.2.15 AEP Fluorescent Substrate Stability in Plasma</td>
<td>152</td>
</tr>
<tr>
<td>4.2.16 AEP Activity in Plasma (Fluorescence Analysis)</td>
<td>152</td>
</tr>
<tr>
<td>4.2.17 AEP Activity in Plasma pH Buffered to pH 5.8 (SRM Analysis)</td>
<td>153</td>
</tr>
<tr>
<td>4.2.18 SRM Sample Dilution Testing</td>
<td>153</td>
</tr>
<tr>
<td>4.2.19 Is Effect of pH on the Stability of AEP Reversible?</td>
<td>153</td>
</tr>
<tr>
<td>4.2.20 AEP Limit of Detection Testing in Plasma</td>
<td>154</td>
</tr>
<tr>
<td>4.2.21 Whole Cell Lysate Sample Preparation</td>
<td>154</td>
</tr>
<tr>
<td>4.2.22 AEP ELISA Assay</td>
<td>154</td>
</tr>
<tr>
<td>4.2.23 Whole Cell Lysate AEP Activity Assay</td>
<td>155</td>
</tr>
<tr>
<td>4.3 Results</td>
<td>156</td>
</tr>
<tr>
<td>4.3.1 AEP SRM Substrate Design</td>
<td>156</td>
</tr>
</tbody>
</table>
4.3.2 AEP Substrate SRM Generation ................................................................. 157
4.3.3 Target Peptide Digestion and Detection .................................................. 159
4.3.4 GluFib Internal Standard Detection ......................................................... 169
4.3.5 Cleaved AEP Target Substrate Linearity .................................................. 171
4.3.6 AEP Fluorescence Substrate Stability in Plasma ........................................ 175
4.3.7 AEP Activity in Plasma ............................................................................. 175
4.3.8 AEP Activity in Whole Cell Lysates .......................................................... 186

4.4 Discussion ................................................................................................. 192
4.4.1 AEP Activity in Plasma Samples ............................................................... 192
4.4.2 Fluorescence/MS Comparison .................................................................. 193
4.4.3 AEP Activity in Whole Cell Lysate Samples ............................................ 195
4.4.4 Changes to the HPLC Setup for the SRM Assay ....................................... 196
4.4.5 Internal Standards .................................................................................... 198

4.5 Conclusion ............................................................................................... 198

5 Chapter 5: Assaying the Activity of the Protease, Caspase-3 using SRM MS .... 200

5.1 Introduction ............................................................................................... 200
5.1.1 Caspase-3 ............................................................................................... 200
5.1.2 Aim of This Chapter ................................................................................ 204

5.2 Methods and Materials .............................................................................. 205
5.2.1 Chemicals ............................................................................................... 205
5.2.2 Mass spectrometry Analysis ..................................................................... 205
5.2.3 Synthetic Peptide Design ........................................................................ 205
5.2.4 Synthetic Peptide SRM Generation ........................................................ 206
5.2.5 SRM Background Levels in Plasma ......................................................... 208
5.2.6 Target Peptide Digestion, Detection and SRM Optimisation ................. 208
5.2.7 Cleaved Peptide SRM Linearity ............................................................... 209
5.2.8 Caspase-3 Activity in Plasma ................................................................... 210
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.9</td>
<td>Sample Dilution Testing</td>
<td>210</td>
</tr>
<tr>
<td>5.2.10</td>
<td>Caspase-3 Assay Failure Investigation</td>
<td>211</td>
</tr>
<tr>
<td>5.2.11</td>
<td>Caspase-3 Activity with Higher Levels of Caspase-3 in Plasma and/or Shorter Incubation Times</td>
<td>213</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>215</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Caspase-3 Synthetic Peptide Target Design and SRM Transition Generation</td>
<td>215</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Cleaved Peptide SRM Linearity</td>
<td>220</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Caspase-3 Activity in Plasma</td>
<td>221</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Caspase-3 Assay Investigation</td>
<td>224</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Caspase-3 Activity in Plasma Using Higher Levels of Caspase-3 in Samples</td>
<td>235</td>
</tr>
<tr>
<td>5.3.6</td>
<td>Caspase-3 Activity in Plasma Using Shorter Incubation Times</td>
<td>236</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>238</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Caspase-3 Assay Investigation</td>
<td>238</td>
</tr>
<tr>
<td>5.5</td>
<td>Conclusion</td>
<td>239</td>
</tr>
<tr>
<td>6</td>
<td>Chapter 6: Final Conclusions and Future Work</td>
<td>242</td>
</tr>
<tr>
<td>6.1</td>
<td>Final Conclusions</td>
<td>242</td>
</tr>
<tr>
<td>6.2</td>
<td>Future Work</td>
<td>243</td>
</tr>
<tr>
<td>7</td>
<td>Chapter 7: References</td>
<td>246</td>
</tr>
<tr>
<td>8</td>
<td>Chapter 8: Publications and Conference Proceedings</td>
<td>282</td>
</tr>
<tr>
<td>8.1</td>
<td>Publications</td>
<td>282</td>
</tr>
<tr>
<td>8.2</td>
<td>Conference Proceedings</td>
<td>283</td>
</tr>
</tbody>
</table>

Word Count: 70,317
List of Figures

Figure 1.1: The induced fit model of an enzymic reaction .................................................. 28
Figure 1.2: The basic system of fibrin formation ....................................................................... 32
Figure 1.3: The coagulation cascade .......................................................................................... 34
Figure 1.4: The canonical WNT pathway .................................................................................. 36
Figure 1.5: The (a) PCP and (b) WNT-Ca$^{2+}$ pathways ......................................................... 38
Figure 1.6: The differences between apoptosis and necrosis ...................................................... 45
Figure 1.7: The caspase cascade in mammalian cells ................................................................... 47
Figure 1.8: A schematic of a generic mass spectrometry instrument .......................................... 50
Figure 1.9: The process of electrospray ionisation .................................................................... 51
Figure 1.10: Ion stability through a quadrupole mass analyser .................................................. 52
Figure 1.11: A schematic of an ion trap ..................................................................................... 54
Figure 1.12: A schematic showing how components in a mixture can be separated via LC methods ......................................................................................................................................................... 55
Figure 1.13: The terminology of different peptide fragments ..................................................... 56
Figure 1.14: How peptide fragments can be used to sequence a peptide .................................... 57
Figure 1.15: A schematic of the 4000 Q-TRAP ........................................................................ 59
Figure 1.16: An explanation of selected reaction monitoring ..................................................... 61
Figure 1.17: The sample preparation method developed in this project .................................... 64
Figure 2.1: A representation of SBITS ...................................................................................... 66
Figure 2.2: The sample preparation of SBITS .......................................................................... 67
Figure 2.3: The mass spectrometry workflow of SBITS ............................................................. 68
Figure 2.4: Enhanced product ion scanning in a 4000 Q-TRAP .............................................. 70
Figure 2.5: iTRAQ regent structure ........................................................................................... 72
Figure 2.6: The EPI scan detecting iTRAQ reporter ions in SBITS .......................................... 81
Figure 2.7: An example of the EPI scan used to sequence peptides in SBITS ............................ 82
Figure 2.8: The ER scan used to determine charge state in SBITS ........................................... 83
Figure 2.9: Using Q3 as (a) a linear ion trap or (b) a quadrupole in SBITS to detect iTRAQ reporter ions ......................................................................................................................................................... 85
Figure 2.10: (a) Using Q3 as a linear ion trap or (b) performing localised quadrupole scanning to detect iTRAQ reporter ions in SBITS ......................................................................................................................................................... 86
Figure 2.11: The SBITS workflow after optimisation ................................................................. 87
Figure 2.12: SBITS and SRM linearity testing .......................................................................... 88
Figure 2.13: (a) SBITS and (b) SRM limit of detection testing using 20 fmol and 1 fmol of target analyte respectively................................................................. 89
Figure 2.14: The selectivity of SBITS and SRM................................................................. 92
Figure 2.15: A theoretical comparison between different sample populations using SBITS . 94
Figure 3.1: The sample preparation workflow being developed in this study................... 99
Figure 3.2: SRM workflow ........................................................................................... 103
Figure 3.3: The workflow to assess intact protein removal from plasma...................... 105
Figure 3.4: The workflow to assess peptide retention in plasma..................................... 107
Figure 3.5: The data processing workflow to assess peptide retention......................... 108
Figure 3.6: The effect of D-amino acids on proteolysis using trypsin............................ 122
Figure 3.7: SRM responses from endogenous plasma peptides when detecting the peptide etaeaylGGK................................................................. 124
Figure 3.8: etaeaylGGK SRM linearity ........................................................................... 125
Figure 3.9: SDS-PAGE gel (10% acrylamide) showing intact protein removal by each peptide enrichment method................................................................. 126
Figure 3.10: Peptides and proteins identified in plasma samples enriched by filtration or acetonitrile precipitation ........................................................................ 130
Figure 3.11: Proof of concept testing............................................................................. 132
Figure 4.1: The effect of pH on (a) the stability and (b) the activity of AEP................... 139
Figure 4.2: The reaction mechanism of substrate proteolysis by a cysteine protease ......... 140
Figure 4.3: Amino acid sequence of wild type and H150A mutated AEP...................... 140
Figure 4.4: SRM responses from endogenous plasma peptides when detecting the cleaved AEP SRM substrate ........................................................................ 158
Figure 4.5: Total ion chromatogram when detecting hphFAANDVSK9p using LC-MS on a 4000 Q-TRAP ................................................................................. 159
Figure 4.6: MALDI-ToF/ToF MS of hphFAANDVSK9p .................................................. 160
Figure 4.7: MALDI-ToF/ToF MS/MS of hphFAANDVSK9p ............................................. 161
Figure 4.8: MALDI-ToF/ToF MS of (a) the eluate and (b) the flow through of hphFAANDVSK9p following StageTip purification at pH 3. ................................. 163
Figure 4.9: New England Peptides LC-UV trace of hphFAANDVSK9p ......................... 163
Figure 4.10: Offline LC-UV trace of hphFAANDVSK9p .................................................. 164
Figure 4.11: MS/MS data from a Q-STAR XL of hphFAANDVSK9p ............................. 165
Figure 4.12: MS/MS of hphFAANDVSK9p from LC-IDA analysis on a 4000 Q-TRAP.. 166
Figure 4.13: MS/MS of hphFAANDVSKhph dissolved in PBS from LC-IDA analysis on a 4000 Q-TRAP ................................................................. 166
Figure 4.14: LC-SRM analysis of hphFAAN and DVSKhph ........................................ 168
Figure 4.15: (a) SRM responses from endogenous peptides, (b) when detecting GluFib and (c) SRM detection of GluFib at 250 fmol ................................................................. 171
Figure 4.16: Cleaved AEP substrate linearity ............................................................. 173
Figure 4.17: DVSKhph SRM linearity data normalised against the GluFib internal standard ................................................................. 174
Figure 4.18: Fluorescent substrate stability in plasma over a 24 hour period. ............. 175
Figure 4.19: AEP activity in plasma at pH 5.8 ............................................................ 177
Figure 4.20: Rate of reaction of AEP in plasma at pH 5.8 when being analysed by (a) fluorescence and (b) SRM assays generated from the data produced in Figure 4.16 .......... 181
Figure 4.21: AEP activity in plasma at pH 5.8 normalised against GluFib .................. 183
Figure 4.22: Restoring AEP activity in plasma by altering pH .................................. 185
Figure 4.23: AEP ELISA assay linearity curve ............................................................ 186
Figure 4.24: AEP activity in WCL samples .............................................................. 188
Figure 4.25: Rate of reaction of AEP in WCL samples for the (a) fluorescence and (b) SRM assays .............................................................................................................. 190
Figure 5.1: Caspase-3 crystal structure when bound to a tetrapeptide aldehyde inhibitor ... 200
Figure 5.2: Caspase-3 activity at different pHs ......................................................... 201
Figure 5.3: Caspase-3 causing tumour regrowth ...................................................... 203
Figure 5.4: SRM responses from endogenous plasma peptides when detecting (a) the N-terminal and (b) the C-terminal form of the cleaved caspase-3 substrate ......................... 217
Figure 5.5: Cleaved caspase-3 substrate SRM detection ........................................ 219
Figure 5.6: Cleaved caspase-3 substrate linearity ..................................................... 220
Figure 5.7: LC-SRM of cleaved caspase-3 substrate when assaying caspase-3 activity in plasma .............................................................................................................. 223
Figure 5.8: The individual stages of the sample preparation method to be investigated ...... 226
Figure 5.9: MALDI-ToF/ToF data showing the caspase-3 substrate has not degraded over time .............................................................................................................. 227
Figure 5.10: MALDI-ToF/ToF data showing the caspase-3 substrate at each stage of sample preparation .............................................................................................................. 229
Figure 5.11: MALDI-ToF/ToF data showing the effect of formic acid on the caspase-3 substrate .............................................................................................................. 230
Figure 5.12: LC-SRM data investigating carry over in this assay .................................................. 232
Figure 5.13: MALDI-ToF/ToF data investigating the interaction of caspase-3 and its substrate post-acetonitrile precipitation. ........................................................................................................................................ 232
Figure 5.14: MALDI-ToF/ToF data of samples previously analysed by LC-SRM ............ 235
List of Tables

Table 1.1: The main classes of cancer, along with specific examples encompassed by each class.................................................................................................................................................. 22
Table 1.2: Five of the most common types of cancer diagnosed, and the number of deaths caused by each type in the UK in 2008.................................................................................................................................. 23
Table 1.3: A more detailed description of the different levels in Roman numeral cancer staging.............................................................................................................................................. 24
Table 1.4: How proteases are linked to a variety of diseases, including various types of cancer........................................................................................................................................................ 30
Table 2.1: The amount of 114 and 117 iTRAQ labelled GluFib target peptide present in each sample in SRM and SBITS linearity testing........................................................................................................ 78
Table 2.2: The amount of iTRAQ 114 labelled GluFib in MIDAS testing............................................................................................................................... 79
Table 3.1: Optimised SRM transitions to detect etaeaylGGK............................................................................................................................................ 113
Table 3.2: The relative amounts of etaeaylGGK used in SRM linearity testing................................................................. 114
Table 3.3: The amounts of trypsin added to plasma in proof of concept testing................................................................. 118
Table 3.4: The time points at which samples were taken in proof of concept testing................................................................. 118
Table 3.5: Target peptide recoveries of plasma samples enriched by filtration and acetonitrile precipitation.................................................................................................................................................. 127
Table 3.6: Statistical analysis of peptide retention data.................................................................................................................................................. 128
Table 3.7: Statistical analysis of the data shown in Figure 3.11 when analysed by single factor ANOVA testing .................................................................................................................................................. 133
Table 3.8: Statistical analysis of the data shown in Figure 3.11 when analysed by single factor ANOVA testing, excluding data from samples containing 200 ng trypsin ........................................................................................................ 133
Table 4.1: SRM transitions to detect hphFAANDVSKhph, hphFAAN and DVSDKhph........................................................................................................................................ 144
Table 4.2: Offline LC-UV gradient to detect hphFAANDVSKhph........................................................................................................................................ 149
Table 4.3: Online LC-UV gradient to detect hphFAAN and DVSDKhph............................................................................................ 149
Table 4.4: Optimised SRM transitions to detect DVSDKhph................................................................................................. 150
Table 4.5: Optimised SRM transitions to detect GluFib.................................................................................................................. 151
Table 4.6: The amounts of DVSDKhph used in SRM linearity testing.................................................................................................................. 151
Table 4.7: The amounts of rAEP added to plasma during proof of concept testing............................................................................................. 152
Table 4.8: Sample dilution sample amounts.................................................................................................................................................. 153
Table 4.9: Known AEP cleavages in the MEROPS database (10\textsuperscript{th} July 2010)........................................................................................................ 156
Table 4.10: Single factor ANOVA analysis of the data produced in Figure 4.19 (b) at each time point. .................................................................................................................. 178
Table 4.11: Calculated $k_2$ values for each time point when determining rAEP activity in plasma. ........................................................................................................ 179
Table 4.12: GluFib SRM responses when monitoring AEP activity in plasma. .......... 183
Table 4.13: Single factor ANOVA analysis of the data produced in Figure 4.24 (b) at each time point. ........................................................................................................ 189
Table 4.14: A comparison of the fluorescence and SRM based assays used to quantify AEP activity in plasma. .................................................................................................................. 194
Table 5.1: SRM transitions to detect hphDEVDGAGShph, hphDEVD and GAGShph. ..... 206
Table 5.2: Gradient conditions to detect hphDEVD and GAGShph. ......................... 209
Table 5.3: Optimised SRM transitions to detect hphDEVD. ........................................ 209
Table 5.4: The amounts of hphDEVD used in linearity testing. ................................. 209
Table 5.5: The amounts of caspase-3 added to plasma in proof of concept testing. ........ 210
Table 5.6: Sample dilution sample amounts. ................................................................. 211
Table 5.7: Overnight substrate storage conditions to investigate the effect of acid on hphDEVDGAGShph stability. ................................................................................. 211
Table 5.8: Samples prepared post-acetonitrile precipitation to investigate hphDEVDGAGShph stability. ........................................................................................................ 212
Table 5.9: Samples prepared to investigate each stage of the sample preparation method on the stability of hphDEVDGAGShph. ......................................................... 213
Table 5.10: The samples previously analysed by LC-SRM analysed by MALDI-ToF/ToF. 213
Table 5.11: The amount of caspase-3 added to plasma in adjusted proof of concept testing. .................................................................................................................. 214
Table 5.12: The amounts of caspase-3 added to plasma in proof of concept testing using increased levels of caspase-3. ................................................................. 214
Table 5.13: Known caspase-3 cleavages in the MEROPS database (10th July 2010). ....... 215
Table 5.14: SRM responses of ten randomly analysed samples in proof of concept testing. 221
Table 5.15: SRM responses of ten randomly analysed samples in adjusted proof of concept testing. .................................................................................................................. 236
Table 5.16: SRM responses of ten randomly analysed samples in proof of concept testing using increased levels of caspase-3. ......................................................................... 237
Abstract
The University of Manchester, Doctor of Philosophy (PhD) degree, 25th September 2012
David N. Potier
The Development of Mass Spectrometry Based Approaches to Monitor Protease Activity in Biological Fluids

When treating patients with cancer, the ability to predict a patient’s response to treatment is an important tool to allow therapy to be tailored for best outcome. Therefore, the need exists for a test to forecast a patient’s response using a sample that is readily accessible, and provides an accurate reflection of a patient’s response to a disease or treatment. Profiling biological fluids, such as plasma or urine, has gained considerable interest in recent years. This is because these fluids are readily available and are expected to provide an accurate representation of a patient’s response to treatment. As such, much effort has been put into finding biomarkers or prognostic indicators.

Abnormal protease activity has been linked to the progression of cancer due to their involvement in several processes vital to the survival and proliferation of the disease. These include metastasis, resistance to apoptosis and angiogenesis, amongst others. In addition, dysregulated protease activity has been linked to poor response to chemotherapy as well as tumour regrowth following radiotherapy. Therefore, an increased understanding regarding the activity of a patient’s proteases may provide the clinician with more information as to how best to treat the patient. Therefore, monitoring protease activity has been suggested as a potential marker to predict a patient’s response to cancer treatment.

Most enzyme activity assays are currently performed by fluorescence spectroscopy. However, these workflows suffer from limited sensitivity and linear range. Therefore, an alternative, more sensitive assay is required. Mass spectrometry (MS) is a highly sensitive analytical technique routinely used to quantify changes in biological systems. As such, MS has the potential to be used in enzyme activity assays.

This study illustrates the development of a novel MS based method to monitor the activity of target enzymes in plasma; specifically asparaginyl endopeptidase (AEP) and caspase-3 using mass spectrometry. These enzymes have been linked to poor chemotherapeutic response in childhood leukaemia and tumour regrowth post-radiotherapy respectively. This project will describe the development and optimisation of each stage of a five step sample preparation and analysis method. This includes the design of enzyme substrates designed to be cleaved by the target enzyme, whilst reducing the effect of other enzymes acting on this substrate, how best to enrich samples for these target peptides, as well as determining the best MS technique to monitor these peptides.

In addition, this project describes a comparison between this assay and an existing fluorescence assay when monitoring AEP activity in biological samples such as plasma and whole cell lysates. The application of this method in quantifying caspase-3 activity in plasma samples is also examined.
Declaration
No portion of the work referred to in this thesis has been submitted in support of an application of another degree or qualification of this or any other university or other institute of learning.
I hereby state that this thesis and its contents, to the best of my knowledge, except where due reference is made, is original and my own work.

Copyright Statement
The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on presentation of Theses.
List of Abbreviations

a.u. – Arbitrary unit
AC – Alternating Current
AEBSF – 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride
AEP – Asparaginyl Endopeptidase/Legumain
ALL – Acute Lymphoblastic Leukaemia
ALT – Alanine Transaminase
AML – Acute Myeloid Leukaemia
amu – Atomic Mass Unit
amu/s – Atomic Mass Units per Second
AP-MALDI – Atmospheric Pressure Matrix Assisted Laser Desorption/Ionisation
APMA – 4-aminophenylmercuric acetate
Asp – Aspartic Acid
AQUA – Absolute Quantitation
AXIN1 – Axis inhibition protein 1
an – a ion containing n number of amino acids
bn – b ion containing n number of amino acids
cn – c ion containing n number of amino acids
xn – x ion containing n number of amino acids
yn – y ion containing n number of amino acids
zn – z ion containing n number of amino acids
Ca2+ – Calcium Ion
CHAPS – 3-[(3-Cholamidopropyl)dimethylammonio]1-propanesulfonate
CHCA – α-cyano-4-hydroxycinnamic acid
CI – Chemical Ionisation
CID – Collisionally Induced Dissociation
CML – Chromic Myeloid Leukaemia
cps – Counts per second
CT – Computerised Tomography
CID – Collisionally Induced Dissociation
Da – Dalton
DC – Direct Current
DNA – Deoxyribonucleic Acid
EDTA – Ethylene Diamine Tetraacetic Acid
EI – Electron Impact
ELISA – Enzyme-Linked Immunosorbent Assay
EMS – Enhanced Mass Scan
EPI – Enhanced Product Ion Scan
ESI – Electrospray Ionisation
eV – Electron Volt
FDA – Food and Drug Administration
fmol – Femtomole
FWHM – Full Width at Half Maximum
GranB – Granzyme B
HIV – Human Immunodeficiency Virus
HMWK – High Molecular Weight Kininogen
HPLC – High Performance Liquid Chromatography
IDA – Information Dependent Acquisition
iTRAQ - isobaric Tagging for Relative and Absolute Quantitation
kDa – Kilo-Dalton
LC – Liquid Chromatography
LC-MS – Liquid Chromatography Mass Spectrometry
LIT – Linear ion trap
M (in TNM Staging) – Description of the level of metastasis of the tumour
m/z – Mass to Charge Ratio
MALDI – Matrix Assisted Laser Desorption/Ionisation
MIDAS – MRM-Initiated Detection and Sequencing
mmol – Millimole
MRI – Magnetic Resonance Imaging
MS – Mass Spectrometry/Spectrometer
ms – Millisecond
MS/MS – Tandem Mass Spectrometry
N (in TNM Staging) – Description of the lymph nodes involved in the tumour
NH₂ – Amine group
nm – Nanometre
NMR – Nuclear Magnetic Resonance
PAGE – Polyacrylamide gel electrophoresis
PBS – Phosphate Buffered Saline (pH 7.4)
PCP – Planar Cell Polarity pathway
PET – Positron Emission Tomography
PMD – Photomultiplier Detector
pmol – Picomole
pmol/µl – Picomoles per microlitre
PTM – Post Translational Modification
RPR – Rapid Plasma Reagin
Q0 – Area in which ions can be held if the linear ion trap is full
Q1 – The 1st quadrupole in a mass spectrometer
Q2 – The 2nd quadrupole in a mass spectrometer
Q3 – The 3rd quadrupole in a mass spectrometer
QC – Quality Control
r.f. – Radio Frequency
rAEP – Recombinant human AEP
rpm – Revolutions per minute
S/N – Signal to Noise ratio
SBITS – Signal Boosting and Ion Trap Scanning
SDS – Sodium Dodecyl Sulphate
SDS-PAGE – Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SRM – Selected reaction monitoring
T (in TNM Staging) – Measure of tumour size in TNM staging
TCC – The Caspase Cascade
TEMED – Tetramethylethylenediamine
TFA – Trifluoroacetic acid
TNM – Classification of Malignant Tumours staging system
ToF – Time of Flight
Tris – Tris (hydroxymethyl)aminomethane
UK – United Kingdom
µl/min – Microlitres per Minute
UV – Ultra Violet
v/v – Volume to volume
V – Volt
WCL – Whole Cell Lysate
WHO – World Health Organisation
XIC – Extracted Ion Chromatogram
Acknowledgements

Firstly, I want to thank my supervisor, Prof. Tony Whetton for taking me on. I would especially like to thank Tony for his continual advice, support and feedback on my project. I also want to thank Prof. Vaskar Saha for acting as my advisor and allowing me to work in his lab and Dr. Robert Graham for the critical reading of this thesis.

I would like to offer the greatest thanks to Dr. Michael Walker, Dr. John Griffiths and Dr. Duncan Smith. Thank you for teaching me pretty much everything I know about mass spec as well as all your help and advice in areas ranging from technical expertise to sharing your biological knowledge with me, ensuring I did the right experiments at the right time and in the right way – I could not have done it without you! I owe the three of you a tremendous amount and am extremely grateful for the incredible support and advice you’ve all given me over the past four years! I’ll never forget the three of you and how you’ve helped me during my time here.

I also want to express my gratitude to Dr. Richard Unwin, Dr. Andrew Williamson and Yvonne Connolly for their knowledge and insight which saved many wasted hours in the lab, as well as offering me advice about the finer details of my work which has helped me get the results I have. I’d also like to thank Dr. Mark Holland for all his help, support, advice and insight into the work I performed in his lab.

I would like to thank the members of the SCALPL and Meyer groups, both past and present; Thanks for making me feel welcome and putting up with a chemist trying to do a little bit of biology!

I acknowledge BBSRC & Philips for funding this work.

Finally, I would like to thank my friends and family, especially my grandparents, for the influence they’ve had on me. If it wasn’t for you, I wouldn’t have been doing this for the past four years. I’ll never forget you and this thesis is dedicated to you.
1 Chapter 1: Introduction

1.1 Project Background
Cancer is a disease causing hundreds of thousands of deaths every year. In 2008, the number of people diagnosed with cancer was 408,381 in the UK, with 156,723 deaths from the disease being reported that year (1). The rate of incidence of cancer is increasing by around 1.5% each year (1). The World Health Organisation (WHO) has shown that the number of cancer related deaths worldwide in 2008 was 7.6 million. This equates to 13% of all deaths in that year, with this number set to rise to 12 million by the year 2030 (2). Despite £4.5 billion being spent on diagnosing and treating cancer in England alone in 2008 (1) the number of deaths is increasing (3). Therefore better diagnosis and treatment is required.

1.2 Cancer Classification
Cancer is classed by the type of cell that the tumour resembles or its site of origin. From this, the origin of the tumour is inferred (4-8). When classifying cancer, there are five main categories (Table 1.1).

Table 1.1: The main classes of cancer, along with specific examples encompassed by each class.

<table>
<thead>
<tr>
<th>Type of Cancer</th>
<th>Origin</th>
<th>Examples of this Cancer</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carcinoma</td>
<td>Epithelial Cells</td>
<td>Breast, Lung, Colon</td>
<td>(4, 9, 10)</td>
</tr>
<tr>
<td>Sarcoma</td>
<td>Connective Tissue</td>
<td>Osteosarcoma, Angiosarcoma</td>
<td>(11-13)</td>
</tr>
<tr>
<td>Lymphoma/Leukaemia</td>
<td>Hematopoietic Cells</td>
<td>Acute Lymphoblastic Leukaemia (AML), Chronic Myeloid Leukaemia (CML)</td>
<td>(14-16)</td>
</tr>
<tr>
<td>Germ Cell Tumour</td>
<td>Pluripotent Cells</td>
<td>Testicular, Ovarian</td>
<td>(17-19)</td>
</tr>
<tr>
<td>Blastoma</td>
<td>Immature/Embryonic Tissue</td>
<td>Nephroblastoma, Medulloblastoma</td>
<td>(20-22)</td>
</tr>
</tbody>
</table>
Table 1.2: Five of the most common types of cancer diagnosed, and the number of deaths caused by each type in the UK in 2008. *Taken from UK cancer statistics, reference (1).*

<table>
<thead>
<tr>
<th>Number of Diagnoses in 2008</th>
<th>Breast Cancer</th>
<th>Lung Cancer</th>
<th>Colorectal Cancer</th>
<th>Prostate Cancer</th>
<th>Pancreatic Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>48,034</td>
<td>40,806</td>
<td>39,991</td>
<td>37,051</td>
<td>8,085</td>
<td></td>
</tr>
<tr>
<td>Percentage of Total Diagnoses</td>
<td>16</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>Rank in Highest Number of Diagnoses</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Number of Deaths</td>
<td>12,116</td>
<td>35,261</td>
<td>16,259</td>
<td>10,168</td>
<td>7,781</td>
</tr>
<tr>
<td>Percentage of Total Deaths</td>
<td>8</td>
<td>22</td>
<td>10</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Rank in Highest Number of Deaths</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

### 1.3 Stages of Cancer

Staging is a method clinicians use to describe the size and progression of a cancer or tumour. There are different types of staging system for different cancers, but the TNM staging system is used to describe the majority of carcinomas and sarcomas (23-33). The TNM system is used to diagnose cancers including breast cancer (34, 35), lung cancer (36-38) and kidney cancer (39-41) amongst others (42-44). The TNM staging system is divided into three sections. The first section (T) is a measure of the tumour size and whether it has invaded its surrounding tissue. The second section (N) describes the lymph nodes involved in the tumour. This is split into four categories ranging from N0, meaning that tumour cells are not present in regional lymph nodes, to N3, meaning the tumour has spread to distant and/or multiple lymph nodes. The third section (M) describes the degree of metastasis of the tumour. This has two categories, M0, meaning no distant metastasis has taken place, or M1, meaning the tumour has spread to distant organs (45).

However, cancer progression is commonly grouped using Roman Numeral Staging. These range from Stage 0 to Stage IV. This staging is described in more detail in Table 1.3.
Table 1.3: A more detailed description of the different levels in Roman numeral cancer staging. Taken from Mountain et. al, reference (46).

<table>
<thead>
<tr>
<th>Roman Numeral Stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 0</td>
<td>Carcinoma in situ</td>
</tr>
<tr>
<td>Stage I</td>
<td>Cancer localised to one area</td>
</tr>
<tr>
<td>Stage II</td>
<td>Advanced cancer in one local area</td>
</tr>
<tr>
<td>Stage III</td>
<td>Advanced cancer beginning to spread to the lymph nodes</td>
</tr>
<tr>
<td>Stage IV</td>
<td>Advanced metastatic cancer</td>
</tr>
</tbody>
</table>

1.4 Current Diagnosis
The time taken to diagnose cancer depends on several things. These include the type of cancer (47-49), the geographical location of the patient and also how advanced the cancer is. There are several common ways to diagnose cancer and the stage of the cancer. These include computerised tomography (CT) (50, 51), magnetic resonance imaging (MRI) (52, 53), positron emission tomography (PET) scans (54, 55) plus biopsies (56, 57). However, despite the range of existing ways to diagnose various types of cancer, hundreds of thousands of people still present at a late stage of cancer each year, with over 50% of all cancer-related deaths being caused by five types of cancer (Table 1.2).

Lung cancer is reported to be the most common cancer in the world, with over 1.6 million new cases being diagnosed every year (58). Lung cancer corresponds to 15% and 10% of all cancer diagnoses in North America and Western Europe respectively (2). It has been reported that the prognosis of lung cancer patients is exceptionally poor, with average 5 year survival rates of 7% for men and 9% for women in England and Wales (1). This is attributed to the late diagnosis of lung cancer (1). Current methods to diagnose lung cancer include bronchoscopy (59), CT scanning (60), PET scanning (61) and biopsy (62). However, bronchoscopies and biopsies are often painful processes (63, 64). In addition, the high cost of PET and CT scanners often limit their availability to patients, although this is improving (47, 65). Therefore, there is a clear need to develop methods to achieve early diagnosis and thereby optimise the treatment of disease. The problem of late diagnosis and thus patients responding poorly to cancer therapies is not localised to lung cancer alone and is common in other cancers (66).

1.5 Risk Factors Associated with Cancer
Most cancers are considered to be sporadic (i.e. not inherited). These are typically caused by cells becoming mutated which can lead to the development of cancer.
There are many risk factors associated with the development of cancer (67-75). These include a range of factors including age (76-78), body weight (73-75, 79, 80) and Additionally, other external factors such as smoking (including exposure to second-hand smoke) (73-75, 81) and sun exposure (82, 83) can also increase the likelihood of being diagnosed with cancer. As a result, much work has been done to increase the public’s knowledge and awareness of these links to cancer (84-91). A particular effort has been undertaken to educate the public about the link between smoking and exposure to the sun and cancer (86-88, 92, 93).

However, some cancers are due to genetic predisposition (94-97). These are caused by mutations carried by sperm and/or ova which are therefore passed from parents to children. Additionally, spontaneous mutations may occur due to the risk factors associated with the development of sporadic cancer. This in combination with inherited mutations increase the likelihood of the development of cancer. A typical hereditary cause of cancer is Li-Fraumeni syndrome which causes a mutation of the p53 tumour suppressor gene, which is involved in tumour suppression (98, 99).

1.6 Molecular Causes of Cancer
There have been many molecular links to cancer. In 2000, six “hallmarks of cancer” were described. This states that all cancers share six common features, allowing them to transition from normal cells to cancer cells, and allow these cancer cells to proliferate. These hallmarks are that cancer cells can stimulate their own growth without the need for growth signals, they are insensitive to growth-preventing signals from neighbouring cells, they are apoptosis resistant or have reduced levels of apoptosis when compared to normal cells, they can proliferate and reproduce indefinitely, they can stimulate and maintain angiogenesis and they can escape their site of origin and metastasise (100). In 2010 this was expanded to include four additional parameters – the use of abnormal metabolic pathways to generate energy, the evasion of the immune system, the presence of unstable DNA and local chronic inflammation at the site of cancer (101).

Oncogenes are genes which have become mutated. In tumours, these oncogenes are commonly overexpressed when compared to healthy cells. Oncogene activation has also been linked to apoptosis resistance (102). Due to this overexpression and the effects these mutations can have on tumour proliferation, oncogenes have been the target of gene therapy (103). Normal genes which become overexpressed or mutated are termed proto-oncogenes. The proteins these proto-oncogenes encode are typically linked to cell growth, cell differentiation and mitosis (104, 105). As a result of this, proteins encoded by oncogenes and
proto-oncogenes have been identified as drug targets. Examples of oncogenes include BCR/ABL and RAS. BCR/ABL is an oncogene generated by a Philadelphia translocation, causing the BCR gene from chromosome 22 to bind to the c-ABL gene from chromosome 9 (106). The activation of the BCR/ABL oncogene leads to uncontrolled cell proliferation in chronic myeloid leukaemia (107). BCR/ABL has been shown to be activated by ionising radiation (106).

Additionally, mutations of tumour suppressors such have been linked to cancer progression. Tumour suppressors are genes which protect a cell from one step on the path to cancer. Proteins encoded by tumour suppressors can inhibit cell division in mutated cells, repair damaged DNA and initiate apoptosis if damaged DNA cannot be repaired for any reason. In addition, proteins encoded by tumour suppressors can increase cell adhesion, thereby inhibiting metastasis (108). BCRA1 and 2 as well as p53 are examples of tumour suppressors. BRCA1 or 2 mutations account for between 5 and 10% of all breast cancers as well as being linked to ovarian cancer (109). BRCA1 and 2 are genes responsible for encoding proteins involved in the repair of damaged DNA (110). As well as links to breast and ovarian cancer, mutations in BCRA2 have been linked with Fanconi anaemia. Patients with Fanconi anaemia often go on to develop cancers such as acute myeloid leukaemia (AML) as well as cancers of the neck, skin or reproductive organs. BRCA1/2 mutations have been linked to exposure to UV light (111).

Resistance to apoptosis is caused by defects in the apoptosis pathway. As apoptosis is a cascade pathway (section 1.12.1), the disruption or inhibition of the proteases (caspases) involved in this process can lead to cells becoming apoptosis resistant. Apoptosis resistance has been linked to the dysregulation of the tumour suppressor protein, p53 as well as increases in the naturally occurring inhibitor of apoptosis proteins (IAPs) which irreversibly bind to caspases, stopping apoptosis from taking place. If a cell is mutated and apoptosis resistant, it can proliferate and cause cancer (112). The human papilloma virus (HPV) has been linked to apoptosis resistance (113).

A great deal of work has been done by many groups around the world to search for biological molecules to act as drug targets in different types of cancer (114-125). In addition, much work has been done to search for biomarkers to either diagnose cancer or monitor its progression and response to treatment (126-131). These biological targets include genes such as CYP17 (132), BRCA1 (133) and GSTM1 (134), amongst others (135-137). Additionally, proteins such as actinin-4 (138), phosphoinositide-3 kinases (139) and α-fetoprotein (140) have been put forward as potential drug targets or biomarkers. Amongst these potential
biological targets, many proteolytic enzymes have been suggested as molecules of interest (141-144).

1.7 Biomarkers
Biomarkers are molecules which are indicative of a response to a disease or treatment. These can be proteins (145), lipids (146), antibodies (147) or metabolites (148). Biomarkers can be used for both diagnostic and prognostic purposes (149). This allows the development of tests to diagnose a patient or predict their response to a treatment (150). Additionally, biomarkers have been used as targets for drugs.

Biomarkers are typically discovered by a variety of techniques, with several being used in combination to further validate the biomarker. These techniques include mass spectrometry (151, 152), northern or western blots (153) or microarrays (154). However, the most commonly used method in biomarker assay development is the Enzyme-linked immunosorbant assay (ELISA). This is an antibody-based assay developed to quantify a protein of interest. This is performed by immobilising a sample to a solid support and using an antibody specific to the antigen of interest to form a complex between the antibody and the antigen. At this point, if a detection antibody has not been already linked to the primary antibody, a secondary (detection) antibody with an affinity to the primary antibody can be added and the amount of detection antibody bound to the antigen-primary antibody complex can be quantified using spectrophotometry.

If a biomarker can be found providing clinicians with additional information about a disease or predict how a patient will respond to a treatment, then treatment can be tailored to the individual patient.

1.8 Personalised Medicine
Personalised medicine is the application of a treatment programme specifically tailored to the needs of the patient. This is currently performed by clinicians but the features examined by the physician are primarily based on the disease, rather than the patient and most treatments are made from empirical data and much more needs to be done in order to personalise treatments to individual patients (155). One exception to this in oncology is breast and ovarian cancer treatment. In this case, patients are screened for BRCA1/2 mutations and based on this genetic testing, the feasibility of treatment with PARP inhibitors is assessed (156). Members of the PARP family of proteins are involved in repairing damaged DNA. When BRCA1 or 2 are mutated, this can lead to mutations in these PARP proteins. This in turn can cause errors in the DNA repair process, which can cause breast and ovarian cancers.
Therefore, the effect of these mutated proteins can be reduced by inhibiting these PARP proteins. When these PARP proteins are inhibited, the cell cannot repair this damaged DNA and will undergo apoptosis (157). As a result, patients presenting with breast or ovarian cancer showing BRCA1/2 mutations are treated with PARP inhibitors. There is a clear need for personalised therapy to be applied to other diseases. For example, less than 1 in 3 children with acute lymphoblastic leukaemia (ALL) survive ten years if initial chemotherapy is unsuccessful (66). Therefore, if a link between biological differences relating to patients who respond well or poorly to therapy can be found, this link can be exploited in order to tailor treatment to the individual patient.

1.9 Enzymes

Enzymes are biological catalysts. These are fundamental in performing reactions essential to life at the speed required to maintain the organism’s survival and are present in every organism without exception. Enzymes have various roles ranging from regulating cell death (158), to assisting in the growth of new blood vessels (159), to synthesising DNA as is the case in DNA polymerase (160). Enzymes have a very specific amino acid sequence to give them the unique three-dimensional arrangement of their active site, where their substrate binds (Figure 1.1). The shape and properties of the amino acids at the active site are important to determine the substrate specificity of the enzyme.

![Figure 1.1: The induced fit model of an enzymic reaction. This shows that the shape of the active site does not remain rigid, as was originally proposed in the “lock and key model”. In reality, the active site shape changes slightly as the substrate binds and reverts back to its original shape as the products leave (161, 162).](image-url)
Some enzymes are found naturally in an inactive state. This allows the activity of the enzyme to be regulated by the addition of co-factors or post-translational modifications. If an enzyme is activated by the removal of a region it is called a proenzyme. Examples of this will be discussed in further detail in sections 1.10.1 and 1.12.1.

1.10 Proteases
Proteases are a specific type of enzyme which cleave or break down proteins into smaller peptides. The genes encoding these account for approximately 2% of most genomes (161). There are six protease families, based on which molecules are present at their active sites. These are serine, cysteine, threonine, metallo-, aspartic acid and glutamic acid proteases. Proteases have several important functions in the cell including breaking down proteins into peptides and eventually into their individual amino acids. Proteases are also integral to more complicated, highly ordered systems where they often regulate other proteases or signalling networks. As discussed before, enzymes are substrate specific, meaning different proteases will cleave at different positions on the protein. This will produce protein fragments with known termini. This is especially important in proteomics applications as this process regularly involves discovering the exact sequence of amino acids within a protein. However, in the case of proteases, their substrate profiles can overlap. This means that one protease can cleave multiple substrates and one protein can be broken down in the same place by more than one protease (161). Proteases can also act as “switches” in biological systems, meaning that a protease can cleave a pro-enzyme into its active form (162). An example of this is discussed in further detail in section 1.10.1. The levels of proteases within a cell or system must be kept within very strict limits in order to maintain homeostasis within the cell. Because proteases are themselves proteins, they can be subject to proteolysis. This helps remove proteases which are no longer needed and control their level within cells.
Monitoring protease activity rather than quantifying the amount of protease present has several advantages. Firstly, if a protease is mutated for any reason, this may affect its activity and as such its impact on the biological system(s) it is involved in. By measuring the activity of the protease, the effect of these mutations (if any) on the biological activity of the protease is known. Other groups have attempted to quantify protease activity by looking for a “fingerprint” of peptides produced by a target protease (175, 176). However, there are problems with this approach. If detecting a series of ten peptides, for example, gives a confident diagnosis of a disease, it may be unclear if the detection of six of these targets indicates the presence of this disease. Additionally, if additional weight is given to the presence or absence of some of these target peptides, this may lead to further confusion when diagnosing a patient with a disease. For example, if ten peptides give a confident diagnosis, with two of these peptides seen in 90% of cases and a patient produces seven of these diagnostic peptides, with only one of two of these commonly seen peptides, it may be difficult to confidently diagnose this patient. By determining protease activity using a single surrogate peptide and quantifying the amount of this cleaved peptide, its biomarker/prognostic indicator potential can be more easily assessed.
There are two possible ways to measure protease activity using a single peptide. One is to monitor how much of a particular peptide becomes lysed over a fixed time period. Alternatively, the time taken to break down a fixed amount of peptide could be monitored. However, when performing enzyme assays, the amount of substrate is usually several orders of magnitude higher than that of the target enzyme. This ensures that the enzyme is saturated with its target substrate, leading to a more accurate measure of total enzyme activity in the sample. Therefore, in this project, the amount of cleaved substrate produced by each target enzyme will be measured. This could potentially provide biomarker information by comparing these results from different patient populations (i.e. patients who respond well vs. patients who respond poorly to a treatment). By comparing the amount of peptide broken down over a fixed time period between these two sample groups, a protease’s potential as a prognostic indicator can be assessed.

In this instance, this will be done by adding a known amount of a synthetic peptide whose amino acid sequence is definitively known into a serum sample and monitoring its fragmentation over time. The synthetic peptide will be designed to be cleaved in a specific place producing two fragments of known amino acid structure and molecular weight. This enables these fragments to be monitored via mass spectrometry (MS) and quantitation of the fragments can be assessed. Once this has been assessed, the amount present can be compared to the amount that would be present if all the synthetic peptide had been cleaved and as such, protease activity can be assessed. A proposed model of the way the synthetic peptide and its fragments are isolated is shown in Figure 1.16. MS analysis is ideal for this assay due to its sensitivity, selectivity and flexibility (177, 178). This has been shown by many studies using MS to monitor enzyme activities (179-190).

Abnormal protease activity has been linked to several diseases including cancer, Alzheimer’s disease and arthritis (see Table 1.4 for references). As a result of this, fragments of proteins which have undergone proteolysis have been used as biomarkers (162). Table 1.4 shows a table linking various processes involving proteases to different diseases.

1.10.1 Coagulation
A well understood example of a disease resulting from protease disregulation is haemophilia. Haemophilia is a disease where a patient cannot form blood clots at the site of a wound efficiently. Clotting is the primary way to prevent blood loss and infection from bacteria entering a wound. The process of initiating and stopping coagulation is known as haemostasis. Haemostasis is regulated by three components. These are the vascular wall,
platelets and the coagulation cascade. Coagulation involves a protease cascade and results in the protease thrombin breaking down soluble fibrinogen into insoluble fibrin. Thrombin is naturally present in its inactive prothrombin form and only becomes active in the presence of calcium and thromboplastin. Once this takes place, thrombin can break down fibrinogen into fibrin and form a clot, as shown in Figure 1.2.

![Diagram of fibrin formation](image)

**Figure 1.2: The basic system of fibrin formation.** This shows that thrombin must be produced from prothrombin before it can cleave fibrinogen into fibrin and form a clot (taken from Riddel et. al., reference (170)).

Several coagulation factors have been identified and the coagulation cascade is now understood in great detail. The discovery of these factors allowed the elucidation of two different pathways of coagulation. These are termed the Intrinsic and Extrinsic Pathways, and are shown in Figure 1.3. The intrinsic pathway of coagulation begins when blood comes into contact with a negatively charged surface such as the membrane of an activated platelet. Platelets become activated after contact with collagen (168). Once this has happened, factor XII becomes activated XIIa. Factor XII gets anchored to the platelet with the aid of the molecule high molecular weight kininogen (HMWK). Once XIIa has been produced, this protease breaks down prekallikrein into kallikrein. Once kallikrein has been produced, more factor XIIa is produced via a positive feedback system. The process shown in Figure 1.4 then takes place, forming prothrombinase. Prothrombinase breaks down prothrombin into the protease thrombin. The extrinsic pathway is started by forming a complex between thromboplastin tissue factor (TF or factor III) on the surface of a cell and factor VII outside the cell system. When this complex is formed, factor VII becomes the activated factor VIIa and the cascade shown in Figure 1.3 takes place, forming prothrombinase.

After activated factor Xa has been produced from either pathway, the rest of the cascade takes place in an identical fashion. This part of the cascade is known as “the common pathway”. This is shown in Figure 1.3, culminating in cross linked fibrin (clot) formation. Once enough fibrin has been produced to form a clot, the coagulation cascade must be
terminated. This is done by the use of anticoagulants. An example of an anticoagulant is antithrombin, which is activated by binding to heparin-like molecules on endothelial cells. If the coagulation cascade is not functioning correctly and haemostasis is not maintained, several diseases may present themselves including, amongst others, haemophilias A and B (169, 170). Haemophilia is most commonly caused by the mutation of the X chromosome, leading to the deficiency of the clotting factor VIII (Haemophilia A). Therefore, management of haemophilia is currently treatment with recombinant clotting factors such as factors VII and VIII (191). In this and other instances, understanding protease action directly links to biomedical research.
Figure 1.3: The coagulation cascade. This shows the intrinsic, extrinsic and common pathways of coagulation, leading to cross linked fibrin production and clot formation (taken from Riddel et. al., reference (170)).
1.10.2 **WNT Signalling**

A proteolytically regulated pathway that has been shown to be involved in cancer is the WNT signalling pathway. WNT signalling is involved in controlling asymmetric cell division as well as the self-renewal of haematopoietic stem cells, T-cell activation and leukaemia development (192). There are three possible WNT signalling pathways. These are known as the canonical WNT pathway (also known as the WNT-β-catenin signalling pathway), the planar cell polarity pathway (PCP) and the WNT-Ca^{2+} pathway. The canonical WNT pathway is the most well known and characterised pathway. Unless a WNT ligand is bound to its (Frizzled) receptor and low density lipoprotein receptor-related proteins (LRP) 5 and/or 6 to activate the cascade shown in Figure 1.4(b), the protein β-catenin is broken down by the proteasome via a destruction complex. This complex consists of two proteins involved in tumour suppression known as adenomatous polyposis coli (APC) and axis inhibition (AXIN1) protein. Serine/threonine kinases, casein kinase 1 and glycogen synthase kinase 3β (GSK3β) are also involved in this pathway by phosphorylating β-catenin in four places. This provides receptors for β-transducin-repeat containing protein. Once this has happened, β-catenin gets ubiquitinated and broken down by the proteasome. If a WNT ligand is bound to its Frizzled receptor, the kinases in the β-catenin destruction complex are inactivated, so β-catenin does not get broken down and enters the nucleus of the cell. Anti-apoptotic genes in the nucleus have been identified as WNT targets (193).
Figure 1.4: The canonical WNT pathway. (a) in the absence of WNT signalling, and (b) in the presence of WNT signalling (taken from Staal et al., reference (192)). In the absence of WNT signalling, β-catenin becomes tetraphosphorylated and polyubiquinated. This provides modifications that lead it to be broken down by the proteasome. However, if a WNT ligand is bound to its Frizzled receptor, β-catenin does not become phosphorylated and can enter the nucleus. Here it can interact with genes in the nucleus and act as an oncogene. This can stimulate oncogenesis (174).
The PCP and WNT-Ca\textsuperscript{2+} pathways do not require β-catenin or LRP5/LRP6 expression. The PCP pathway is activated by the genes WNT5A and WNT11 through Frizzled receptors. This pathway affects the cytoskeleton of the cell and consequently its shape. The PCP pathway has been linked with inhibiting the canonical pathway by down regulating β-catenin levels as well as being involved with cell adhesion and migration (194, 195).

The WNT-Ca\textsuperscript{2+} pathway influences both the canonical and the PCP pathways. The pathways of PCP and WNT-Ca\textsuperscript{2+} are shown in Figure 1.5.
Figure 1.5: The (a) PCP and (b) WNT-Ca\(^{2+}\) pathways. In the PCP pathway, DVL and DAAM activate RHOA-ROCK pathway. This leads to reorganisation of the cytoskeleton. RAC1 is also activated by DVL and both RHOA and RAC1 activate the JNK stress response pathway. This pathway affects the cytoskeleton of the cell and consequently its shape. In the WNT-Ca\(^{2+}\) pathway, WNT-Frizzled complex activates PLC via a G protein, leading to InsP\(_3\) and DAG formation due to the cleavage of PtdInsP\(_2\). DAG then activates PKC. InsP\(_3\) binding to intracellular calcium store receptors elevates Ca\(^{2+}\) levels. This in turn activates PKC and phosphatase calcinerurin (Calcineurin). Calcineurin in turn activates nuclear factor of activated T-cells (NFAT) inside the nucleus. (taken from Staal et. al., reference (192)).
Overly active WNT signalling has been linked to several cancers including, amongst others, colon carcinoma, acute myeloid leukaemia, chronic myeloid leukaemia and acute lymphoblastic leukaemia (192). Several of the above examples have shown links to cancer amongst other diseases (Table 1.4). Therefore, proteases and their activities are potential prognostic or diagnostic markers for disease.

1.11 Enzyme Kinetics

As mentioned previously, enzymes are biological catalysts used to speed up reactions in organisms to rates essential for cellular survival. Much work has gone into understanding the kinetics of enzyme based catalysis. In 1913, a model of enzyme kinetics known as Michaelis-Menten kinetics was developed, building on work published by Henri et. al. in 1903 and is explained in this section.

Enzymes catalyse reactions by binding to a target molecule (a substrate), performing a chemical reaction on it, and releasing the products of this reaction whilst returning to its original (native) state. A scheme of this is shown in Equation 1.1.

\[
\begin{align*}
E + S & \rightleftharpoons k_1 ES \rightarrow k_2 E + P \\
\end{align*}
\]  
(Equation 1.1)

Where:

E = Enzyme

S = Substrate

ES = Substrate bound to Enzyme

P = Product

\(k_1\) = rate constant for the substrate binding to the enzyme

\(k_{-1}\) = rate constant for the substrate being released from the enzyme without undergoing a chemical reaction

\(k_2\) = rate constant for product formation

By assuming the rate of reaction \(k_{-1}\) to be negligible, the rate of product formation (V) can be described as a first order equation based on the concentration of ES and the value of the rate constant \(k_2\) (Equation 1.2).

\[
V = k_2[ES] 
\]  
(Equation 1.2)
However, as it is usually difficult to measure [ES] and subsequently to work out $k_2$, the total amount of enzyme (free and bound to substrate, Equation 1.3) as well as the amount of product can be measured.

$$[E]_t = [E] + [ES]$$  \hspace{1cm} (Equation 1.3)

Where:

$[E]_t$ = total enzyme
$[E]$ = free enzyme
$[ES]$ = enzyme bound to complex

Therefore, it is necessary to express the rate of reaction, $V$, in terms of substrate concentration and total enzyme concentration (Equation 1.4).

$$K_S = k_{-1} = \frac{[E][S]}{k_1 [ES]}$$  \hspace{1cm} (Equation 1.4)

Where $K_S$ is an equilibrium dissociation constant.

However, Equation 1.4 does not always hold true, as $E$, $S$ and $ES$ are not truly in equilibrium. This is because $ES$ is continually being removed to reform $E$ and create $P$. In 1925, Briggs and Haldane proposed a model avoiding the assumption that $ES$, $E$ and $S$ are in equilibrium (196). This model states that the more $ES$ present, the faster it will dissociate either into $E$ and $S$ or to $E$ and $P$. Therefore, when an enzyme catalysed reaction starts, $ES$ concentration will build up and quickly reach a steady state, where $[ES]$ is approximately constant until almost all $S$ has been reacted. This steady state accounts for almost all the reaction time, therefore reaction velocity can be calculated using these steady state conditions.

In this steady state, the rate of formation of $ES$ and the rate of its breakdown are equal (Equation 1.5).

$$k_1[E][S] = k_{-1}[ES] + k_2[ES]$$  \hspace{1cm} (Equation 1.5)

Where:

$k_1[E][S]$ = rate of $ES$ formation from enzyme and substrate
$k_{-1}[ES]$ = rate of $ES$ breakdown into enzyme and substrate
$k_2[ES]$ = rate of $ES$ breakdown into enzyme and product

This can be re-arranged to give Equation 1.6.

$$[ES] = \left(\frac{k_1}{k - 1 + k_2}\right) [E][S]$$  \hspace{1cm} (Equation 1.6)
The ratio of the rate constants in Equation 1.6 can be combined to give equation 1.7. In this equation, KM is defined as the ratio of the rate constants of ES degradation (either back to E + S or E + P) divided by the rate of ES generation.

\[ KM = \left( \frac{k_1 - 1 + k_2}{k_1} \right) \]  
(Equation 1.7)

By assuming KM is constant, as described in Equation 1.5, equation 1.6 can be re-written as Equation 1.8.

\[ K_M[ES] = [E][S] \]  
(Equation 1.8)

[ES] is now defined in terms of [E] and [S] – parameters that can be measured experimentally. However, as stated previously, [ES] is difficult to measure. Re-arrangement of Equation 1.3 shows that \([E] = [E]_0 - [ES]\). By incorporating this re-arranged form of Equation 1.3 into Equation 1.8, Equation 1.9 is produced.

\[ K_M[ES] = [E]_0[S] - [ES][S] \]  
(Equation 1.9)

Re-arrangement of this Equation 1.9 produces Equation 1.10.

\[ [ES](K_M + [S]) = [E]_0[S] = [E][S] \]  
(Equation 1.10)

This can be re-written in the form of Equation 1.11.

\[ [ES] = \frac{[E]_0[S]}{K_M + [S]} \]  
(Equation 1.11)

By relating Equation 1.11 back to Equation 1.2, Equation 1.12 is produced. This equation is known as the Michaelis-Menten equation.

\[ V = \frac{k_2[E]_0[S]}{K_M + [S]} \]  
(Equation 1.12)

Where KM is known as the Michaelis constant.

Because KM is a ratio of rate constants, it is unique to each specific reaction. Therefore, the reaction of a specific enzyme acting on a specific substrate has its own individual KM value. At high substrate concentrations (such as those to be used in this project), where \([S] \gg KM\), the reaction producing ES reaches its maximum velocity, Vmax. This is because all the enzyme active sites are filled by substrate and the reaction is said to be saturated. When
[S] >> KM, so KM + [S] ≈ [S], Vmax can be expressed by the equation shown in Equation 1.13.

\[ V_{\text{max}} = k_2[E]_t \]  

(Equation 1.13)

By expressing \( k_2[E]_t \) as Vmax, the Michaelis-Menten equation can be re-written (Equation 1.14).

\[ V = \frac{V_{\text{max}}[S]}{KM + [S]} \]  

(Equation 1.14)

By knowing the amount of substrate [S] put into samples in this project, and calculating the rate of reaction, V, by measuring [P] generated over time, and assuming \( V = V_{\text{max}} \), KM can be calculated. For each enzyme assay performed in this project.

Michaelis-Menten kinetics can also be used to describe multi-step enzyme reactions, similar to the scheme in Equation 1.15.

\[ E + S \xrightleftharpoons[k_1]{k_2} ES \xrightarrow{k_3} EP \rightarrow E + P \]  

(Equation 1.15)

In this multi-step equation, the term \( k_2 \) is defined by Equation 1.16.

\[ V = \frac{k_{\text{cat}}[E]_t[S]}{KM + [S]} \]  

(Equation 1.16)

Where \( k_{\text{cat}} \) encompasses all the reactions between ES and E+P. The value of \( k_{\text{cat}} \) encompasses the rate constants for the rate limiting steps in the reaction. In Equation 1.15, this is usually defined as \( k_2 \).

KM and \( k_{\text{cat}} \) are the two parameters which describe an enzyme’s activity. As described previously, KM is commonly associated with a specific enzyme’s affinity for a specific substrate. In an example where \( k_2 \ll k_1 \), then equation 1.7 can be defined as Equation 1.17.

\[ KM \equiv \frac{k - 1}{k_1} = K_s \]  

(Equation 1.17)

In this case, a large value for KM means \( k_1 \gg k_1 \). This means that the enzyme has a low affinity for the substrate in question, and binds to it very weakly. Therefore, under these conditions, KM is an inverse measure of substrate binding strength. However, a very large value of \( k_2 \) can lead to a large value of KM (Equation 1.7). Based on this, KM can be described as a measure of the instability of [ES], leading to the generation of E + S or E + P.
Unfortunately, this definition is not particularly useful, because a high KM can mean either P is formed quickly (high k2 value) or the ES complex dissociates back to E and S quickly (high k-1). The most useful way to view KM is as a plot of reaction rate versus substrate concentration (Figure 1.8). this shows KM is equal to the substrate concentration at which the reaction rate has reached half of its maximum. In other words, KM is the concentration of substrate required for effective catalysis to take place. Therefore, an enzyme reaction with a high KM requires a higher concentration to achieve a specific reaction rate than a reaction with a low KM, but the same kcat.

kcat is a direct measure of P generation under optimum conditions (i.e. all enzyme sites are saturated). \( \frac{1}{k_{\text{cat}}} \) can be defined as the time taken for one enzyme molecule to generate one product molecule, or the number of product molecules generated by one enzyme molecule per second.

In the case of low substrate concentrations, \([S] \ll KM\) and most of the enzyme is free (unbound to substrate). Therefore, \([E]_t \equiv [E]\). At this point, Equation 1.16 becomes Equation 1.18.

\[
V \equiv \frac{k_{\text{cat}}}{KM} [E][S] \quad \text{(Equation 1.18)}
\]

At this point, the ratio of kcat:KM acts as a second order rate constant for the reaction between E and S. This is important as it gives a direct measure of an enzyme’s specificity for a given substrate and its efficiency of processing it into P. This ratio shows allows the reaction rate to be calculated when the majority of enzyme sites are empty, and it allows a direct comparison of the efficiency of a given enzyme to bind and process a variety of different substrates. For example, if an enzyme, E, has a choice of two different substrates, A and B, present at an equimolar but dilute concentration, then Equation 1.19 holds true. In this case, the two substrates, A and B are competing against each other for enzyme sites and as such, their rates are dependent on each others’ concentrations.

\[
\frac{V_A}{V_B} = \frac{\left(\frac{k_{\text{cat}}}{KM}\right)[E][A]}{\left(\frac{k_{\text{cat}}}{KM}\right)[B][E][B]} = \frac{k_{\text{cat}}}{k_{\text{cat}}(KM)} A \quad \text{(Equation 1.19)}
\]

By knowing the amount of substrate \([S]\) put into samples in this project, and calculating the rate of reaction, V, by measuring \([P]\) generated over time, and assuming the rate of reaction is at its maximum \(V_{\text{Max}}\), the activity of the enzyme can be calculated for each enzyme assay.
performed in this project. By ensuring the amount of substrate in the assay is several times greater than the value required for rate of reaction to be 50% of $V_{\text{Max}}$ (the $K_M$ value), this ensures that $V_{\text{Max}}$ is reached. Because the enzymes to be tested in this project are physiologically at very low levels in peripheral blood, target substrates for these enzymes can be specifically designed to be cleaved by their target enzyme. This will allow [P] to be measured more easily, increasing the probability of differences in target enzyme activity between patient populations to be seen. The design of these target substrates will be discussed further in Chapter 3.

1.12 How do Proteases Enter the Bloodstream?
Proteases, have been documented as being detected in biological fluids such as plasma (197). Some of these proteases have shown biomarker potential, including cathepsin D (198), prostate-specific antigen (PSA) (199) and various kallikreins (141, 200, 201). Proteases can enter the bloodstream in several different ways. Firstly, cells can secrete these proteases in the form of exosomes (202), small vesicles (30 – 90 nm) secreted by cells which contain RNA and various proteins, including proteases (203). These are released from the cell when endosomes (vesicles secreted by the Golgi apparatus) bind with the plasma membrane (203). Exosomes have been proposed as a target for cancer vaccinations (204), as well as being used as diagnostic and prognostic indicators (205, 206). Proteases can also enter the bloodstream when a cell dies. Cell death typically takes place in one of three ways; either by apoptosis (type I cell death), autophagy (type II cell death) or necrosis (type III cell death) (207). These will be discussed in further detail below.
Figure 1.6: The differences between apoptosis and necrosis. During apoptosis, cells shrink and chromatin condenses. A process known as budding then occurs. During this, the plasma membrane is ruffled and bound apoptotic bodies containing organelles, nuclear fragments and cytosol are formed. These are then phagocytosed without triggering inflammatory processes. In contrast, during necrosis, the cell swells, becomes leaky and finally is disrupted. Its contents are then released into the surrounding tissue, typically leading to an inflammatory response. Modified from Van Cruchten et. al., reference (207)).

1.12.1 Apoptosis

Apoptosis, or type I cell death, is an active process of cell death. In apoptosis, a cell is broken down in an organised fashion before it is removed by phagocytosis. In apoptosis, the caspase cascade (TCC) is involved in killing a cell in an organised and ordered fashion. Caspases can be sorted into two classes, initiator and executioner caspases. Initiator caspases, such as caspases 2, 8, 9 and 10 become activated by binding to proteins such as cytochrome c (208). These activated caspases then activate executioner caspases, such as caspases 3, 6 and 7. These executioner proteases then digest a variety of intracellular proteins, allowing a cell to die (209).

When a cell undergoes apoptosis, its contents are broken down proteolytically and packaged into smaller bodies. This includes the degradation of chromatin, altering the plasma
membrane and compacting organelles within the cell (210). Caspases have also been linked to DNA topomerisomerase I fragmentation (211). These processes allow the contents of the cell to be removed by either macrophages or neighbouring cells, preventing an inflammatory response. It is believed that during apoptosis, rather than one caspase working on a substrate, a series of reactions activating several different caspases takes place. This process is commonly known as the TCC (210). In the TCC, a hierarchy of caspases is present, with one caspase, termed an “initiator” activates an “amplifier” caspase. This then activates “executioner” caspases, such as caspase-3 (Figure 1.7).
Figure 1.7: The caspase cascade in mammalian cells. In all of these cases, a biological stimulus activates an initiator caspase (e.g. caspase-12 or caspase-2). Once this caspase has been triggered, it in turn activates an amplifier caspase (e.g. caspase-9). Once these amplifier caspases have been stimulated, the next stage of the cascade takes place. This involves activating the executioner caspases caspase-3, caspase-6 and caspase-7. There are several pathways for this process to take place – the intrinsic and extrinsic pathways as well as the granzyme B (GranB) pathway. However, in all cases, the caspases 3, 6 and 7 become activated (taken from Chowdhury et. al., reference (158)).
If apoptosis is disregulated then several diseases can result. These include Alzheimer’s disease and various types of cancer (166, 167). Caspases typically exist in their inactive proenzyme form until they become activated. Studies on caspase-1 show the enzyme only becomes active upon the removal of an 11 kDa N-terminal peptide and a 2 kDa linker peptide. It has been suggested that this takes place via autoproteolysis (212). There are three main pathways by which apoptosis can be initiated. These are termed the intrinsic and extrinsic pathways and the granzyme-B pathway. The intrinsic pathway involves activating BH3 proteins which are signals to indicate cell stress or damage. Once these have been activated, this stimulates the activity of initiator caspases and starts the TCC. The extrinsic pathway involves the binding of death ligands from outside the cell such as tumour necrosis factor-α to receptors attached to the plasma membrane. This leads to the activation of initiator caspases which starts the caspase cascade. The granzyme-B path involves delivery of the protease from outside the cell, through specialised granules in the cell membrane and into the cell (213). Because caspases are linked to cell death, inefficient caspase activity could be a possible link to cancer. As apoptosis is believed to be a highly ordered process involving cellular content being packaged into apoptotic bodies which are subsequently engulfed by phagocytosis, cells undergoing death by apoptosis may not explain why proteases enter the bloodstream. However, cells which cannot undergo apoptosis have been linked to a variety of diseases, including cancer (214-218). This is because if a cell becomes mutated and is apoptosis resistant, this mutated cell will proliferate and not be removed, leading to the formation of a tumour (213, 219-221). Cells which cannot undergo apoptosis must find another way to die, such as necrosis (section 1.12.2). Based on the links between apoptosis disregulation and caspase-3 in apoptosis, caspase-3 activity in plasma will be investigated to assess its potential as a prognostic biomarker.

1.12.2 Necrosis
Necrosis, also known as type III cell death, was long thought of as being a totally uncontrolled process (222), however that has been debated in recent years (223). It has been shown that in some situations, ordered procedures such as the generation of reactive oxygen species and mitochondrial disruption take place in necrosis, indicating that necrosis is not always a totally disordered process (223). This ordered necrosis is known as necroptosis (224). Unlike type I or II cell death, only parts of necrotic cells are removed by phagocytosis
As the plasma membranes of cells become ruptured during necrosis, and not all cellular contents are removed by phagocytosis, this offers one possible route for proteins and proteases entering the bloodstream.

### 1.12.3 Secondary Necrosis

Aside from being removed by phagocytosis, apoptotic bodies can also undergo secondary necrosis. Secondary necrosis is when a cell has undergone apoptosis and its contents are enclosed in apoptotic bodies, these apoptotic bodies undergo necrosis, and their contents are spilled out into the surrounding environment (i.e. the bloodstream) (224).

The ways and means to measure changes in protease levels and this cell death require the most sensitive methods. Therefore MS approaches for such research have been developed.

### 1.12.4 Autophagy

Autophagy, or type II cell death, is a process of controlled cell death where cellular components are degraded by the cell’s own lysosomes and autophagosomes. These fuse together to form autolysosomes which can break down a cell’s components. Autophagy is typically stimulated by nutrient deprivation, but has also been linked to neurodegenerative diseases and cancer. Autophagic cell death is defined as a cell which undergoes programmed cell death without chromatin condensation, but with large autophagic vacuolarisation of the cytoplasm (226).

### 1.13 Mass Spectrometry

Fluorescence assays are commonly used to quantify protease activity (227-230). However, other methods using mass spectrometry (MS) based workflow allow for the monitoring of hundreds of peptides in one single analysis (231). This allows the potential to assay the activity of many proteases in a single multiplexed assay. This combined with the sensitivity of mass spectrometry means that MS is the ideal platform to monitor protease activities in this project.

Mass spectrometry (MS) involves the detection of charged ions which are separated out with the assistance of an electrostatic or magnetic field. In this procedure, a compound is vaporized, ionised, isolated and finally detected, as shown in Figure 1.8. This gives a mass:charge ratio ($m/z$) of the compound’s precursor ion. This technique can also cause fragmentation of the ion so individual pieces of its structure can be detected. From this, the structure of these fragments can be worked out based on their masses. As a result of the fragmentation, MS is known as a destructive technique. Mass spectrometry has been used in a
variety of applications, from simple compound analysis and purity checks to determining the amino acid sequence in proteins (231).

Figure 1.8 A schematic of a generic mass spectrometry instrument. This shows the sequence of generating ions from a sample and the ways this can take place in a proteomics study. This is followed by mass selection using various types of mass analysers, such as quadrupoles or ion traps. Ions are then detected and this data processed by the instrument’s software or a database. Based on the information received from the detector, further MS analysis can be performed on target peptides if necessary.
1.13.1 Electrospray Ionisation

Electrospray ionisation (ESI) was developed by Fenn et al. (232) at a similar time to Matrix Assisted Laser Desorption/Ionisation (MALDI). Collectively, these techniques are known as “soft” ionisation processes. This is because less fragmentation of the precursor takes place when compared to other “harder” ionization methods such as Electron Impact (EI) or Chemical Ionisation (CI) (233). ESI takes place by forcing a liquid through a very small capillary which has a high voltage between the tip of the capillary and an extraction plate. If necessary, an inert gas such as nitrogen can be used to desolvate these droplets. As the solvent evaporates and gets removed, the intact ions remaining in the droplet are forced together. These ions then Coulombically repel and break away from each other, forming smaller droplets still. This is shown in Figure 1.9. This process of ion repulsion is repeated until individual ions remain which are then sent into the mass spectrometer (MS) for analysis. This is a very reproducible and flexible method as this can be linked to any mass analyser. The main drawback of this technique is the production of multiply charged species, leading to more peaks being produced from each fragment. This, whilst often being advantageous, leads to a more complicated spectrum. The charge produced depends on several factors including molecule size and structure, the number of places where a charge can reside and the solvent used. However these charge numbers can be calculated and assigned to a particular fragment (234).

![Electrospray Capillary Tip](image)

**Figure 1.9: The process of electrospray ionisation.** The larger droplet breaks apart into smaller droplets when the size of the droplet reaches its Rayleigh limit. This is governed by the charge:surface area ratio. At this limit, the forces due to Coulombic repulsion exceed the surface tension of the liquid and the droplet breaks apart, forming smaller particles and finally individual charged entities. (Adapted from Kebarle et al., reference (235)).
1.13.2 Quadrupole Mass Analysers

Quadrupole devices were devised by Paul et. al. (236) and work using an arrangement of four cylindrical rods arranged as shown in Figure 1.10 below. Each opposing rod pair is connected electrically as shown. Each pair has an alternating current (AC) running across it 180° out of phase with the other set of rods. This creates a hyperbolic (quadrupolar) field. As well as this, both sets of rods can have a Direct Current (DC) flowing through them. To let every single ion through the quadrupole, this DC value can be set to zero. This is known as a radio frequency (r.f.) scan. As described above, the four sets of rods create a hyperbolic field. The distance between the edges of the opposite poles facing each other is 1/1.148 times the diameter of the rods themselves. The mid-point between these faces gives an ideal hyperbolic field.

Target ions are focused down the centre of the quadrupole. This is done by careful control of the DC & AC voltages applied to the quadrupole. All other (undesired) ions are unstable at these voltages. Due to their instability at these voltages, they hit the rods of the quadrupole and as such, do not reach the detector. It is the DC component which pushes the ions through the quadrupole and the r.f. components which provide a stable pathway for only a small selection of ions. All other ions have an unstable path through the quadrupoles and hit the sides and do not reach the detector. By adjusting the r.f. values for each set of quadrupoles, ions of different m/z can be filtered out. This is shown in Figure 1.10.

![Figure 1.10: Ion stability through a quadrupole mass analyser. Only three of the four rods have been shown for the purposes of clarity (taken from Glish et. al., reference (237)). The four rods create a quadrupolar field which ions travel down. This field makes the ions take a “corkscrew” like motion through the quadrupole. As a result of this, very few ions are stable through the quadrupole. This allows very good resolution, but limited sensitivity.](image-url)
Quadrupoles, and in particular triple quadrupole systems, form an essential component of selected reaction monitoring (SRM) (also known as multiple reaction monitoring, or MRM). This is explained in more detail in section 1.14.1. As a result of this, quadrupoles will be used in this project in a combined triple quadrupole-linear ion trap MS. This apparatus will be described in section 1.13.7.

1.13.2.1 Collisionally Induced Dissociation (CID)
CID is a fragmentation method typically used in triple quadrupole and quadrupole-time of flight mass spectrometers. This involves introducing a collision gas into a central quadrupole at such a pressure that ions entering the quadrupole undergo at least one collision with this gas. The gas used is typically an inert gas such as nitrogen, argon or helium. When ions collide with molecules of this inert gas, the kinetic energy in the gas is transferred to the ion. This kinetic energy then causes the ion to fragment. These ions can then be detected by quadrupoles or time of flight mass analysers. These fragments can then be used to sequence a peptide of interest (238, 239).

1.13.3 Ion Trap Mass Analysers
The ion trap devised by Paul et al. (240) consists of two curved (hyperbolic) electrodes, one at each end, and a ring-shaped electrode (usually called a ring electrode), as shown in Figure 1.11. The hyperbolic end electrodes usually have no DC voltage and an r.f. component of zero. The ring component, however, has an alternating r.f. component. This can be adjusted to filter out ions based on their m/z ratios after all ions have been “trapped”. The amplitude of the r.f. voltage is then increased in order to scan out these ions in order of increasing m/z values (241, 242).
Figure 1.11: A schematic of an ion trap. The ring electrode is circular and would be seen behind and in front of the ions being held in the trap, but this has been omitted for reasons of clarity (taken from Glish et. al., reference (237)). When ions enter the ion trap, they are held until the trap is full. Once the trap is full, ions are slowly scanned out towards the detector. This allows a high degree of sensitivity but limited resolution.

1.13.3.1 Resonance Excitation
As well as acting as a mass analyser, an ion trap can also perform multiple stages of mass spectrometry. This means the ion trap can perform fragmentation multiple times before scanning the product ions out of the trap. This can be performed by Collisionally Induced Dissociation (CID) (243), or resonance excitation (244). Fragmentation by resonance excitation takes place by altering the potential difference between the end caps of the ion trap very quickly. This causes the target ions to resonate and break apart. This process is known as “tickling” (244). The amplitude of this a.c. voltage to cause resonance excitation is low. If the amplitude of this voltage is too great, resonance ejection takes place, leading to ions leaving the ion trap.

1.13.4 High Performance Liquid Chromatography (HPLC)
Due to the nature of biological mixtures, they are often very complex. This makes it essential to separate components in a mixture so each component of interest can be analysed individually. This provides greater accuracy and simplicity in identifying and quantifying individual components in the biological mixture. This is particularly important when investigating protease activity as the role of proteases is to break down proteins in a sample. This means the sample will inevitably become more complicated as the time taken to perform
a protease activity experiment increases. This is the case until all possible areas of all protein substrates have been cleaved by enzymes. At this point the mixture is at its most complex so separating peptides becomes a necessity. Separation is performed by a variety of chromatographic techniques as described below. Chromatography relies on the principle that every compound has a unique affinity for individual solvents and surfaces. This was first performed using Thin Layer Chromatography (TLC) and then column chromatography. High Performance Liquid Chromatography (HPLC) was then developed to increase automation and obtain a high degree of separation on very low levels of analyte.

The way peptides are separated via LC methods is the peptides both stick to the coating of the column until a specific percentage of organic solvent is present in the mobile phase. Once this percentage has been reached, a peptide will have more affinity for the solvent than the column and will desorb from it. This level of organic solvent is known as the “hydrophobic foot”. The hydrophobic foot for a peptide depends on its structure. This allows peptides to desorb at different times and separate, as different peptides will have different organic levels at their hydrophobic foot. Figure 1.12 below describes the principle of HPLC.

**Figure 1.12:** A schematic showing how components in a mixture can be separated via LC methods. *First the mixture is injected into the system. It then flows through the tubing to the column. Here the components adsorb to the surface of the column until their hydrophobic foot is reached. Once this has been reached, the component desorbs and gets eluted from the column. Due to every component in a mixture having a different hydrophobic foot, they will desorb from the column at different times. This allows separation of components in a complex mixture.*

### 1.13.5 Liquid Chromatography-Mass Spectrometry (LC-MS)

Instrumentation advances now mean that once a component has been separated from a mixture, it can now be immediately analysed on MS platforms. Whilst it may be typical to have flow rates of 3 ml/min when performing HPLC analysis (245), LC-MS testing can often have flow rates of around 500 nl/min (246). The advantage of this is that as a nanospray of analyte is coming out of the LC system, it is easier to atomise the smaller droplets. Recent
advances involving novel drying techniques have allowed MALDI ionisation to take place. This allows molecules to be analysed (247, 248) via MALDI & ESI. This can also be performed under atmospheric pressure via Atmospheric Pressure MALDI (AP-MALDI) (248).

1.13.6 Ion terminology
As peptides and proteins are made up of chains of amino acids, invariably at one end of the molecule there will be a free amine group. This is known as the “N-terminus”. At the other there will be a free carboxylate group, known as the “C-terminus”. This is illustrated in Figure 1.13.

Figure 1.13: The terminology of different peptide fragments.
Figure 1.14: How peptide fragments can be used to sequence a peptide. In this spectrum, the y6 ion (m/z 699.41) has the amino acid sequence RSINLP and the y7 ion (m/z 812.50) has the sequence RSINLPL. This spectrum shows that by calculating the mass from m/z differences between neighbouring y ions, the amino acid next in the structure can be calculated. In this case, the difference in mass between y7 and y6 is 113.05, which corresponds to leucine acid. Not all a, b and y ions have been identified in this spectrum. The original amino acid sequence for this peptide is GVVDSEDLPNISR. This shows that by working out the mass differences between ion fragments, the complete amino acid sequence of a peptide can be obtained.
When an ion fragments, the positive charge it possesses will reside on one half of the molecule. If the positive charge is on the N-terminus fragment, the molecule is termed an “a_n” ion. If this fragmentation takes place one bond further in the molecule, it is called a “b_n” ion. If the fragmentation takes place one bond further in the molecule than the b_n ion, the fragment is termed a “c_n” fragment. The “n” here indicates the number of amino acids present in the fragment.

If the positive charge is on the C-terminal fragment, it is deemed either an “x_n”, a “y_n”, or a “z_n” fragment depending on where the fragmentation takes place. The “n” has the same meaning in these ions as it does in “a/b/c_n” ions (234). Figure 1.13 shows how these ions are named. An example of a mass spectrum showing these ions is shown in Figure 1.14, with each ion labelled in the method described above.

1.13.7 4000 Q-Trap MS Instrument

The 4000 Q-Trap mass spectrometer is a triple quadrupole MS device with the added feature of the third quadrupole being able to act either as a quadrupole or as a linear ion trap (249, 250). The purpose of the 4000 Q-Trap is to utilise the benefits of both the selectivity of triple quadrupole MS/MS and the sensitivity of an ion trap scan. When performing a standard triple quadrupole scan, very good quantitative analyses can take place when only a few target analytes are required. The drawbacks to performing scans using a triple quadrupole are that very few ions are stable throughout the entire path length at any one time. By using an ion trap MS, this problem is overcome when performing scans as ion traps offer a high degree of sensitivity. This sensitivity can be improved by holding ions in Q0 (Figure 1.15) until the ion trap has been emptied. This allows ion trap scans to take place over a longer period of time, increasing sensitivity. This process is known as “Q0 Trapping”. This cannot be performed on SRM scans however, as product ions are generated in Q2. If the intact peptide needs to be detected however, Q0 trapping can be employed, increasing sensitivity. This is useful when attempting to detect low abundance peptides. However, the drawbacks of using an ion trap are the low mass cut-off (25-30% of the m/z value of the precursor ion). Because of this, ion traps are often unable to ascertain a complete peptide sequence.
When performing a survey scan using the Q-Trap, ions flow all the way into quadrupole Q3 at a low energy (Figure 1.15). They are then collisionally “cooled” before being scanned out towards the detector. This type of scan is also known as an Enhanced MS (EMS) scan. If a specific molecule is to be analysed, instead of performing a survey scan, a product scan can be performed. In this instance, ions enter Q3 where they are isolated by removing all extraneous ions except the precursor ion which is then fragmented and cooled collisionally before leaving Q3 to be detected. The problem with performing a product scan is the size of Q3 means that there is a fixed capacity of the chamber. A consequence of this is that some ions are removed without being processed. Therefore some desirable ions may be unintentionally removed due to this limited capacity. Another main limitation of this type of scan is a mass cut-off of typically ~ 0.3 x m/z of the precursor ion (251).

By having the option of the final quadrupole in the Q-Trap acting as an ion trap or a quadrupole, these problems are overcome. This allows a different type of scan called an Enhanced Product Ion (EPI) scan. This combines the previous two types of scan to use the sensitivity of an ion trap and the resolution properties of a quadrupole simultaneously. Here, instead of ions being selected and fragmented at Q3, ion selection is performed at Q1 (Figure 1.15). They are then fragmented via acceleration from Q1 to Q2 which is pressurised with either nitrogen or a Noble gas. The ions collide with the gas molecules and break apart via CID. Because the fragmentation process is the same as the triple quadrupole MS method, a near identical fragmentation pattern is produced. This allows fragmentation pattern to be produced...
more quickly. As a result of the shortened time of selecting and fragmenting ions, higher duty cycles are now possible. EPI scanning is discussed in greater deal in section 2.1.5. The benefits of the sensitivity of an ion trap being employed mean that lower limits of detection are possible. The benefits of a triple quadrupole being employed means there is a lower mass cut-off. The result of these two benefits means that lower limits of identification can be reached. For proteomics studies, this approach combines the resolution and selectivity of a triple quadrupole system with the sensitivity of a linear ion trap (249). This allows SRM to take place as shown in section 1.14.1. Due to the advantages in both sensitivity and resolution, as well as the ability to perform SRMs, the 4000 Q-TRAP will be the MS used to perform relative quantitative assays in this project.
Sections 1.14.1 and 2.1.3 – 2.1.5 describe the various types of scanning that will be employed.

1.14 Quantitative Mass Spectrometry
In addition to providing qualitative data and identifying peptides of interest as well as their sites of modification (e.g. phosphorylation, ubiquitination), mass spectrometry can also be used to quantify peptides of interest, as described below.

1.14.1 Selected Reaction Monitoring
A Selected Reaction Monitoring (SRM) assay is a quantitative method typically performed on triple quadrupole mass spectrometers and involves an ionisation process followed by two separate stages of mass selection. The first event is the selection of the precursor ion. This is then passed onto the next stage where fragmentation takes place as described in an EPI scan above. The next stage of mass selection selects specific fragments of this precursor ion and analysing each one in turn, as shown in Figure 1.16. Both Q1 and Q3 are “static”. This means that they constantly select one ion of a given $m/z$, rather than scanning through all possible ions. This ensures the same precursor is selected each time and the product ion is the same each time for each SRM.
SRM assays can generate data leading to the absolute structural specificity of the precursor ion. SRMs can also calculate exact quantities of each precursor ion, provided their quantities are above the limits of detection, with the aid of an internal standard (252). Anderson & Hunter used 137 SRMs to monitor 53 high and medium abundance proteins in human plasma. SRMs have also been used by Kuhn et. al. to search for biomarkers in serum samples from patients with rheumatoid arthritis (231, 253, 254).

Quantitative information can be generated from SRM data by comparing either the intensity of the SRM response, or as is more frequently performed, the peak area from an LC-SRM analysis. Comparing SRM responses between samples allows relative quantitation to be performed. However, if an internal standard is used (typically of the same amino acid sequence but incorporating stable isotopes), the amount of a target peptide relative to this standard can be absolutely determined.

The advantages of SRM analysis are that it requires lower amounts of sample than other methods such as chemical labelling and it can be performed on a variety of sample types, unlike SILAC which can only be used in cell lines and small animals. The major drawback of SRM is its inability to analyse multiple samples in the same MS analysis.

As this assay is being developed on a triple quadrupole instrument, SRM with the use of an internal standard will be the primary method of quantifying protease activity in this project.
1.14.1.1 Internal Standards
The presence of an internal standard has long been used to increase the robustness of any biological assay. This is beneficial for several reasons. The first is this allows inter- and intra-lab variation to be corrected for. In addition, internal standards can be used to correct for matrix effects. Due to the inherent variation of plasma samples, the impact of these matrix effects cannot be predicted in any way. Therefore, additional steps to ensure the validity of the data produced in this assay must be taken in order to produce a valid result. The presence of an internal standard will allow this to take place.

Usually, a synthetic peptide containing “heavy” amino acids (i.e. amino acids containing higher amounts of $^{13}$C, $^{15}$N and/or $^{18}$O) of the same amino acid sequence as the target peptide would be used in an assay of this kind. These peptides are known as AQUA peptides and the approach was first developed by Gerber et. al. (255). Because the “heavy” peptide standard has the same chemical properties as the target peptide they will elute from the HPLC system at the same time, will ionise as efficiently but can be differentiated from the target peptide by a known increase in m/z due to the presence of these heavy stable isotopes of carbon, nitrogen and oxygen. In the absence of a “heavy” peptide, another peptide standard can be used but this may not necessarily elute at the same time as the target peptide and therefore may not necessarily correct for the matrix effects in samples.

1.14.2 Chemical Labelling
Chemical labelling can be used to perform relative and absolute quantitation between different samples on a multitude of different peptides. This involves digesting each sample individually, labelling each sample with a different chemical label, then pooling the samples together for MS analysis. The major advantages of chemical labelling is that allows the analysis of up to hundreds of peptides from up to eight different samples in one LC-MS analysis. A variety of chemical labelling techniques are available and all work in a similar fashion. These include iTRAQ, TMT and dimethyl labelling (256-258).

1.14.3 Stable Isotope Labelling with Amino Acids in Cell Culture (SILAC)
SILAC was developed by Ong et. al. and is typically used in cell line-based proteomics analyses (259). SILAC works by feeding different cell lines with growth
media either containing amino acids incorporating stable isotopes, or growth media with no stable isotope-encorporated amino acids. During protein synthesis, cells will use the amino acids in this media to synthesise new proteins. Therefore, by ensuring that in one of the cell lines, stable isotopes are encorporated, these “heavy” amino acids will be used in protein synthesis. This will lead to an increase in the mass of these proteins at specific amino acids, depending on what “heavy” amino acids have been used. These cell lines can undergo different treatments (e.g. treatment with a drug vs. treatment with a placebo) and the effect these treatments have on a cell line can be determined using mass spectrometry. The main advantage of this technique is that samples can be pooled before trypsin digestion. Therefore, the effect of the remainder of the workflow being performed on peptide losses will be the same on each cell line. This also allows hundreds of peptides from multiple samples to be relatively quantified in a single LC-MS analysis. The major drawback of SILAC is that is can only be used in cell line analysis and at present a maximum of three samples can be assayed in one analysis. Additionally, absolute quantitation cannot be performed using SILAC.

1.15 Project Outline
As discussed previously new biomarkers for cancer are required for earlier diagnosis and prognosis to increase survival rates. The many links between proteases and different types of cancer indicate that protease activity may be good biomarkers for different cancers. Therefore, the aim of this project is to develop a high throughput, low cost, sensitive assay method to monitor protease activity quantitatively in plasma. This assay could then be used to establish whether protease activity can be used as a potential prognostic indicator in order to predict whether a patient will respond to a particular form of cancer treatment. This assay involved adding a synthetic peptide of known sequence, designed to be cleaved by a target protease at a specific site to a plasma sample. This sample was then incubated at 37°C for a fixed period of time before being enriched for peptides. An internal standard was then added to this sample in a fixed amount. The amount of cleaved substrate produced by the protease of interest was then monitored by MS (Figure 1.17).
Figure 1.17: The sample preparation method developed in this project. This involved spiking in a non-endogenous (i.e. consisting of an amino acid sequence not found in humans) into a plasma sample and incubating it for a fixed amount of time at 37°C, allowing the target protease to act on this synthetic peptide substrate. Following this, the reaction was quenched before the sample is enriched for peptides. This enrichment process involves removing as much intact protein as possible, whilst leaving the smaller peptides in the sample. After the samples have been enriched, an internal standard was added to each sample in a fixed amount, allowing relative quantitation to take place, using a mass spectrometer as a detector. From this, patient samples can be assayed and compared to the results when assaying the same enzyme’s activity in healthy control samples. From this, the target protease’s biomarker potential can be assessed.

This report describes the various stages undertaken when developing this assay for monitoring enzyme activity in patient plasma samples.
Chapter 2: A Comparative Study between Selected Reaction Monitoring and a Novel Signal Boosting with Ion Trap Scanning (SBITS) Method to Detect Low Abundance Analytes on a 4000 Q-TRAP Mass Spectrometer

2.1 Introduction
This chapter describes the development and assessment of a novel mass spectrometric method designed for a 4000 Q-TRAP instrument, termed Signal Boosting and Ion Trap Scanning (SBITS). This method is compared against selected reaction monitoring (SRM), the current gold standard to quantitatively monitor target analytes on this mass spectrometer.

2.1.1 Signal Boosting and Ion Trap Scanning (SBITS)
A novel method has been developed in this chapter in order to provide an alternative way to monitor target peptide fragments. This has been termed Signal Boosting & Ion Trap Scanning (SBITS). If a target peptide is present in a sample at a very low level, the analyst may not know if any target analyte is present in the sample or if it is at a level too low to trigger MS/MS. This may lead to target peptides not being analysed by the mass spectrometer. By setting a threshold for MS/MS acquisition to a high intensity and adding in a synthetic peptide of identical structure at a high concentration, this threshold is guaranteed to be met whilst not acquiring data on other high abundance peptides which are not of interest. This is because the analyte present in the sample and the synthetic peptide will elute from the LC system at the exact same time, they will have identical m/z values and the synthetic peptide will exceed the threshold to acquire MS/MS. As a result, MS/MS will take place on both of these peptides (Figure 2.1).
Figure 2.1: A representation of SBITS. In this sample, the analyte is not present at a level great enough to trigger MS/MS and be detected. By spiking in a synthetic peptide of identical structure, this threshold is forced to be met and as such, MS/MS will take place on the analyte. MS/MS will also take place on the synthetic signal boosting peptide, but the benefits of iTRAQ labelling enables the two peptides to be differentiated. By keeping the amount of signal boosting synthetic peptide at a known, constant level, relative quantitation can also take place.

In order to differentiate between the synthetic peptide and the analyte present in the sample, the analyte and the synthetic peptide can be labelled with different iTRAQ labels and pooled together. The sample preparation is shown in Figure 2.2. This allows relative quantitation to take place between patient and control samples respectively.
Figure 2.2: The sample preparation of SBITS. This targeted approach to peptide monitoring allows MS/MS and quantitative data to be generated on levels of analyte normally too low to trigger MS/MS. In order to differentiate between the patient sample, the control sample and the synthetic sample, each is labelled with a different iTRAQ label. When MS/MS takes place on these samples, these diagnostic iTRAQ reporter ions will be released from the peptide and can be detected by the MS. This allows not only analyte origin identification, but it also enables relative quantitation between patient and control samples. By using other iTRAQ labels on the target synthetic peptide and spiking each of these in at different concentrations, a calibration curve can be generated. This allows absolute quantitative data to be generated at the same time.

In order to detect these ions, the mass spectrometer must be set up with scans to detect the precursor ion, obtain sequence information and detect iTRAQ reporter ions. This MS setup is shown in Figure 2.3.
Figure 2.3: The mass spectrometry workflow of SBITS. By performing a precursor ion scan, focusing on the signal boost iTRAQ reporter ion in Q3, the intact precursor ion of the analytes and the signal boosting peptide will be found. Due to the presence of the signal boosting peptide, the threshold to trigger MS/MS will be met. This can be deliberately set so high that the only component(s) in the sample to trigger MS/MS are the target peptide(s). Following this, an enhanced resolution scan will be performed to confirm the charge state of the peptide which has triggered MS/MS. This allows the analyst confirmation that a peptide of the correct m/z is triggering MS/MS and is not a singly charged contaminant (for example) present at such a high level, that it is triggering MS/MS. Following this, two Enhanced Product Ion (EPI) scans will take place, focusing on the most abundant precursor ion present in the system. The first will be performed very quickly and over a large mass range. This is to generate sequence data so the analyst can confirm the target found is indeed the peptide desired. The second scan will be performed more slowly over a shorter mass range in order to detect iTRAQ reporter ions and generate quantitative data using the sensitivity of the ion trap.
SRM has generated considerable interest in recent years. This is because of its ability to scrutinise individual peptide entities. As a result of this, it can be used in this project extensively. The fragment of peptide substrates will be known as will their molecular weights. This allows multiple SRMs to be set up to monitor all of these fragments in one experiment. This will provide fast, accurate data about protease activities in samples. This will allow the potential for protease activity to be assessed quickly and readily.

Due to the sensitivity and specificity of SRM analysis, SRMs will be set up to monitor the synthetic peptide fragments produced in this project and to obtain quantitative data on many fragments of interest very quickly.

2.1.2 MRM-Initiated Detection and Sequencing (MIDAS)

MIDAS analysis is a targeted peptide analysis method and involves using SRM to search for a peptide present in a sample such as a phosphopeptide (254). MIDAS works by monitoring specific SRM transitions. If the level of analyte exceeds a pre-given threshold, an EPI scan will trigger (260). This EPI scan involves fragmenting the precursor ion selected in Q1 via CID in Q2, and scanning out all fragments using Q3 as a linear ion trap. As SRMs involve one stage of mass selection, fragmentation, then an additional stage of mass selection, they do not have the desired full specificity as several ions present may provide these transitions (261, 262). However, the EPI scan in MIDAS analysis allows the analyst to sequence the fragment ions of the precursor ion, providing confirmation to the analyst that the SRM transition corresponds to a desired target ion (section 3.5). MIDAS has been shown previously to detect a target analyte at 10 fmol when in a background of a 6 protein mix at a level of 500 fmol (254). MIDAS was developed by Unwin et. al. in order to find sites of phosphorylation in peptides (254), but as quadrupoles have been proven to show a linear response to analyte concentration (263), this method can also be used to quantify the amount of analyte present in the sample.

2.1.3 Precursor Ion Scanning

Precursor ion scanning in the Q-TRAP involves fixing Q3 transmission on one m/z value and scanning Q1 across a given mass range. All ions stable through Q1 at any given time will enter the collision cell in Q2 where they are fragmented via CID. Following this, they enter Q3 where ions with only one particular m/z are selected (238). This is shown in Figure 1.16 (238).
2.1.4 Enhanced Resolution (ER) Scanning
ER scanning in the Q-TRAP uses Q3 as an ion trap. Q1 transmits all ions with a mass range of 6 – 8 amu around the target of interest. They pass through the collision cell in Q2 without undergoing fragmentation and enter Q3, which acts as an ion trap. Here, they are stored and scanned very slowly over a mass range of approximately 30 amu. This allows the charge state of the peptide to be determined. As some $^{13}$C will be naturally present in the peptide, a small amount of it will have a mass increase of 1 Dalton. If the peptide is doubly charged, the $m/z$ of this heavier isotope will increase by 0.5. If the peptide is triply charged, the $m/z$ of this isotope will increase by 0.3. The ER scans operates at a sufficiently high resolution to detect these mass differences. This allows the analyst to obtain the charge state of the analyte being observed in the mass spectrometer.

2.1.5 Enhanced Product Ion (EPI) Scanning
EPI scanning in the Q-TRAP uses Q3 as an ion trap as is likened to the inverse of a precursor ion scan. In an EPI scan, Q1 is static, only allowing ions of one given $m/z$ to pass through into the collision cell. Here, the selected precursor is fragmented via CID. Following this, all fragment (product) ions are detected over a mass range defined by the analyst, using Q3 as an ion trap to detect them (238). A schematic of this scan is shown in Figure 2.4.

![Figure 2.4](image-url)

**Figure 2.4:** Enhanced product ion scanning in a 4000 Q-TRAP. Here, Q1 is focussed on one $m/z$ and all others are excluded. All precursor ions with this $m/z$ enter the collision cell where they are fragmented. All product ions from this precursor ion are passed into Q3 where they are detected in the ion trap (238).
2.1.6 isobaric Tagging for Relative and Absolute Quantitation (iTRAQ)

iTRAQ is a technique for labelling peptides (264) or proteins (265) at their N-termini and the primary amines present on lysine side chains. iTRAQ consists of three main groups, as shown in Figure 2.5. The peptide reactive group forms an amide bond with free amines of peptides, allowing the iTRAQ reagent to bond to the peptide. The reporter group is released upon fragmentation to give diagnostic reporter ions. As iTRAQ is isobaric, the total molecular mass of each iTRAQ reagent must be the same. This is done via the use of stable, heavy isotopes being incorporated into iTRAQ to generate reporter ions of different m/z values. In order to ensure that each iTRAQ tag has the same total mass, a balance group is required in the molecule. This balance group uses stable isotopes to a different extent to the reporter group in order to ensure that the total mass of each iTRAQ label is the same. For example, the iTRAQ reporter tag of mass 121.1 amu contains a high number of heavy carbon, nitrogen and oxygen atoms. Therefore, in order to ensure the total mass is the same as other labels containing fewer stable isotopes (e.g. 113.1 amu), the balance group for the 121.1 tag will contain no stable isotopes, whereas the 113.1 balance group will contain many. This ensures MS/MS takes place on all iTRAQ labelled peptides with the same chemical structure at the same time and allows all diagnostic reporter ions to be generated at the same time (266), as their precursor masses are the same.
Figure 2.5: iTRAQ regent structure. This shows the placement of the peptide reactive group, the balance group, and the reporter group moieties. When the reagent attached to a peptide fragments, the bonds between the balance group and the reporter group and between the balance group and the peptide reactive group fragment. This ensures the peptide itself has the same mass after fragmentation, as all iTRAQ has been removed.

iTRAQ offers a maximum of eight channels to label peptides or proteins. These channels give diagnostic reporter ions of masses 113 – 119 and 121. Mass 120 has been deliberately omitted, as this is the mass of the immonium ion generated from phenylalanine fragmentation.

These reporter ions enable a calibration curve to be generated to obtain absolute quantitation of a target (267). This also allows relative quantitation of peptides from up to eight different samples pooled together. Pierce et. al. successfully used iTRAQ to compare the activities of six different tyrosine kinases associated with various leukaemias (256).

2.1.7 Project Description

The aim of this project is to monitor protease activity in biological samples using mass spectrometry. As these proteases are expected to be present in these samples at very low levels, a highly sensitive mass spectrometry method is required. Currently, the gold standard for targeted analysis using triple quadrupole mass spectrometers is
SRM. This chapter describes the development of a novel MS assay, SBITS, and a comparison of this novel MS method to SRM analysis in order to provide the most sensitive MS assay to detect the peptides used to assay protease activity in the remainder of this study.
2.2 Materials & Methods

2.2.1 Mass Spectrometry Equipment
All mass spectrometric (MS) analysis, unless stated otherwise, was performed on a 4000 Q-TRAP (Applied Biosystems, Foster City, CA) coupled to a liquid chromatography system consisting of a Famos autosampler, a Switchos trap system and an Ultimate Plus dual pump (LC Packings, Dionex, Sunnyvale, CA). The column used on the liquid chromatography system was an Acclaim pepmap 100, 15 cm, C18 column, pore size 100 Å. The software used for the 4000 Q-TRAP was Analyst 1.4.1 (Applied Biosystems). The software for the liquid chromatography system was Chromeleon 6.50 SP4 Build 1000.

To create the LC buffers A and B, Switchos buffer and needle wash solution, water was HPLC grade (Rathburn, Walkerburn, UK), acetonitrile was HPLC grade (Fisher Scientific UK Limited, Leicestershire, UK) and formic acid was for mass spectrometry grade (Fluka, Sigma-Aldrich, St. Louis, MO).

Buffer A & Switchos buffer consisted of 97.9% water, 2% acetonitrile and 0.1% formic acid.

Buffer B consisted of 79.9% acetonitrile, 20% water and 0.1% formic acid.

The needle wash solution was 100% water.

2.2.2 isobaric Tagging for Relative and Absolute Quantitation (iTRAQ) Labelling
First, the sample was dried using low pressure, elevated temperature centrifugation to remove all solvent. The sample was then dissolved in tetraethylammonium bicarbonate (1.0M, pH 8.5, Sigma, St. Louis, MO) (30 µl). Following this, the iTRAQ reagent (Applied Biosystems, Foster City, CA) was dissolved in propan-2-ol (HPLC grade, 99.9% purity, Sigma, St. Louis, MO) (35 µl) and this was added to the sample.

The iTRAQ container was rinsed with propan-2-ol (35 µl) and this was also added to the sample. This mixture was pulse spun and left at room temperature for a minimum of 2 hours. Following this, all solvent in the sample was removed via low pressure, elevated temperature centrifugation. The sample was dissolved in buffer A and the pH of the sample was tested. If necessary, the sample’s pH was adjusted until it reached 3 with 10% formic acid (mass spectrometry grade, Fluka, Sigma-Aldrich, St. Louis, MO) in water (HPLC grade, Rathburn, Walkerburn, UK). A typical iTRAQ: substrate ratio of 40:1 was used.
2.2.3 StageTip Sample Purification

In order to remove salts generated in the iTRAQ labelling process, samples can be purified via the StageTip purification process (268). To perform this purification, first a 1 ml volume of both the loading and elution buffers needed to be prepared. Once these had been made up, the StageTip (Proxeon, Odense C, Denmark) (200 µl tip containing a C_{18}) disk was conditioned by filling it with 40 µl of loading buffer, attaching a syringe to the tip, and using it to slowly push the buffer through the tip to waste. This was performed three times. Following this, the dry sample was dissolved in 120 µl of loading buffer. Forty microlitres of this was loaded into the StageTip and a syringe was used to slowly push the sample through the tip into a clean eppendorf marked “Flow Through”. The sample continued to be loaded and pushed through the tip into the “Flow Through” eppendorf until all the sample had passed through the tip. An additional 40 µl of loading buffer was added to the tip and passed through it into the “Flow Through” eppendorf. For high salt samples, this additional buffer washing was performed a minimum of three times. Following this, the StageTip was loaded with 10 µl of the elution buffer and a syringe was attached to the tip and used to push this into a separate clean eppendorf labelled “Eluted”. The eluted eppendorf had its solvent removed via low pressure, elevated temperature centrifugation. Once all solvent had been removed, the sample was re-dissolved in buffer A to a concentration predetermined by the analyst.

The loading buffer consisted of water (999 µl) (HPLC grade, Rathburn, Walkerburn, UK) and trifluoroacetic acid (1 µl) (99+ % purity, spectrophotometric grade, Sigma, St. Louis, MO)

The elution buffer consisted of water (499 µl) (HPLC grade Rathburn, Walkerburn, UK), acetonitrile (500 µl) (HPLC grade, Fisher Scientific UK Limited, Leicestershire, UK) and trifluoroacetic acid (1 µl) (99+% purity, spectrophotometric grade, Sigma, St. Louis, MO).

2.2.4 Signal Boosting & Ion Trap Scanning (SBITS) Sample Preparation for Method Development

First, an aliquot of GluFib (Sigma, St. Louis, MO) (1 μM) was iTRAQ labelled with reporter ion 114 in the method described in section 4.2. This was repeated on a separate aliquot of GluFib, labelling this aliquot with iTRAQ reporter ion 117. As these samples were to be sprayed from a static needle (Waters, Manchester, UK) in
some of this testing, all salts in the sample needed to be removed. This was performed by StageTip purification as described in section 4.3. Following this, all solvent was removed via low pressure, elevated temperature centrifugation. The sample was then re-suspended in buffer A to a concentration of 1 pmol/µl. iTRAQ labelled GluFib was used in all SBITS testing.

2.2.5 SBITS Mass Spectrometry Scan Settings
The first scan to be performed in SBITS was a precursor ion scan with Q3 focussed on the iTRAQ reporter ion of the trigger analyte. The scan range was 400 – 1200 amu. The collision energy of this scan was 28.8 eV. The Q1 resolution was Low and the Q3 resolution was Unit. The time taken for this scan was 3 seconds.

The second scan performed in SBITS was an enhanced resolution scan. One scan was summed in this scan with a scan rate of 250 amu/s. Dynamic fill time was used in this scan. Following this, an Information Dependent Acquisition (IDA) criterion was set, focussing on the most abundant peak over a mass range of 400 – 1200 amu with a charge state of 2+ or 3+ with an intensity of 5000 counts per second. Following this, the former target ion was excluded after 2 occurrences for 180 seconds. Once this IDA threshold had been met, an Enhanced Product Ion (EPI) was performed over a mass range of 100 – 1200 at a scan rate of 4000 amu/s, summing one scan. A dynamic fill time was used in this scan using a collision energy of 28.8 eV. Following this, a second EPI scan was performed on this peak over a mass range of 100 – 120 amu with a scan rate of 250 amu/s, summing 20 scans. Q0 trapping was incorporated in this second EPI scan. For this second EPI scan, the collision energy used was 60 eV. For all scans, the ion spray voltage was 2300.0 V and the interface heater temperature was 155.0°C. The total time this scan was performed for was 50 minutes. The total LC run time for this method was 70 minutes. This included a 20 minute window of washing the sample to remove any residual salt present in the sample before triggering the MS. The percentage of buffer B in the LC system began at 8% and peaked at 70 minutes at 100%. 60 µl of sample was injected onto the system. The Ultimate pump flow rate was 0.320 µl/min. The Switchos pump flow rate was 0.030 ml/min.

2.2.6 SBITS area of Fragmentation Investigation
The method “SBITS Mass Spectrometry Scan Settings” involves fragmentation at Q2. To investigate the effects of fragmentation in Q3, the following changes to the above
mentioned scan took place. In place of a second EPI scan, an MS2 Product Ion scan, causing fragmentation in the Q3 ion trap took place. This was performed over a mass range of 100 – 120 amu. The time taken for this scan was 5.20 seconds. The collision energy for this scan was 60 eV. The resolution for Q1 was Low and the resolution for Q3 was Unit. This was done to detect iTRAQ reporter ions and not to obtain analyte amino acid sequence. The EPI scan used to obtain the analyte amino acid sequence was the same as that described in section 4.5. A comparison between this method and the method described in section 4.5 was done.

2.2.7 Q3 Acting as a Quadrupole or an Ion Trap in SBITS
The method described in section 4.5 involves Q3 acting as an ion trap. To investigate Q3 acting as a quadrupole, the following changes to the method described in section 4.5 were made. In place of the second EPI scan, a Q3 scan was performed in its place. This scan was over the mass range of 110 – 120 amu. The time taken for this scan was 0.0810 seconds. Q3 resolution was Unit with a step size of 0.1 amu.

2.2.8 Localised Scanning at Q3 in SBITS
In an attempt to increase sensitivity for iTRAQ reporter ions whilst still using Q3 to perform MS/MS fragmentation, “localised” detection was performed and compared to section 4.6. The changes to the method in section 4.7 above are as follows. In place of a scan of 100 – 120 amu, the mass range in this scan was split in two. These ranges were focussed on the iTRAQ reporter ions used. These reporter ions were 115 and 117 amu. Therefore, the first scan was performed over the mass range of 114 – 116 amu, and the second scan was performed over the mass range of 116 – 118 amu. The time taken for to scan both these mass ranges was 2.5000 seconds each. Q1 resolution was Low and Q3 resolution was Unit. This was compared to the results generated in section 4.7.

2.2.9 Linearity Testing in SBITS
To perform linearity testing, the method for the MS used is a single EPI scan looking for iTRAQ™ reporter ions. The scan time on the MS was 20.009 mins, the total number of cycles was 333, and the precursor ion scan was fixed on 626.000 amu. The LC method was adjusted, keeping the percentage of buffer B at 40% at all times. The total run time for the LC method was 24 minutes. This included a 4 minute wash time. The trap wash method was also adjusted so the percentage of buffer B was kept at
40% at all times. The total time for this trap wash was 10 minutes. Samples consisting of GluFib labelled with 114 and 117 iTRAQ reporter. Section 4.2 describes the method for iTRAQ labelling. These samples were then purified via StageTip purification as described in section 4.3 before being pooled as described below. Samples of varying concentrations of 114 labelled GluFib were pooled with samples of 117 labelled GluFib. The 117 labelled GluFib was at 1 pmol/µl in every case. The 114 labelled GluFib was added to the individual samples as described in Table 2.1.

Table 2.1: The amount of 114 and 117 iTRAQ labelled GluFib target peptide present in each sample in SRM and SBITS linearity testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of 117 labelled GluFib in sample (pmol)</th>
<th>Amount of 114 labelled GluFib in sample (fmol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>500</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1000</td>
</tr>
</tbody>
</table>

Once sample 5 had been prepared, a serial dilution series originating from this sample 5 was prepared by pipetting 20.0 µl from the previous sample into the next eppendorf. These samples were then diluted using buffer A to reach the levels described in the table above. These samples were prepared three times and each sample injected. The area under the peak of the 114 amu peak was recorded and a plot of peak area vs. amount of 114 labelled GluFib added to the static needle was created. The linearity of this plot was examined by assessing the equation and R² value of the line of best fit. The area under this peak was calculated by first obtaining an extracted ion chromatogram of all ions with m/z over the range of 113.5 and 114.5. Following this, a background subtraction was performed followed by two Gaussian smoothes of the peak.

2.2.10 Multiple Reaction Monitoring-Initiated Detection And Sequencing (MIDAS) Limit of Detection
To perform MIDAS, selected reaction monitoring (MRM/SRM) transitions for iTRAQ labelled GluFib were set up. The peptide was iTRAQ labelled as described in section 2.2.2. Following this, a dilution series of this peptide was created as described in table 2.2. These samples were run via LC-MS. The total LC run time for this method was 70 minutes. The MS method for this involved setting up an SRM
transition of 626.0 \(\rightarrow\) 114.0 with a dwell time of 200 ms. Once a threshold of 100 counts per second was reached, an EPI scan over a range of 140 – 1400 amu was performed with a scan rate of 4000 amu/s and a dynamic fill time. This was to obtain amino acid sequence data, allowing confirmation of the target analyte. This included a 20 minute window of washing the sample to remove any residual salt present in the sample before triggering the MS. The percentage of buffer B in the LC system began at 8% and peaked at 70 minutes at 100%. Sixty microlitres of sample was injected onto the system. The Ultimate pump flow rate was 0.320 µl/min. The Switchos pump flow rate was 0.030 ml/min.

Table 2.2: The amount of iTRAQ 114 labelled GluFib in MIDAS testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of iTRAQ Labelled GluFib in Sample (fmol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
</tr>
</tbody>
</table>

Before these samples were run, two blank samples consisting of 100% buffer A were injected. An identical blank was injected following the injection of all 5 samples shown in the table above. From these blanks, the standard deviation of the noise (i.e. standard deviation of the ion count when no analyte is present) was calculated. The limit of detection was defined as any signal whose intensity is greater than three times the standard deviation of the noise.

### 2.2.11 SBITS and SRM Selectivity Testing

iTRAQ labelled GluFib (10 pmol) was added to an *E. Coli* whole cell lysate tryptic digest and analysed by both SBITS and SRM. From this, the contamination from other peptides present in the sample leading to either an SBITS or SRM response could be assessed.
2.3 Results
In order to develop the SBITS method and assess whether SBITS or SRM is the most efficient at detecting specifically targeted analytes, GluFibrinopeptide (GluFib) was used as a peptide standard for targeted assays. The method which detects GluFib with the least variance, greatest dynamic range, and lowest limit of quantitation will be applied to the remainder of this project in order to detect the cleaved form of bespoke protease targets to assay the activities of specific proteases in plasma.

2.3.1 SBITS Development and Optimisation
In order to develop and optimise this SBITS method several mass spectrometry parameters needed to be optimised. These included the fragmentation method (CID vs. resonance excitation) and determining whether Q3 of the 4000 Q-TRAP should be used as either a quadrupole or a linear ion trap in order to detect iTRAQ reporter ions. In this development work, iTRAQ labelled GluFib was used as a target analyte. This was labelled with two different iTRAQ tags and pooled in an equimolar amount.

Initial SBITS testing showed that several important criteria are provided in order to identify a target of interest. These include monitoring the iTRAQ reporter ions in order to differentiate between the response provided by the signal boosting peptide and the endogenous peptide in the sample (Figure 2.6), a full range MS/MS scan, allowing the peptide to be sequenced (Figure 2.7) and identifying the charge state of the peptide (Figure 2.8). An attempt to improve the resolution in the iTRAQ reporter region, using Q3 as a quadrupole rather than an ion trap proved unsuccessful. When using Q3 as a quadrupole for the endogenous iTRAQ signal (114 tag), the peak width was 0.725 Da at the peak’s full width at half maximum (FWHM) (Figure 2.9) as opposed to a resolution of 0.28 FWHM when Q3 is used as a linear ion trap (Figure 2.6). In addition, using Q3 as a linear ion trap proved to be more sensitive than using it as a quadrupole, with a signal:noise ratio (S:N) of $2.5 \times 10^7$ when using it as an ion trap, compared to 58.63 when using Q3 as a quadrupole. This was calculated by comparing the signal of the peak to the noise generated 30 seconds before and after the peak (Agilent Technical note, www.chem.agilent.com/Library/technicaloverviews/Public/5990-7651EN.pdf).

A further attempt to improve resolution using Q3 as a quadrupole was unsuccessful when the mass range was isolated to a small window (1 amu either side of the iTRAQ
reporter ions of interest). This led to a FWHM peak width of 0.75 and a S:N of 88.33 (114 Da). These values do not outperform using Q3 as an ion trap. Attempts to improve iTRAQ reporter ion sensitivity by performing peptide fragmentation in the linear ion trap by resonance excitation proved unsuccessful, as no iTRAQ reporter ions were detected. This correlates with ion trap fragmentation results published previously (269, 270). Therefore, the optimum conditions for the SBITS method are to perform fragmentation in Q2 and use Q3 as a linear ion trap (Figure 2.1).

![Figure 2.6: The EPI scan detecting iTRAQ reporter ions in SBITS. This scan was employed to obtain reporter ion sensitivity.](image)

The EPI scan in Figure 2.6 is focussed on detecting iTRAQ reporter ions only. This potentially allows relative quantitation between the analyte (115 peak) and the trigger (117 peak) to take place. Alternatively, if the 117 peak is solely used to force MS/MS to take place and two samples are present in the sample, relative quantitation can potentially be assessed between the two samples. As this scan is used for obtaining quantitative data, this scan is deliberately performed over a short mass range with a very slow scan rate. This scan is performed 20 times and each scan summed in order to improve the signal:noise (S/N) ratio and gain as much reporter ion sensitivity as possible.
The EPI spectrum in Figure 2.7 shows sequence data for this peptide. This provides confirmation that the analyte is the desired one. As SBITS has the potential to multiplex (i.e. analyse more than one peptide in the same analysis), it is important that definitive structural confirmation of the analyte is obtained. This EPI scan enables this data to be obtained. Before performing this scan, the fragmentation conditions must be optimised in a separate experiment to ensure optimal fragmentation and sequence coverage.
Figure 2.8: The ER scan used to determine charge state in SBITS. This shows an ER scan of GluFib performed in SBITS testing. This not only shows the m/z of the precursor ion, but it also allows the charge state of the molecule to be obtained. Due to isotopic effects in carbon, some $^{13}C$ will be present in the molecule. Therefore, some of this peptide will have a mass gain of 1 amu. If the target peptide has a charge state of 2+, this heavier natural isotope will have a mass increase of 0.5 m/z. If the target peptide has a charge state of 3+ (as seen here), the heavier natural isotope will have a mass increase of 0.3 m/z.

The scan shown in the spectrum in Figure 2.8 is an enhanced resolution scan. This not only shows the m/z of the target analyte, but also allows the charge state of the peptide in question to be ascertained. By knowing this information, the mass of the intact precursor ion can be obtained, providing additional information confirming that the peptide being analysed is of interest.

Following these initial results, the method of fragmenting peptides was investigated. The ability of the 4000 Q-TRAP to act as both a triple quadrupole and a linear ion trap allows analytes to be fragmented in two ways. These are either fragmentation in Q2 using CID, or fragmentation in the linear ion trap using resonance excitation. Using the SBITS initial results (Figures 2.6 and 2.7) where fragmentation was
performed using CID in Q2 as a benchmark, a comparison was made into peak width, peak intensity, and sensitivity when performing fragmentation in the collision cell in Q2 and in the linear ion trap in Q3. The resulting iTRAQ reporter ion scan spectrum was compared to that of the results shown in Figure 2.6.

By performing fragmentation in Q3, all ions with an $m/z$ of approximately 1/3 or less of the precursor ion’s $m/z$ (i.e. all ions with $m/z < 208.6$ amu) are excluded. As the precursor ion $m/z$ of this analyte is 626.0, no iTRAQ reporter ions were generated in the spectrum. This corresponds with results reported previously (251). Therefore, fragmentation should be performed using the collision cell in quadrupole Q2.
When comparing these data to those generated in Figure 2.6, whilst the noise appears lower in these results, the signal also appears lower. Therefore, the S/N is greater when using an ion trap as opposed to a quadrupole. Based on this, it would appear more beneficial to use Q3 as an ion trap rather than a quadrupole, as this offers more sensitivity.

In an attempt to improve sensitivity, localised Q3 scanning in the SBITS method was performed. The resulting spectrum for iTRAQ reporter ions shown in Figure 2.10 was compared to that of the scan performed in Figure 2.6 in terms of signal intensity, peak width and sensitivity.
Figure 2.10: (a) Using Q3 as a linear ion trap or (b) performing localised quadrupole scanning to detect iTRAQ reporter ions in SBITS.

The rationale behind this is that performing this scan using a quadrupole 1 amu either side of the target reporter ions, this scan would be more sensitive. However, this spectrum shows very wide peaks with a signal: noise ratio intensity lower than when using Q3 as a linear ion trap. Wider peaks are potentially problematic because it may be necessary to obtain absolute quantitation for the proteolytic peptide fragment generated by proteases. This would involve using all channels available in iTRAQ to generate a calibration curve. By having wide peaks as shown in (b), it would be difficult to see where one peak ended and the next began. This is a potential source of error when assessing the peak areas of target analytes. Based on this and increased sensitivity offered by an ion trap, this type of scan will not be used in SBITS analysis. Therefore, the optimal way to perform SBITS analysis on target peptides is to perform fragmentation by CID in Q2, and to use Q3 as a linear ion trap (Figure 2.11).
2.3.2 SBITS Linearity Testing and Comparison to SRM

Linearity testing was performed on iTRAQ labelled GluFib with the signal boosting peptide (119 tag) kept at a high amount (10 pmol) and the amount of simulated endogenous peptide (114) in the sample was varied. This peptide was spiked into samples at 1, 5, 10, 50, 100, 500, 1,000, 5,000 or 10,000 fmol. This created a concentration range consisting of nine points spanning four orders of magnitude. In addition, the variance at each data point was examined. These samples were also analysed using an SRM workflow (Figure 2.12). This shows both methods provide a linear response over a wide concentration range.

SBITS linearity testing showed that the GluFib cannot be detected below 10 fmol. Therefore, the limit of detection using SBITS was at least 10 fmol. The limit of
quantitation appeared to be 50 fmol. SRM variance exceeded 50% at a level of 1 fmol (73.5%). Therefore, the limit of detection for SRM analysis was at least 1 fmol, but the limit of quantitation was between 1 and 5 fmol.

When performing limit of detection/quantitation testing in a 6 protein mix background, the average variance for the SBITS method was 25.0%, with values ranging from 11.4% at 10 fmol endogenous GluFib to 17.4% at 10 pmol endogenous GluFib (Figure 2.26). Therefore, the LOD for the SBITS method in a 6 protein mix background was at least 1 fmol, whilst the LOQ lies between 10 and 50 fmol. In the 6 protein background, the SRM method outperforms the SBITS method in terms of sensitivity and variance, with an LOD of at least 1 fmol, an LOQ of 5 fmol, and an average variance of 28.2%, ranging from 28.7% at 5 fmol to 27.5% at 10 pmol (Figure 2.12).

![Graph showing SBITS and SRM linearity testing](image)

**Figure 2.12: SBITS and SRM linearity testing.** *Error bars represent the standard deviation of each sample population (n = 3).*

Both methods show a linear response over a concentration range of three orders of magnitude with similar variances ($r^2$ values of 0.9860 and 0.9818 for SBITS and SRM respectively). However, SRM offers lower limits of detection and quantitation than SBITS.

Due to the lower limit of detection of the SRM method, if a small number of analytes are to be analysed in an experiment, SRM is better suited to detect these low level targets.
When performing MIDAS analysis monitoring for one transition (626.0 → 114.0), 1 fmol was easily detected. This LOD testing was not performed at levels of analyte below one femtomole. This testing proved that MIDAS™ is a preferential way to detect low abundance analytes. It has previously been proven that using quadrupoles to detect a molecular species provides a linear response when comparing peak intensity against analyte concentration (263). A comparison of the linearity of SBITS and MIDAS also confirms this, as well as showing a lower limit of detection than SBITS. As a result of this testing, MIDAS has been proven to not only be able to provide low levels of detection, but also a linear response, allowing quantitative data to be obtained. Figure 2.13 shows a comparison between (a) SBITS at 20 fmol, and (b) MIDAS at 1 fmol.

![Figure 2.13: (a) SBITS and (b) SRM limit of detection testing using 20 fmol and 1 fmol of target analyte respectively. As this shows, the analyte peak (114 amu) cannot be detected above background noise when this scan is performed. (b) The SRM transition for MIDAS at 1 fmol.](image)

The data in Figure 2.13 show that the SRM transition for the analyte has exceeded the threshold for MS/MS and as such can offer the analyst confirmation of the target analyte. This shows that MIDAS has a lower limit of detection that SBITS. Based on these observations, using SBITS to detect low abundance species is not a viable method when compared to methods such as MRM-Initiated Detection and
Sequencing (MIDAS). This is because MIDAS can detect analytes in the same 6 protein mix background of 10 fmol (254).

### 2.3.3 SBITS Selectivity Testing

When performing selectivity testing, iTRAQ labelled GluFib (10 pmol) was screened for in the background of an *E. Coli* tryptic digest of 500 ng. A response was only seen at the elution time of GluFib, with no other response (Figure 2.14(a)). Further examination showed that the peptide detected was GluFib, and not a different peptide (Figure 2.14(b) and (c)). This demonstrated the selectivity of the SBITS method in a complex background. SRM was able to detect GluFib at its retention time with a minimal amount of contamination around the retention time of GluFib (Figure 2.14(d)). This shows that the SBITS method offered comparable selectivity to the SRM method.
(b) Monoisotopic iTRAQ labelled GluFib

(c)
Figure 2.14: The selectivity of SBITS and SRM. (a) Total ion chromatogram detecting iTRAQ labelled GluFib, as well as (b) Enhanced resolution scan to determine the charge state of the peptide and (c) a product ion scan to sequence the peptide in an SBITS workflow, confirming it is GluFib. This shows that SBITS has a similar selectivity to (d) an SRM workflow when detecting GluFib.
2.4 Discussion

SBITS was developed to perform a signal boosting workflow on low-level peptides using iTRAQ labelling and synthetic analogues of peptides of interest at high concentrations. Setting a threshold for MS/MS ensured MS/MS only takes place on peptides of interest. However, at this stage of the project, the amount of proteolytic fragment which will be generated in both patient and control samples is unknown. Therefore, the amount of proteolytic fragment in the sample may be insufficient to trigger MS/MS. By spiking in a high amount of a synthetic peptide of identical structure, it is guaranteed that this threshold will be met and MS/MS will take place. This is because the fragment in the sample and the synthetic peptide spiked into the sample as a signal boost will have the same retention time as each other in any LC method. Therefore, these peptides will co-elute and the threshold for MS/MS will be met. MS/MS will then take place on both the sample fragment and the synthetic signal boost. In order to differentiate between the sample and the signal boost, the two fragments would be iTRAQ labelled with different reporter ions (Figure 2.15). To detect the limit of analyte present, the iTRAQ reporter ion for the analyte would be monitored in the EPI scan designed to monitor the presence of these ions. This allows the sample fragment to be detected at lower levels (Figure 2.6). By comparing the amount of analyte present against a standard or another sample (e.g. comparing patient and control samples), the amount of the analyte peptide present in the sample can, in theory, be calculated. Figure 2.16 shows how a comparison between patient and control samples could take place.

MIDAS is performed by using SRMs to monitor peptide fragments, using very small mass transmission windows around Q1 and Q3 (typically < 0.5 Da). This removes a great deal of background noise, as most of it is not transmitted through the quadrupoles. Once a target peptide is identified by SRMs and a threshold transmission rate is met, an EPI scan will take place allowing peptide sequence information to be obtained. This will provide the analyst with confirmation that the peptide being monitored is a peptide of interest.
Figure 2.15: A theoretical comparison between different sample populations using SBITS. In this case, the patient’s protease in question has a higher activity than the control sample. Therefore, more proteolytic fragment is generated in the patient sample. This MS/MS scan allows a comparison between the samples and relative quantitative data to be obtained.

Before performing linearity testing for SBITS, the most concentrated sample was injected using a dynamic fill time. When using a dynamic fill time, the mass spectrometer allows the ion trap to fill until a certain level of ions are present. This provides optimum sensitivity whilst preventing the trap from being over-filled, which would lead to space charging effects. These space charging effects include reduced resolution, mass accuracy, sensitivity and dynamic range (250). Analysing the sample containing iTRAQ labelled GluFib at the highest concentration under these conditions
and obtaining the fill time used in this experiment, this allowed the fill time to be fixed at this level when performing this testing. This kept all conditions in the mass spectrometer consistent whilst avoiding space-charging effects from the most concentrated samples (271). By keeping this and all other parameters constant, the only variable in this experiment was the concentration of the analyte. This allowed a valid comparison between analyte concentration and target reporter ion peak area to take place.

When performing a comparison between SBITS and MIDAS, the total cycle time was different. This is because the SBITS method is a more intricate method, involving many more types of scan. However, fill time for the SBITS EPI scan detecting iTRAQ reporter ions was as great as possible whilst keeping the cycle time below 10 seconds. It was noticed in preliminary work that the peak width of the target peptide was 30 seconds. By keeping the cycle time below 10 seconds, this ensured that MS/MS is triggered at least once at a point near the apex of the peak, allowing as much sensitivity as possible. The dwell time for the SRM transition in MIDAS was 200 ms. This allows a long dwell time to look for a target peptide. This increases the sensitivity of the MIDAS scan.

The theoretical benefit of SBITS is that as SBITS is a targeted approach, fragmentation of the target peptide in both EPI scans can be optimised beforehand. This allows the analyst to optimise fragmentation for any peptide in order to generate good sequence coverage in the first EPI scan and good iTRAQ reporter ion generation in the second. This also allows the analyst to know beforehand (i.e. whilst optimising the fragmentation conditions) the retention time of the peptide, so he has a very small time window in the run in which he needs to process the data. The potential to multiplex this method does exist by having separate periods in the method, each period tailored to the elution of one specific target. Once this has passed, a separate period can begin for the next target analyte. This was attempted at the outset of this project, however software limitations prevented this. By working with the manufacturer, this problem can very easily be overcome. This has been proven by the manufacturer releasing a software package allowing SRMs to be scheduled over specific times, allowing this process to take place in MIDAS. SBITS provides a detection method allowing qualitative detection of any target peptide providing it ionises and is detectable in a mass spectrometer. This limiting factor is not unique to
SBITS and is an identical problem in all mass spectrometric methods. Therefore, it is not relevant to consider this when comparing SBITS against other MS methods. The LOD in SBITS is not as low as that of MIDAS or SRM. Data generated in Figure 2.13 show that SRM has a lower limit of detection and a greater linear range. As mentioned earlier, if many targets are to be examined and the ability to schedule SRM transitions is not present, the duty cycle of the MS will become too long to generate reliable quantitative data. Too large a number of targets may even cause a workflow such as MIDAS to be unsuitable if there are a large number of targets and very sharp chromatographic peaks. SBITS, however, will only generate MS/MS data when a threshold has been met. This allows many more targets to be analysed and reproducible quantitative data to be generated on each one. This study shows that SBITS can be used in a targeted manner when a large number of targets need to be analysed. For studies involving only a small number of targets, an SRM method offers lower limits of detection. As only a small number of targets are to be monitored in this project and the ability to schedule SRM’s exists on the current instrumentation setup, SRM analysis will be used to monitor all target peptides for the remainder of this project.

2.5 Conclusion
This chapter describes testing of a novel mass spectrometry method, termed Signal Boosting and Ion Trap Scanning (SBITS) using a 4000 Q-TRAP mass spectrometer. In this chapter, we have determined that in the SBITS workflow, in order for this method to be as sensitive as possible when detecting iTRAQ labelled peptides, fragmentation of precursor ions must be performed in Q2 using CID, and Q3 must be used as a linear ion trap.

The primary objective in this project is to monitor the quantity of a synthetic peptide fragment as it becomes cleaved by proteases in serum or plasma. Work done up to this point has illustrated that the most sensitive MS method for detecting samples of this nature is SRM. As the amount of synthetic target produced by proteases in the workflow described in Chapter 1 (Figure 1.17) is unknown, it was decided that the most sensitive method of detection was to be used for this project. Therefore, SRM will be used for the remainder of this project.
Chapter 3: The Development and Optimisation of a Multi-Step Sample Preparation Method to Assay Protease Activity in Biological Fluids by SRM MS

3.1 Introduction
This chapter describes the development of a five step method to monitor protease activity in plasma samples. When developing this assay, trypsin was used as a model protease. This is because it is a well characterised enzyme which specifically cleaves proteins at the C-terminus of lysine and arginine residues, except when either is followed by proline. This chapter outlines the design of a target trypsin substrate which has been engineered to undergo proteolysis at a specific location. In addition, an investigation into the enrichment of small molecular weight (< 10 kDa) peptides in plasma has been performed. Finally, we assess the ability of this five step method to monitor the digestion of a target substrate using selected reaction monitoring (SRM) in order to assay trypsin activity in plasma.

3.1.1 Proteases
Proteases are known to participate in multiple pathways performing important functions in the cell, including breaking down proteins into peptides and eventually into their constituent amino acids. Proteases are also integral to more complicated, highly ordered systems such as the initiation of programmed cell death (210, 211, 272-274). This process, known as apoptosis is discussed in greater detail in section 1.12.1. Regulation of apoptosis is particularly important to avoid the growth of tumours or the proliferation of mutated or damaged cells (275-278). As discussed before, enzymes are substrate specific, therefore different proteases cleave at specific sites within a protein and produce protein fragments with known termini. This is especially important in proteomics as this area of research regularly involves discovering the exact sequence of amino acids within a protein. However, in the case of proteases, their substrate profiles can overlap. This means that a protease can cleave multiple substrates and one protein can be broken down in the same place by more than one protease (161, 279-282). Proteases can also act as “switches” in biological systems, meaning that a protease can cleave a proenzyme into its active form (162, 283, 284). Proteins undergo proteolysis for several reasons, including recycling of their amino acids for further protein synthesis or to modify enzymatic activity. The level and activity of proteases within a cell or system must be kept...
within very strict limits in order to maintain homeostasis (285-288). Because proteases are themselves proteins, they can be the subject of protease digestion. This helps remove proteases which are no longer needed and can control their level within cells. Abnormal protease activity has been linked to several diseases. These diseases including cancer (277, 289-294), Alzheimer’s disease (295-297) and arthritis (298-300).

As biological mass spectrometry advances, it is envisaged that the simultaneous screening of a variety of diseases within a single experiment on a single sample may be a realistic and valuable possibility (179, 301). Monitoring the activities of proteases linked to diseases may provide insight into prognostic indicators and potential treatment strategies (302-305). One potential strategy is monitoring the presence, absence or amount of a known proteolytic cleavage fragment (231). This would allow a comparison between healthy (disease free) subjects and those with a specific disease.

Mass spectrometry (MS) coupled to advanced separation techniques such as High Performance Liquid Chromatography (HPLC) has been successfully applied to many proteomic and metabolomic investigations (306-313) and can be applied to protease analyses.

3.1.2 Project Description and Aim

The aim of this chapter is to develop a broad method to monitor protease activity in plasma samples. Several important studies need to take place to optimise each stage of the method being developed in this project (Figure 3.1).
Figure 3.1: The sample preparation workflow being developed in this study. This method will involve adding a synthetic peptide of known sequence, specifically designed to be cleaved at one site, to a plasma sample in a fixed amount. The plasma sample will then be incubated at 37°C for a fixed time period. Following this, the plasma sample will then be quenched and enriched for peptides. An internal standard will be added to the sample in a known amount. From this, an absolute amount of cleaved peptide generated by the target protease in plasma can be determined. This method allows the comparison of protease activity between different populations (i.e. samples from patients within a specific disease compared to apparently healthy individuals).
The target peptide must be cleaved specifically by the protease of interest for the assay to work. If this peptide is further cleaved by multiple proteases, the amount of cleaved product detected will be reduced. This may indicate reduced target protease activity and potentially provide a false result. Therefore, the target peptide must be cleaved at a pre-designed location by a specific protease. There may be some overlap between different proteases, most likely from members of the same enzyme family (314, 315), such as the ADAM or caspase families of proteases (Source: MEROPS database, http://merops.sanger.ac.uk). However, if the cleavage of the target peptide only takes place at one specific site in the molecule, the total amount of cleaved peptide can be monitored.

The cleaved form of each target peptide must be detected in a quantitative manner. As this project involves the detection of a small number of target peptides, SRM will be used to detect these targets, as it offers lower limits of detection and quantitation than the SBITS method described in the previous chapter.

Due to the complexity of plasma, samples need to be enriched for target peptides. This reduces the complexity of the sample and removes intact protein that can lead to HPLC columns becoming blocked. As several different methods have been published describing how to enrich for peptides (316), a comparative study of these methods will be performed.

### 3.1.3 Synthetic Peptides for Protease Activity Analysis

In order to assay protease activity, a target substrate must be monitored, in this case a synthetic peptide. When designing these peptides, several rules were created. Firstly, the cleaved forms of the peptide substrates to be monitored must be non-endogenous in humans. If the fragment to be monitored is an endogenous peptide in humans the amount of this fragment present in the sample at the start of the assay may vary. Therefore, there is no way to determine how much of this peptide was produced by monitoring protease activity in this assay and may be a source of erroneous results. By ensuring the target peptide being monitored is non-endogenous in humans, this potential source of error is avoided.

As SRM was used to monitor the cleaved form of the tryptic target peptide, the cleaved peptide of interest must not be susceptible to post translational modifications (PTMs) such as methionine oxidation or asparagine or glutamine deamidation. SRMs are specific in terms of their precursor and product ion mass selection; therefore, if the
peptide undergoes a PTM, this will cause the mass of the peptide to alter. As a result, some of the synthetic peptide to be monitored in this assay would be excluded in either the Q1 or Q3 stages of mass selection (Figure 3.2). Therefore, the synthetic peptide should not contain amino acids subject to oxidation (317) or deamidation (318).

Thirdly, the target peptide must be designed in such a way as to be cleaved in one specific place. If a target peptide can be cleaved by a variety of proteases more active than the target protease in question, the peptide may be cleaved in several different places. As a result of this, an accurate measure of the activity of the target protease in question may not be taken and incorrect results may be generated in this assay. In addition to proteases having overlapping specificities and substrate profiles (MEROPS database), exopeptidases are naturally present in plasma (175, 319, 320). Exopeptidases are enzymes which remove amino acids individually from the ends of peptides (321). These can be broken down in to two classes, aminopeptidases (enzymes cleaving single amino acids from the aminoterminus of a peptide), and carboxypeptidases (enzymes cleaving single amino acids from the carboxyterminus of a peptide). Therefore, the synthetic peptide target used in this experiment could also be digested by these enzymes. The problem of other proteases acting on the target peptide can be overcome by using the other stereoisomer of non-target amino acids (175). Villanueva et. al. demonstrated that by using the non-naturally occurring isoform of an amino acid (the D-isoform), proteolytic digestion is inhibited at those amino acids (175). By end-capping both ends of the target peptide with three D-amino acids the effect of exopeptidases on the target synthetic peptide will be minimised. This increases the specificity of proteolysis of the peptide and, combined with making the central L-amino acid sequence one recognised by the protease of interest (MEROPS database), maximises the chance of proteolysis taking place in one area by the target protease.

In theory, by having every amino acid in the synthetic peptide as the opposite enantiomer (the D-isoform) to that which is naturally present, proteases cannot act on these acids. By having L-amino acids only at the desired cleavage site, only the protease of interest can act upon the synthetic peptide. As a result of this, all the synthetic peptide which has been cleaved in the sample can only be acted on by the target protease. Therefore, by having all amino acids present in the molecule with the exception of amino acids at the protease’s substrate consensus sequence, the problems
of the synthetic peptide being acted on by exopeptidases or other proteases is
minimised. This targeted digestion will mean that the substrate is only acted on by its
target protease. This will increase the signal to noise of the assay, leading to a more
sensitive method. The use of D-amino acids surrounding the L-amino acid consensus
sequence for hydrolysis has been an effective means of ensuring peptide hydrolysis
specificity (175).
In order to absolutely quantify the amount of cleaved peptide present in the sample,
an internal standard can be employed. Several important parameters to be considered
when selecting the internal standard to be used in this assay. An example of this is the
requirement for a minimal SRM response from endogenous plasma peptides as this
could lead to erroneous results when quantifying the amount of cleaved peptide
produced in the sample against the standard. Additionally, the internal standard must
be cheap and easy to obtain. As the internal standard will be added post sample
quenching and enrichment, the effect of endogenous peptidases is not a consideration.
This is because these peptidases will be removed by the sample enrichment process
before the addition of the internal standard. As such, there will be no proteolysis
taking place on the internal standard. Internal standards will be investigated in
Chapters 4 and 5.
Figure 3.2: SRM workflow. This shows that a molecule of one specific m/z is selected in Q1. The molecules of this specific m/z enter Q2, where they are fragmented by collisionally induced dissociation (CID). Following this, all fragments are passed into Q3. In Q3, another stage of m/z selection takes place, so only molecules of one specific m/z can travel through Q3 and reach the detector.

3.1.4 Target Peptide Recovery and Enrichment

As described in Chapter 1, plasma is a very complex mixture, with a protein concentration range spanning at least ten orders of magnitude (322). In order to reduce the complexity of the sample, it is desirable for the proteins to be removed. It has also been shown that overloading a liquid chromatography column can lead to blocking of the column or reduction in its performance making it not fit for purpose. By removing these larger proteins, this problem is overcome.

Amongst these proteins will also be proteases, including the protease(s) being monitored. By removing these proteases before the sample is injected on to the LC-MS system and ensuring this is done in a reproducible manner, this source of potential sample variation (varied substrate/enzyme incubation time) is removed.

Work by Kay et. al. (323), as well as Chertov et. al. (316, 324) has shown that precipitation can remove larger proteins. This is confirmed by Alpert et. al. who showed that adding 1 volume of acetonitrile to 1 volume of a serum sample can apparently remove all proteins with a molecular weight of > 20 kDa. This molecular weight cut-off is lowered to 6 kDa if two volumes of acetonitrile is used (325). Villanueva et. al. also demonstrated that C8 trapping can be used to enrich peptides in serum (175). Another alternative to precipitation or C8 trapping is to perform filtration with a molecular weight cut off, as demonstrated by Tirumalai et. al. (322).
As the proteases acting on the target peptides will have molecular weights greater than 20 kDa (326-328), all of these methods can all be used to remove high molecular weight proteins, including the target proteases.

Several different methods, including those described above, for enriching peptides from complex samples have been published in recent years (316), it is important to use the method which offers the most reproducible results with the minimum peptide losses. If more than one method offers similarly low variance, plainly the method which offers the greatest peptide retention with the most intact protein removed will be favoured.

In order to assess which method offers the most efficient removal of intact protein, the samples generated by each enrichment method will be analysed by SDS-PAGE, followed by silver staining. By running each method alongside a standard of plasma which has not undergone any enrichment, the amount of protein remaining in each sample can be assessed (Figure 3.3). As described previously, intact protein can cause damage to nano-LC systems. Therefore, only methods showing minimal amounts of intact protein will be examined for peptide retention.
Figure 3.3: The workflow to assess intact protein removal from plasma. In this experiment, aliquots of equal volume will be taken from a stock of plasma from healthy volunteers and enriched by one of the methods published. Following this, the enriched plasma will be dried to completion, re-suspended in PBS and run on a SDS-PAGE gel. This gel will then be silver stained to detect intact protein. The enriched samples showing the least amount of staining indicate the most efficient removal of intact protein. These methods showing the least staining will be investigated further to determine their ability to retain smaller peptides (reproduced from Potier et. al., reference (329) with permission from Analytical Chemistry).
As peptides have their own individual properties such as hydrophobicity based on their own individual amino acid sequences, the chemical properties of different peptides may be considerably different. Therefore, a range of different peptide standards need to be used in order to find the best method to retain peptides in a plasma sample. Because of this, a series of peptides derived from a QconCAT protein will be used. QconCAT was used by Pratt et al. to quantify a variety of target peptides in a single assay (330). QconCAT is a protein consisting of a series of several tryptic peptides (330). Following digestion of this protein with trypsin, a range of peptides are released in an equimolar amount. The QconCAT protein used in this study will generate a series of thirteen peptide standards, each with a different amino acid sequence and all in an equimolar amount.

In order to determine quantitatively the losses from each enrichment method, an internal standard for each peptide must be added to the sample post-enrichment. However, the standards added post-enrichment must be added in a fixed amount and these post-enrichment standards must be differentiated from the peptides added pre-enrichment. Therefore, mTRAQ labelling will be used in this study. mTRAQ was first used by Holzmann et al. to accurately determine the amount of MP1-p14 complex present in standard samples as an alternative to the AQUA strategy (331). mTRAQ is a non-isobaric labelling technique similar to iTRAQ, leading to a mass shift of 4 Da between each label. As a result of this, quantitation between a standard peptide and a sample peptide can be made. In this experiment the use of mTRAQ enabled a set of standards to be made in order to assess the loss of peptide from each enrichment method. A workflow of the experiment to determine peptide losses is shown in Figure 3.4.
Figure 3.4: The workflow to assess peptide retention in plasma. This involves digesting a QconCAT protein, producing a series of target peptides in an equimolar amount. This QconCAT protein digest was split into two equal aliquots and each labelled with a different mTRAQ tag. The $\Delta_4$ labelled peptides were spiked into plasma to be tracked through each enrichment process, whilst the $\Delta_0$ peptide series were used as internal standards added to the sample post-enrichment. In order to assess labelling efficiency, a series of samples containing both mTRAQ sets of peptides in an equimolar amount were prepared. This allows samples to be normalised against these standard samples, allowing any difference in labelling efficiency between mTRAQ stocks to be accounted for (reproduced from Potier et. al., reference (329) with permission from Analytical Chemistry).
Figure 3.5: The data processing workflow to assess peptide retention. By finding the ratio between the $\Delta_4$ and $\Delta_0$ set of mTRAQ labelled peptides without any enrichment taking place, any difference in labelling efficiency between the two sets of peptides can be accounted for. The samples prepared to assess peptide recovery in plasma can then be normalised to the $\Delta_0:\Delta_4$ ratio for these standard samples. From this, the peptide loss from each enrichment method can be accurately determined.
In order to determine the losses from each process, the SRM transition with Q3 focused on the m/z of the mTRAQ reporter ion of each peptide will be examined. The method offering the most efficient intact protein removal with the more reproducible and efficient peptide retention based on this study will be used for the remainder of this project.

3.1.5 SRM Transitions
When analysing each target peptide, as well as generating reliable quantitative data, it is important that the peptide in question is identified confidently. Whilst selected reaction monitoring is a selective process, there is still the possibility of more than one co-eluting peptide having the same SRM transition (262). Therefore, more than one SRM transition is required to comprehensively identify a peptide of interest. The MIDAS method developed by Unwin et. al. performs an MS/MS scan over a wide mass range once a given SRM intensity threshold has been met (254), allowing the peptide to be sequenced. Nonetheless, Sherman et. al. stated that the minimum number of transitions required to confidently identify a peptide is four (332). Domon et. al. have previously shown that three SRM transitions correspond to approximately 50% of the total ion count of a peptide, whilst six transitions cover around 80% of a peptide’s product ion spectrum (333). As well as providing more comprehensive peptide identification, using six SRM transitions allows five internal replicates for quantitation if necessary, as opposed to three in Sherman’s criteria this means that if there is a problem with SRM interference in one transition, more internal replicates are present to perform single, summed or averaged SRM quantitation. Therefore, six SRM transitions will be used to identify each peptide of interest, with the most intense transition being used to generate quantitative data. In order to increase the selectivity of this assay, several of these transitions will have the m/z of Q3 greater than the m/z of Q1. When a peptide undergoes fragmentation, the charge state of these product ions is often lower than that of the precursor ion (334). This means that product ions containing large numbers of amino acids will have a higher m/z than the precursor ion of that peptide. However, shorter product ions containing only a few amino acids will have a lower m/z than the precursor ion. Because these lower mass product ions consist of fewer amino acids, the likelihood of other peptides not of interest containing this short amino acid sequence is relatively high. This makes an identification using these shorter amino acid product ions less confident. By using
product ions consisting of longer amino acid chains, the chance of a contaminant peptide having the same amino acids present is reduced, leading to a more confident peptide identification (262).
3.2 Materials & Methods

3.2.1 MALDI-ToF/ToF Analysis
Peptide digestion testing was performed on a 5800 MALDI-ToF/ToF (Applied Biosystems, Foster City, CA) mass spectrometer using an Opti-TOF® 384 well MALDI plate insert (123 x 81 mm) (Applied Biosystems, Foster City, CA). All samples were spotted manually. The matrix used was α-cyano-4-hydroxycinnamic acid (≥ 98% purity, TLC grade) (Sigma-Aldrich, St. Louis, MO) dissolved in 79.9% acetonitrile, 20% water and 0.1% trifluoroacetic acid (reagent grade, ≥98% (titration)) (Sigma-Aldrich, St. Louis, MO).

3.2.2 LC-SRM Analysis
All mass spectrometric (MS) analysis, excluding peptide digestion testing, was performed on a 4000 Q-TRAP (Applied Biosystems, Foster City, CA) coupled to a liquid chromatography system consisting of a Famos autosampler, a Switchos trap system and an Ultimate Plus dual pump (LC Packings, Dionex, Sunnyvale, CA). The column used on the liquid chromatography system was an Acclaim pepmap 100, 15 cm, C18 column, pore size 100 Å. The software used for the 4000 Q-TRAP was Analyst 1.4.1 (Applied Biosystems, Foster City, CA). The software for the liquid chromatography system was Chromeleon 6.50 SP4 Build 1000 (LC Packings, Dionex, Sunnyvale, CA). Data were processed using MultiQuant 2.0 software (AB Sciex, Toronto Canada) using the summation method of integrating peak areas.
To create the LC buffers A and B, Switchos buffer and needle wash solution, water was HPLC grade (Rathburn, Walkerburn, UK), acetonitrile was HPLC grade (Fisher Scientific UK Limited, Leicestershire, UK) and formic acid was mass spectrometry grade (Fluka, Sigma-Aldrich, St. Louis, MO).
Buffer A & Switchos buffer consisted of 97.9% water, 2% acetonitrile and 0.1% formic acid.
Buffer B consisted of 79.9% acetonitrile, 20% water and 0.1% formic acid.
The needle wash solution was 100% water.

3.2.3 Shotgun Peptidomics Analysis
All shotgun peptidomic peptide enrichment samples were analysed on an LTQ Orbitrap Velos (Thermo Fisher Scientific, Hemel Hempstead, UK) coupled to a nanoAcquity UHPLC system (Waters, Manchester, UK). In this experiment, CID was
performed in the linear ion trap on the top 20 precursors at 30,000 resolution (precursor m/z 300 – 2000) with a minimal signal of 5000 cps and a charge state of 2+ and greater after being detected in the orbitrap. Precursors were excluded for 30 seconds after two occurrences. Product ions were scanned over m/z 100.0 – 1700.0 at 7,000 resolution. Orbitrap source voltage was 1500.0 V and source temperature was 200.0 °C. The software controlling the orbitrap was Xcalibur version 2.1 (Thermo). The nanoAcquity was controlled by MassLynx version 4.1. The loading flow was 15 µl/min and the gradient flow rate was 300 nl/min. Buffer A was 0.1% formic acid in water, and buffer B was 0.1% formic acid in acetonitrile. The loading buffer was 20 mM citric acid. The gradient increased from 3% B to 40% B over 90 minutes. Following this, the % B increased to 90% over 3 minutes where it was held at this level for 15 minutes. After this, the % B was reduced to 3% where it is held constant for 21 minutes. The analytical column was a 1.7 µm BEH130 C18 column, 75 µm x 250 mm (Waters) and held at a constant temperature (40°C).

3.2.4 Synthetic Peptide Design
An E. Coli tryptic digest was analysed in an Information Dependent Acquisition (IDA) experiment using the 4000 Q-TRAP. This was Mascot searched and a peptide of seven amino acids identified with ≥ 95% confidence was selected. This sequence was then BLAST searched against the human proteome in order to ensure this peptide sequence was not present in any known human protein. Following this, the sequence was extended up to its next tryptic site based on the amino acid sequence of this peptide’s parent protein. All amino acids were chosen to be the D-isoform except the site of the tryptic cleavage site for peptide SP1. For peptide SP2, the next amino acid towards the C-terminus was also an L-isoform. The sequences for SP1 and SP2 were etaeaylGGkvdtav and etaeaylGGKVdtav respectively, where amino acids in lower case correspond to the D-isoform (glycine is achiral and as such only has one isoform).

3.2.5 Peptide Digestion
First, an aliquot of each the peptides SP1 and SP2 (1 pmol) were prepared. The synthetic peptide SP1 consists of the amino acid sequence etaeaylGGKvdtav, and the peptide SP2 has the sequence etaeaylGGKVdtav. In the case of these peptides, all amino acids written in lower case have the non-naturally occurring isoform (i.e. the D-isoform) of the amino acid. These two peptides were both split into four equal aliquots. To the first three, 1 µl trypsin was added, whilst the fourth had 1 µl of PBS
added to it. The first aliquot of both peptides underwent an overnight incubation at 37°C. The second underwent a microwave digestion at 55°C for 15 minutes. The third was subjected to a microwave digestion followed by overnight incubation at 37°C, whilst the fourth was left on the bench overnight. Each sample was then dried to completion, re-suspended in Buffer A, and analysed by mass spectrometry to find the cleaved and uncleaved forms of each peptide.

3.2.6 SRM Generation & Optimisation
The MS/MS spectrum for each peak in the Total Ion Chromatogram (TIC) of the partially digested peptide was manually sequenced to confirm peptide identity and to obtain six product ions to be used in further analysis. Whilst some of these Q3 m/z values were lower than the Q1 m/z, these transitions were used for identification purposes only. In addition to this, the precursor m/z of each peptide, along with their retention times were obtained.

In order to optimise the collision energy of each transition, the same sample was re-injected, this time performing an SRM only experiment. This SRM experiment contained the Q1 & Q3 m/z for each target transition (Table 3.1), along with the collision energy used in the IDA experiment (26.0 eV). In addition to this, transitions were generated over a range of ± 5 eV, with steps of 0.5 eV to optimise their collision energies (Table 3.1). The dwell time for all transitions is 100.0 ms.

Table 3.1: Optimised SRM transitions to detect etaeaylGGK

<table>
<thead>
<tr>
<th>Q1 (m/z)</th>
<th>Q3 (m/z)</th>
<th>Collision Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>519.8</td>
<td>608.3</td>
<td>26.5</td>
</tr>
<tr>
<td>519.8</td>
<td>261.2</td>
<td>27.0</td>
</tr>
<tr>
<td>519.8</td>
<td>537.3</td>
<td>24.5</td>
</tr>
<tr>
<td>519.8</td>
<td>374.2</td>
<td>24.0</td>
</tr>
<tr>
<td>519.8</td>
<td>737.4</td>
<td>25.0</td>
</tr>
<tr>
<td>519.8</td>
<td>808.4</td>
<td>25.0</td>
</tr>
</tbody>
</table>

3.2.7 SRM Linearity
First, the sample shown in section 4.4: SRM Generation underwent a serial dilution across four orders of magnitude. The dilution of each sample relative to the most concentrated is shown in Table 3.2.
Table 3.2: The relative amounts of etaeaylGGK used in SRM linearity testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Relative Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>500</td>
</tr>
<tr>
<td>7</td>
<td>1,000</td>
</tr>
<tr>
<td>8</td>
<td>5,000</td>
</tr>
<tr>
<td>9</td>
<td>10,000</td>
</tr>
</tbody>
</table>

Each sample was prepared in triplicate and analysed by the optimised SRM method. The linearity of this SRM method was assessed by plotting the peak area of the most intense transition against relative sample concentration. This allowed the linearity of SRM transitions monitoring the N-terminal cleaved form of the SP2 peptide synthetic to be assessed.

3.2.8 Plasma Sample Preparation
The plasma was obtained from healthy volunteers and centrifuged at 16,000 g for 5 minutes at 4°C. The supernatant was removed, diluted four-fold with PBS and spin filtered (0.22 µm cut-off) at 1500 g for 30 seconds at 4°C. The resulting filtrates were pooled and aliquotted out into 25 µl samples.

3.2.9 Peptide Enrichment
All methods for this investigation have been reproduced from Potier et. al., reference (329) with permission from the journal.

3.2.9.1 Sample Preparation to Monitor Protein Depletion
To prepare samples by precipitation, first an aliquot of plasma (25 µl) was removed from the stock of plasma prepared previously. Following this, either one volume (25 µl) or two volumes (50 µl) of either acetone or acetonitrile was added to the sample and the resulting mixture was vortexed for ten seconds. The sample was then centrifuged at 10,000 rpm for 30 seconds. The supernatant was removed and dried to completion by low pressure, elevated temperature centrifugation and re-suspended in PBS (25 µl). The supernatant was re-suspended in PBS by repeated pipetting up and down until all protein was re-suspended.
To prepare samples by filtration, an aliquot of plasma (25 µl) was diluted to a total volume of 1000 µl with PBS and placed in a 5 kDa molecular weight cut-off, 4 ml spin concentrator for proteins (Agilent, Santa Clara, CA). The filter was then centrifuged at 3,200 g at 10°C for 30 minutes. The filtrate was then removed and placed in an eppendorf. The residue was placed in a separate eppendorf, and all samples were dried to completeness as described earlier. Each sample was then re-suspended in PBS (25 µl).

To prepare samples by solid phase extraction with C₈ coated magnetic beads, Samples were prepared by adding C₈ coated magnetic beads, gentle agitation, removing the beads and treating with 0.1% TFA. These samples were then dried to completion and re-suspended in PBS.

All samples, once prepared, were incubated with 4x Laemmli buffer (2 µl) at 100°C for 2 minutes. All samples were left to cool to room temperature before being loaded onto the SDS-PAGE gel.

3.2.9.2 SDS-PAGE

To perform 1D-PAGE, two gel plates were clamped together and placed on an impervious surface. Resolving gel was then carefully poured between the plates until a gap of 1cm below the end of the sample lane generating comb was left. Then water was added to the top of the resolving gel, ensuring that the gel/water bilayer was straight. Once the resolving gel had set, the water was removed and the plates were topped up with stacking gel. The sample lane generating comb was carefully added to the stacking gel. Once the stacking gel had set, the comb was carefully removed, the plates were unclamped, and the rubber seal was carefully removed. The plate stack was placed in a gel tank and clamped in place on each side of the plates. The tank was filled with running buffer. Samples were loaded into the lanes of the resolving gel. In addition to this, 5 µl of full range rainbow molecular weight markers (GE Healthcare, Buckinghamshire, UK) was added to a separate lane. Electrophoresis was then started with a voltage of 60 V until the entire sample had entered the resolving gel. The voltage then was increased to 200 V until the bromophenol blue had run off the bottom of the gel. The resolving gel consisted of acrylamide solution (3.3 ml), 1.5M Tris (pH 8.8, 2.5 ml), sodium dodecyl sulphate solution (10% w/v) (100 µl), water (4 ml), 10% ammonium persulphate (50 µl), and TEMED (3.3 µl).
The stacking gel consisted of acrylamide solution (0.67 ml), 0.5M Tris (pH 6.8, 1.25 ml), sodium dodecyl sulphate solution (10% w/v) (50 µl), water (3.1 ml), 10% ammonium persulphate (25 µl) and TEMED (5 µl).

The running buffer consisted of Tris (3.1 g), glycine (14.4 g), sodium dodecyl sulphate (1 g) and water (1 litre).

Silver staining of all SDS-PAGE gels was performed as described by Shevchenko et al. (335).

3.2.9.3 Peptide Standard Preparation
Two aliquots of QconCAT protein (100 pmol, 5 µg) were each digested by overnight incubation with trypsin (0.5 µg) at 37°C. Following this, both digested samples were dried to completion and re-suspended in tetraethylammonium bicarbonate (1.0M, pH 8.5, 30 µl). mTRAQ™ reagent (Applied Biosystems, Foster City, CA) was dissolved in propan-2-ol (35 µl) and was added to the sample. The mTRAQ™ container was rinsed with propan-2-ol (35 µl) and this was also added to the sample. This mixture was vortexed for ten seconds and left at room temperature for a minimum of 2 hours. All solvent in the sample was removed via low pressure, elevated temperature centrifugation. The samples were re-suspended in PBS to a concentration of 1 pmol/µl.

3.2.9.4 Sample Preparation to Monitor Peptide Losses
Plasma (25 µl) containing Δ4 labelled peptides (1 pmol) underwent either filtration or precipitation with two volumes of acetonitrile, as described when preparing samples for protein removal. Each depleted sample was pooled with a Δ0 labelled untreated series of standards (1 pmol) and the ratios of the mTRAQ™ reporter ions for each peptide were examined in order to assess peptide losses from each enrichment method.

3.2.9.5 Sample Preparation to Monitor Endogenous Plasma Peptide Retention
Plasma (25 µl) underwent either filtration or precipitation with two volumes of acetonitrile, as described when preparing samples for protein removal. Each depleted sample was dried to completion and re-suspended in Buffer A. Each sample was then analysed using the shotgun proteomics workflow described above.
3.2.10 Data Processing
All LC-SRM data was processed via MultiQuant software Version 2.0 (AB Sciex). Peak integration was manually inspected to confirm the integration had been performed correctly. The peak area for the most intense transition for each peptide was compared between the $\Delta_0$ and the $\Delta_4$ labelled samples and the ratio found between the two for each standard run. An average ratio between the two labels was found by averaging the results for all standards injected. This was performed on every target peptide. Following this, the same process of finding an average ratio between labels was performed on the enriched samples. In order to determine the loss of peptide from each enrichment process, the ratio for these samples was compared to the average standard ratio. If any peptide was lost due to this enrichment method, there would be less $\Delta_4$ labelled peptide in the enriched samples. This would affect the $\Delta_4:\Delta_0$ ratio in these samples. Therefore, when this ratio was compared to the standard ratio, this value would become less than 1.00. Therefore, the nearer this final ratio is to 1.00, the more target peptide has been retained in the enrichment process. From these results, the percentage loss for each peptide, as well as the variance for each method was obtained.

All shotgun peptidomic data were processed by Proteome Discoverer 1.1 (Thermo) using the following criteria; Data were searched against the Homo sapiens proteome using the UniProt database, and a no enzyme specificity search. The strict and relaxed target false discovery rates (FDRs) were 0.01 (1%) and 0.05 (5%) respectively. Precursor mass tolerance was 15 ppm and fragment mass tolerance was 0.8 Da. Methionine oxidation as well as asparagine and glutamine deamidation were set as dynamic modifications. All peptides with an ion score $\geq$ 20 and assigned a protein accession number were counted as being identified, with one or more of these peptides being used to assign the identification of a protein.

3.2.11 Proof of Concept Testing
First, the synthetic peptide SP2 (2 µg) was added to plasma (25 µl), along with trypsin in different enzyme: substrate ratios (Table 3.3) Each enzyme concentration was analysed 24 times, as this study involves eight different time points, each one in triplicate. In order to keep the total volume of each sample the same (45 µl), samples with less trypsin had the corresponding amount of PBS added to them.
Table 3.3: The amounts of trypsin added to plasma in proof of concept testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of Trypsin Added to Sample (µg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>0.04</td>
</tr>
<tr>
<td>4</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Samples were incubated at 37°C. At each time point (described by the table in Table 3.4), Samples were removed and quenched with formic acid (5 µl).

Table 3.4: The time points at which samples were taken in proof of concept testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Time Point (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>24</td>
</tr>
</tbody>
</table>

In the case of samples at the 0 hour time point, formic acid was added to quench the sample before the trypsin was added. Samples were stored at -20°C following the addition of formic acid until all samples for this experiment had been collected. Samples were then enriched for peptides by acetonitrile (100 µl), as described in section 4.6.1 and dried to completion. Each was then re-suspended in Buffer A (1000 µl). 2 µl was placed into a vial and diluted further with Buffer A to a total volume of 55 µl. All 55 µl was then injected onto the HPLC system. All samples were then monitored by SRM analysis, using the SRM transitions and collision energies described in section 4.4. Following this, the peak areas for the most intense transitions were processed using MultiQuant software, version 2.0 (AB Sciex, Cheshire, UK). The peak areas for each time point were averaged and plotted against time for both the cleaved and the uncleaved forms of SP2.
3.2.12 Sample Dilution Testing
First, a portion of an enriched sample taken at 24 hours (2 µl) underwent a serial dilution across four orders of magnitude and each sample analysed by SRM. The peak areas for each sample were then processed using MultiQuant software. Each peak area was examined against the linearity plot of the uncleaved peptide shown in section 5.2. From this the dilution of the neat enriched sample necessary to generate reliable quantitative data was found.

3.2.13 HPLC Column Degradation
Following the injection of all 96 enriched neat plasma samples, three injections of the 6 protein mix digest used to routinely assess the performance of the HPLC column were performed (20 fmol digested protein injected onto the column per injection) and the column’s performance was assessed as per laboratory SOP. This involved Mascot searching the resulting data using the following parameters. The database searched was UniProt and a taxonomy of all entries was used with trypsin as the proteolytic enzyme used to digest the sample. Peptide tolerances of ± 1.2 Da were used for peptides of charge states 2+ and 3+ with an MS/MS tolerance of ± 0.9 Da. If the Mascot search identified all six proteins (serotransferrin, alcohol dehydrogenase, cytochrome c, β-galactosidase, bovine serum albumin and lysozyme c), and the peak widths of five of the identified peptides were assessed and checked to see if they were less than 30 seconds wide, the column is deemed fit for purpose.
3.3 Results

3.3.1 Synthetic Peptide Design and Digestion

A synthetic peptide was designed to be digested by trypsin. The amino acid sequence for this peptide was ETAEAYLGGKVDTAV. This was predicted to be cleaved at the C-terminal side of lysine (K), producing two peptides – ETAEAYLGGK and VDTAV. In order to assess the ability of D-amino acids at inhibiting proteolysis, two peptides of this amino acid sequence; etaeaylGGKvdtav (hereafter known as SP1), and etaeaylGGKVdtav (hereafter known as SP2) were incubated overnight at 37 °C with trypsin. All amino acids in lower case signify the D-isoform of this amino acid. Glycine is achiral and as such, it is always in its naturally occurring isoform. When BLAST searching the amino acid sequences of ETAEAYLGGKVDTAV, ETAEAYLGGK and VDTAV against the human proteome using the Expasy database (www.expasy.org), no protein hits were reported. This indicates that the intact form of this target peptide, as well as the predicted cleavage products of this substrate post-tryptic digestion, are not endogenous in humans.

The results in Figure 3.6(a) show that the presence of D-amino acids inhibits digestion of peptide SP1 by trypsin. However, by changing the amino acid next to the site of trypsin cleavage (towards the C-terminus), digestion can take place, albeit at a reduced rate (Figure 3.6(b)). This is shown by the presence of a peak in the HPLC-UV trace at 36.0 minutes when examining SP2 post-tryptic digestion. This peak is not present in SP1 post-digestion with trypsin. Therefore, this peptide was examined by MALDI–ToF/ToF and was confirmed to be the cleaved form of the tryptic substrate, sequence etaeaylGGK (Figure 3.6(c)). The peptide vdtav was not found.
Figure 3.6: The effect of D-amino acids on proteolysis using trypsin. (a) A UV spectrum showing the results of trypsin digestion of SP1, and (b) the results of trypsin digestion of SP2. (c) Sequence coverage of the “Cleaved” form of SP2, showing this is in fact the cleaved form of the peptide. The presence of the cleaved form of SP2 but not SP1 shows D-amino acids inhibits protease digestion.
3.3.2 SRM Generation & Optimisation

In addition, the background response for the SRM transitions to be monitored in this study from other species naturally occurring in plasma must be minimal. If a peptide naturally present in plasma coincidentally has an identical SRM transition the same as the non-endogenous target peptide, this could lead to a source of error in this assay. This problem would be magnified by the potential sample variation described previously and if the contaminating peptides co-elute with the target peptide of interest. In order to overcome this problem, theoretical SRM transitions were generated using Protein Prospector. These transitions were tested on a pooled plasma sample generated from ten healthy volunteers (Seralab, Haywards Heath, UK). By ensuring the target peptide fragment is not endogenous in humans, with minimal response from endogenous peptides present in plasma these potential sources of sample variation are avoided.

The six transitions with the most intense SRM response were selected and their collision energies optimised (Table 3.2). These transitions were then used to monitor this peptide in all future experiments. The most intense transition will be used for quantitation, whilst the remaining five will be used to provide identification of the peptide. Whilst some of these transitions have a Q3 m/z < Q1 m/z, these transitions are only to be used to assist in the identification of the peptide. It was also shown that these transitions have minimal interference from endogenous peptides in plasma (Figure 3.7) and as such will not be the cause of erroneous results due to their contaminating SRM data.

An enriched plasma sample being analysed by these transitions showed no significant response (Figure 3.7). A significant SRM response is one whose intensity exceeds 100 counts per second (cps). This indicates that all responses seen from these transitions will be as a result of the presence of the target tryptic substrate being monitored in this assay, and not as the result of contamination of endogenous peptides.
Figure 3.7: SRM responses from endogenous plasma peptides when detecting the peptide etaeaylGGK. The lack of a significant SRM response indicates that there are no endogenous peptides present in plasma which can be detected by these transitions. Therefore, all response detected in this assay is as a result of the proteolysis of the target tryptic substrate, and not from endogenous plasma peptides.

3.3.3 SRM Linearity
The cleaved form of SP2, sequence etaeaylGGK showed a linear SRM response spanning four orders of magnitude ($r^2 = 0.9980$) with a limit of detection of 1 fmol, a lower limit of quantitation of 5 fmol and an upper limit of quantitation of >10 pmol (Figure 3.8). Therefore, samples producing a response with a peak area in the range of $5.97 \times 10^4 – 9.03 \times 10^7$ can be quantified in this assay.
Figure 3.8: etaeaylGGK SRM linearity. *Error bars represent the standard deviation of each sample population (n = 3).*

### 3.3.4 Peptide Enrichment Investigation

In order to assess which of the published methods are best at removing intact protein (175, 316, 322-324), samples from a pooled stock of plasma generated from three healthy volunteers were enriched by each method and the resulting samples ran on an SDS-PAGE gel (10%). The gel was then silver stained in order to detect intact protein. If a method is efficient at removing proteins, less staining will be seen. The results in Figure 3.9 show that precipitation with 2 volumes of acetonitrile and filtration offer the best way of removing intact protein from a plasma sample. The high amount of albumin present in lanes 2, 3 and 8 lead to inverted staining of the gel.
Following these results, a series of ten tryptic peptides were analysed by SRM analysis in order to determine their retention in plasma post-enrichment. As filtration and precipitation with two volumes of acetonitrile offer the most efficient protein removal (Figure 3.9), these two methods were assessed for peptide retention using these ten peptide standards.

However, this workflow (Figure 3.5) assumes that the QconCAT peptide standards will be mTRAQ labelled to the same degree. In reality, this may not be the case. Therefore, a series of standard samples were analysed. These standards consisted of each target peptide with both mTRAQ labels injected in an equal amount and monitored by SRM. By calculating the mTRAQ ratios of each peptide in these standard samples, the ratios from enrichment samples can be normalised to these external standard ratios (Figure 3.5).

Following the SRM analysis of these ten target peptides once they have been added to plasma and the plasma enriched by either filtration or acetonitrile precipitation, it is
shown that acetonitrile precipitation offers greater peptide recoveries with less variance than filtration (Table 3.5).

Table 3.5: Target peptide recoveries of plasma samples enriched by filtration and acetonitrile precipitation. (Reproduced from Potier et al., reference (329), with permission from Analytical Chemistry).

<table>
<thead>
<tr>
<th>Peptide Sequence</th>
<th>Filtration</th>
<th>Precipitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recovery (%)</td>
<td>Variance (%)</td>
</tr>
<tr>
<td>AADDEPEYEDGR</td>
<td>29.1</td>
<td>17.5</td>
</tr>
<tr>
<td>AETIGEK</td>
<td>5.1</td>
<td>127.0</td>
</tr>
<tr>
<td>YEAVPADASSSSEVK</td>
<td>65.6</td>
<td>45.0</td>
</tr>
<tr>
<td>LRPLLEK</td>
<td>16.9</td>
<td>10.1</td>
</tr>
<tr>
<td>APGGEDEEEGVGGGGGGELELR</td>
<td>8.9</td>
<td>75.7</td>
</tr>
<tr>
<td>ASLTTGPSEYSSPSVISVSK</td>
<td>2.6</td>
<td>61.1</td>
</tr>
<tr>
<td>SWLPVVVIK</td>
<td>0.4</td>
<td>63.8</td>
</tr>
<tr>
<td>LWIWEK</td>
<td>0.2</td>
<td>101.6</td>
</tr>
<tr>
<td>STGSFPFPYVLEPLGASPSETS</td>
<td>1.8</td>
<td>4.9</td>
</tr>
<tr>
<td>AFYPASLSPPAAGTAASTALLR</td>
<td>0.6</td>
<td>60.8</td>
</tr>
<tr>
<td>AVERAGE</td>
<td>13.1</td>
<td>56.7</td>
</tr>
</tbody>
</table>

This shows that in the majority of cases (8/10), more peptide is recovered using acetonitrile precipitation rather than filtration. Additionally, in 60% of cases, the variance in peptide recovery is lower in acetonitrile precipitation than filtration.

Finally, a discovery experiment was performed using an Orbitrap Velos (Thermo) in order to determine whether filtration or precipitation with two volumes of acetonitrile is best at retaining endogenous, non-tryptic peptides in plasma. This study identified 354 peptides in precipitated samples and 14 in filtered samples. This equates to 95 and 9 degraded proteins identified respectively (using 1 unique peptide or more to identify a protein). This indicates that endogenous plasma peptides can still be profiled following enrichment with acetonitrile precipitation or filtration, with acetonitrile precipitation leading to a greater number of peptides and proteins being identified.

Statistical analysis (Welch’s t-test for two sample sets with unequal variance) on this peptide retention data showed that of the ten target peptides examined, six show a significant difference in peptide retention (p < 0.05) (Table 3.6).
Table 3.6: Statistical analysis of peptide retention data

<table>
<thead>
<tr>
<th>Peptide Sequence</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AADDEPEYEEDGR</td>
<td>0.74</td>
</tr>
<tr>
<td>AETIGEK</td>
<td>0.33</td>
</tr>
<tr>
<td>YEAVPADASSSSEVK</td>
<td>0.40</td>
</tr>
<tr>
<td>LRPLLEK</td>
<td>0.02</td>
</tr>
<tr>
<td>APGGEDEEEGVGGGGLGER</td>
<td>0.36</td>
</tr>
<tr>
<td>ASLTPGEYSSPSVISVSK</td>
<td>0.01</td>
</tr>
<tr>
<td>SWLPVVK</td>
<td>0.00</td>
</tr>
<tr>
<td>LWIWEK</td>
<td>0.01</td>
</tr>
<tr>
<td>STGSFPFPYVLEPLGASVETSK</td>
<td>0.01</td>
</tr>
<tr>
<td>AFYPASLSPPAAAGTAASLSTALLR</td>
<td>0.01</td>
</tr>
</tbody>
</table>

In addition, it appears that acetonitrile precipitation retains endogenous peptides in a more reproducible fashion. When using acetonitrile precipitation, 53% of endogenous peptides are seen in more than one replicate, with 34% of peptides being seen in all three replicates. This compares to 36% of peptides being seen in more than one replicate, and 29% of peptides being seen in all three replicates when enriching plasma samples by filtration (Figure 3.10). Similar results are seen when assessing the reproducibility of protein identifications from each method, with 77% of proteins being identified in more than one sample when enriching plasma by acetonitrile, with 26% of proteins being identified in all three replicates. This compares to 50% of proteins being identified in more than one replicate when enriching plasma samples by filtration, with 20% being identified in all three replicates. This indicates that the plasma peptidome can be more comprehensively and reproducibly examined using acetonitrile precipitation enrichment as opposed to filtration (text reproduced from Potier et. al., reference (329) with permission from Analytical Chemistry).
Figure 3.10: Peptides and proteins identified in plasma samples enriched by filtration or acetonitrile precipitation. The number of peptides identified by enriching samples by (a) filtration, and (b) precipitation. (c) and (d) show the corresponding number of proteins identified in each replicate when enriching samples by filtration and precipitation respectively. The total number of peptides and proteins found by enriching plasma samples by filtration and acetonitrile precipitation, along with the number of peptides/proteins identified in both methods are shown in (e) and (f) respectively (reproduced from Potier et. al., reference (329), with permission from Analytical Chemistry).
Based on the results from this study, precipitation with two volumes of acetonitrile will be used to remove intact protein and enrich for peptides for all plasma samples to be used in this project.

3.3.5 Proof of Concept Testing
Before analysing all samples, a sample with the highest concentration of trypsin at the final time point of this experiment was diluted and each dilution analysed by SRM. This is because this sample is expected to contain the highest amount of cleaved target substrate. The most concentrated sample injected in this experiment gave a peak area of $3.0744 \times 10^6$. This is comfortably within the linear range of the SRM response for this peptide. Therefore, once a sample is re-suspended in 1000 µl Buffer A, no further dilution is necessary.

The amount of cleaved peptide produced over a 24 hour period from each enzyme:substrate ratio is shown in Figure 3.11. The more active protease (in this case, trypsin) that is present in the sample, the more cleaved peptide substrate is produced. The exception of this is when trypsin is present at its highest amount in this assay. At this point, it is believed that trypsin itself has become a substrate. This means that trypsin is being digested either by other trypsin molecules or other proteases present in the plasma sample and as such becomes inactive. Under these conditions, trypsin cannot cleave its target peptide, SP2. As such, less SP2 is produced in these samples.
Figure 3.11: Proof of concept testing. With the exception of the highest enzyme: substrate ratio, a general trend is seen, with the more trypsin present in the sample, the more cleaved SP2 is produced. In addition to this, the longer the sample is left unquenched, the more cleaved SP2 is produced. Error bars represent the standard deviation of each sample population (n = 3).

It is possible that when 200 ng trypsin is in 25 µl of plasma, it is in such a high concentration that trypsin itself is becoming a substrate. As a result, less active trypsin is in the sample, so less target peptide is digested.

Statistical analysis on these data was performed in order to determine whether there is a significant difference between the amount of trypsin present in the sample, and the amount of cleaved tryptic substrate (etaeaylGGK) produced. This testing was performed at each time point using single factor ANOVA testing. If the Null hypothesis holds true, then there will be no significant increase in etaeaylGGK production at each time point, relative to the amount of trypsin in the sample. A p-value lower than 0.05 indicates that there is a less than 5% chance of the null hypothesis (no significant increase in etaeaylGGK production with increased levels of trypsin at each time point) being true; and therefore, significant changes in etaeaylGGK production are seen with different amounts of trypsin in the sample.
Single factor ANOVA testing indicated no significant change in the level of etaerylGGK production with changing levels of trypsin in the sample until 16 hours. At the 16 and 24 hour timepoints, the change in etaerylGGK production was determined to be significant (P < 0.05) (Table 3.7). However, the data in Figure 3.11 show that the level of etaerylGGK is considerably lower in samples containing the most amount of trypsin (200 ng). A potential cause of this has been suggested; stating that trypsin itself has become its own substrate and is therefore no longer active. Therefore, statistical analysis of all data was re-performed using single factor ANOVA testing on the remaining data (Table 3.8).

**Table 3.7: Statistical analysis of the data shown in Figure 3.11 when analysed by single factor ANOVA testing.**

<table>
<thead>
<tr>
<th>Time Point (hrs)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>0.14</td>
</tr>
<tr>
<td>4</td>
<td>0.17</td>
</tr>
<tr>
<td>8</td>
<td>0.36</td>
</tr>
<tr>
<td>12</td>
<td>0.33</td>
</tr>
<tr>
<td>16</td>
<td>0.05</td>
</tr>
<tr>
<td>24</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**Table 3.8: Statistical analysis of the data shown in Figure 3.11 when analysed by single factor ANOVA testing, excluding data from samples containing 200 ng trypsin**

<table>
<thead>
<tr>
<th>Time Point (hrs)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.56</td>
</tr>
<tr>
<td>2</td>
<td>0.63</td>
</tr>
<tr>
<td>4</td>
<td>0.54</td>
</tr>
<tr>
<td>8</td>
<td>0.36</td>
</tr>
<tr>
<td>12</td>
<td>0.29</td>
</tr>
<tr>
<td>16</td>
<td>0.53</td>
</tr>
<tr>
<td>24</td>
<td>0.31</td>
</tr>
</tbody>
</table>

The data shown in Table 3.8 indicate that when data from the 200 ng trypsin samples have been removed from statistical analysis, no significant difference is seen between samples containing different amounts of trypsin at any time point. Therefore, in this case, the Null hypothesis appears to hold true. However, this analysis is hindered by a limited number of replicates of each sample at each time point. In order to confirm with greater certainty that the Null hypothesis is true, this experiment needs to be
repeated using a greater number of replicates \( (n \geq 6) \). However, due to limitations in instrument time, this was not possible in this study.

### 3.3.6 HPLC column degradation

After all 96 samples had been injected, the LC system met all parameters when assessing if it remained fit for purpose. This was the case for all three injections of the six protein mix digest standard. This indicates that the peptide enrichment process is efficient enough to allow up to 96 plasma samples to be enriched in one batch with no loss in column performance. No carry-over of any six protein mix peptides or the peptides etaeaylGGK or etaeaylGGKVdtav was seen.
3.4 Discussion
In this chapter, the establishment of a mass spectrometry-based assay of high specificity was developed using the site-specific degradation of a target peptide substrate. We used trypsin in our paradigm experiments as this enzyme is well characterised and used routinely in proteomics (336-346).

3.4.1 Target Peptide Digestion
It is clear that the presence of D-amino acids hindered digestion with trypsin. This has been shown in all cases of the attempted digestion of SP1. From these data (Figure 3.6), the presence of D-amino acids leads to a complete inhibition of digestion. The fact that it appeared that there was the presence of digested SP2 after an overnight incubation with trypsin also shows that the presence of D-isoforms affected digestion, as this cleaved peptide was not seen in SP1 post-digestion.

3.4.2 Peptide Enrichment
When performing peptide enrichment testing, the reproducibility of the technique is of paramount importance. If the process results in variable loss of a target peptide, this assay is of no value. Also important is minimising the losses of the peptides of interest. By using a process which loses the least amount of synthetic peptide, the most accurate results and highest sensitivity can be obtained in this assay. Also to be considered is the process which takes the shortest amount of time. This is because part of this project is ensuring this test is potentially the basis of a high throughput assay in a clinical environment. By observing many different peptides, this gives a more general view of which enrichment technique is best, rather than which is best for one particular peptide and its individual chemistry. Whilst precipitation with two volumes of acetonitrile and filtration offer similar levels of intact protein removal (Figure 3.9), acetonitrile precipitation (2 volumes) offers greater recoveries and lower variance between samples than filtration (Table 3.6). As a result of this, precipitation with two volumes of acetonitrile will be used to remove intact protein and retain peptides in plasma samples, preparing them for MS analysis.

3.4.3 Proof of Concept Testing
As the Figure 3.11 shows, the amount of synthetic peptide cleaved is directly dependent on the amount of time the enzyme has to digest the peptide. The amount
produced is also dependent on the amount of enzyme present in the sample. These results have also shown that the method of adding a synthetic peptide to a complex environment such as plasma and enriching it after any given time point using acetonitrile precipitation and monitoring it by SRM can produce reproducible results. Unexpectedly, samples with trypsin at the highest concentration appear to have the least amount of cleaved SP2 produced. This may indicate that the trypsin itself is becoming a substrate for proteolytic cleavage by trypsin or other proteases in the plasma sample. The presence of a large amount of trypsin may lead to it getting digested preferentially in the sample. If this is the case, it would result in less active trypsin being present in the sample and less cleaved SP2 being produced.

### 3.5 Conclusion

In this chapter, we have developed a method to assay protease activity in plasma samples using SRM mass spectrometry. We designed a target peptide substrate to be cleaved by a specific protease (trypsin) and demonstrated that the presence of D-amino acids inhibits proteolysis at these specific locations. In addition, we have examined a series of different methods published in a variety of scientific journals in order to determine which of these methods is best at removing intact protein, whilst retaining peptides. From these results, we concluded that precipitation of plasma with two volumes of acetonitrile is the best method to do this, with the majority of target peptides (6 out of 10) showing significantly improved recoveries when using acetonitrile precipitation as opposed to filtration ($p < 0.05$). As such, acetonitrile precipitation will be used to enrich plasma samples for peptides in all future experiments. Finally, we have shown that the five-step method described in Figure 3.1 can be used to assay protease activity in plasma samples. Whilst a general trend is seen showing the more trypsin is in plasma, the more cleaved tryptic substrate is produced (with the exception of samples with the highest amount of trypsin), statistical analysis (single factor ANOVA testing) showed no significant increase in cleaved substrate production with respect to increased levels of trypsin with the exception of the 16 and 24 hour timepoints ($p < 0.05$) (Table 3.7). However, the data shown in Figure 3.11 clearly show that samples with the highest amount of trypsin have the lowest amount of cleaved substrate production. This appears to be the cause of error in the statistical analysis, so the data shown in Figure 3.11 was re-analysed without these 200 ng trypsin samples being used in statistical analysis. When this was
done, no significant increase in cleaved substrate production with increased levels of trypsin was seen at any time point, and the Null hypothesis stating “no significant increase in etaeaylGGK production is seen with increasing levels of trypsin at any time point” holds true (Table 3.8). However, due to limitations in instrument time, this experiment was performed in triplicate at all time points and with all enzyme: substrate ratios. Repeating this experiment with increased numbers of replicates would enable more powerful and robust statistical testing to be performed using ANOVA, which would in turn provide a more accurate understanding into the significance of the change in cleaved substrate production in relation to the amount of trypsin in plasma samples.

By proving not only that the amount of cleaved target synthetic peptide present in a sample is enzyme and time dependent, but that the method of enrichment and SRM analysis is successful, this allows proteases with clinical significance to be monitored. The proteases asparaginyl endopeptidase (AEP) and caspase-3 will be examined in Chapters 4 and 5 respectively. AEP has been shown to degrade the drug L-Asparaginase, a drug essential for treatment of childhood acute lymphoblastic leukaemia (ALL) (347-349), whilst caspase-3 has been linked to tumour recurrence post-radiotherapy (350). These proteases have clear clinical relevance and their activities in plasma in terms of prognostic indicators will be investigated in this study.
4 Chapter 4: Assaying the Activity of the Protease, Asparaginyl Endopeptidase (AEP) using SRM MS

4.1 Introduction

This chapter describes the application of the method developed in “Chapter 3: The Development and Optimisation of a Multi-Step Sample Preparation Method to Assay Protease Activity in Biological Fluids by Selected Reaction Monitoring Mass Spectrometry” to the clinically relevant protease, asparaginyl endopeptidase (AEP). The performance of the new assay can be assessed by comparison with an existing fluorescence method validated by Holland et al. (351). This fluorescence assay has been validated on whole cell lysates (WCLs) of different cell lines expressing AEP. In addition, a cell line had been developed with a missense mutation at the active site of AEP. This allows the effect of mutations on the activity of AEP to be explored, which is not possible in an ELISA assay.

4.1.1 Asparaginyl Endopeptidase (AEP)/Legumain

AEP is a cysteine protease with high substrate specificity, cleaving the Carboxyl termini of asparagine amino acids (352, 353). Appel et al. showed that when treating children with childhood Acute Lymphoblastic Leukaemia (ALL), those with increased levels of AEP in plasma respond poorly to treatment with L-Asparaginase – one of the primary chemotherapeutic drugs used in the treatment of childhood ALL (354). This was supported when Patel et al. showed that AEP degrades L-Asparaginase (355). AEP has been shown to be over-expressed in high risk subsets of childhood ALL (355). If a child does not respond to treatment, their chances of survival is dramatically reduced (66). Of those who do not respond to the first treatment of chemotherapy, less than 1 in 3 will survive five years (66). Therefore, a need exists to predict how a child will respond to treatment with drugs such as L-asparaginase. Therefore, monitoring AEP activity is of paramount importance in order to predict a child’s response to L-asparaginase, as this drug has been shown to dramatically improve patient outcome, when it is effective (347-349). AEP has also been shown to be up-regulated in malignant ovarian cancer tumours and as such been suggested as a potential diagnostic or prognostic biomarker (356). Chen et al. show that AEP is not active at the pH of native plasma (Figure 4.1, reference (357)). As AEP becomes unstable and inactive above pH 6.5, and the pH of
plasma is 7.4, AEP will not be active in plasma. Therefore, AEP activity cannot be assayed in plasma without reducing the pH of the sample.

**Figure 4.1:** The effect of pH on (a) the stability and (b) the activity of AEP. These data show that AEP is not stable at the pH of native plasma (pH 7.4). Therefore, the pH of the sample needs to be lowered in order for AEP to be stable and active (taken from Chen et al., reference (357)).

As with all cysteine proteases, AEP requires a histidine residue to stabilise the intermediate species of AEP bound to a substrate. During proteolysis the thiol group of the cysteine at the active site becomes deprotonated, and a basic residue such as histidine is required to sequester this positive charge and stabilise this reaction intermediate. In order to cleave the peptide bond of the substrate, nucleophilic substitution takes place via an $S_N2$ process, where the now deprotonated cysteine attacks the carbonyl carbon of the substrate. This $S_N2$ reaction forms a tetrahedral intermediate at this carbonyl carbon, before the peptide bond becomes broken, forming a new amine terminus in the substrate. The histidine residue becomes deprotonated, releasing its positive charge and a thioester intermediate bonding the newly formed carboxy terminus of the substrate to the thiol of the cysteine active site is generated. This thioester bond is then hydrolysed, releasing the cleaved substrate and restoring the active site of the protease to its original form (Figure 4.2) (358).
Figure 4.2: The reaction mechanism of substrate proteolysis by a cysteine protease. This shows the histidine stabilising the reaction intermediate by sequestering positive charge during the hydrolysis of a peptide bond in the substrate (adapted from Schneck et. al., reference (358)).

Cell lines have been developed with a missense mutation (H150A) at the active site of AEP. This means the histidine at position 150 has been substituted for an alanine (Figure 4.3), which can’t accept the positive charge. It has been shown that AEP has a reduced activity in cell lines mutated in this way (359). This is because H150 interacts with C191 to form the catalytic domain of AEP.

Figure 4.3: Amino acid sequence of wild type and H150A mutated AEP. This shows the missense mutation at position 150 and the retention of C191 in the cell lines to be used in this project (adapted from the UniProt database, reference (360)).
By incorporating the H150A missense mutation into AEP, the activity of the enzyme is greatly reduced due to the instability of the enzyme-substrate intermediate, as, unlike histidine, alanine cannot sequester positive charge and stabilise this intermediate. Further evidence that H150 and C191 interaction forms the active site of AEP has been shown by creating a missense mutation at C191 as opposed to H150 activity is also reduced (359). Cell lines expressing AEP and H150A mutated AEP will be investigated in this assay in order to assess the effect this H150A mutation has on the amount of AEP SRM substrate cleaved and if the effect of this H150A mutation is seen in both SRM and fluorescence assays.

As described in Chapter 1, profiling biological fluids such as plasma or urine has gained considerable interest in recent years (361-365). This is because these fluids are readily accessible and are expected to provide an accurate reflection of the situation of the patient at any given time. As such, considerable effort has gone into profiling these biological fluids in order to find potential biomarkers or prognostic indicators (366, 367).

Work has been done in profiling AEP activity in cell lines which provide accurate models of childhood ALL (351, 355, 368), but at this point, very little work has been done on monitoring AEP activity in biological fluids such as plasma.

4.1.2 Aim of This Chapter

This aim of this work was to apply an MS-based assay developed in this project to monitor the activity of the clinically relevant protease, AEP in both plasma and whole cell lysate (WCL) samples. Furthermore, we compare this MS approach to the fluorescence-based approach to monitor AEP activity developed by Chen et. al. (368) and validated in house (351) when analysing these types of sample. In this comparison, several parameters will be examined, including correlation between datasets, limit of detection, linear range and variance.
4.2 Methods and Materials

4.2.1 Chemicals
In all experiments, the water used was HPLC grade (Rathburn, Walkerburn, UK), acetonitrile was HPLC grade (Fisher Scientific UK Limited, Leicestershire, UK) and formic acid was mass spectrometry grade (Fluka, Sigma-Aldrich, St. Louis, MO). TFA was mass spectrometry grade (Sigma-Aldrich, St. Louis, MO). AEP was recombinant human AEP, expressed in Murine myeloma cell line, NS0 derived, Ile18 Tyr433, with an N-terminal 7 His tag (R & D Systems, Abingdon, UK). The AEP specific target peptide (99.9% purity) was supplied by New England Peptides (Gardner, MA).

4.2.2 Mass Spectrometry Analysis
All MS samples, unless otherwise stated, were analysed on a 4000 Q-TRAP (Applied Biosystems, Foster City, CA) coupled to a liquid chromatography system consisting of a Famos autosampler, a Switchos trap system and an Ultimate Plus dual pump (LC Packings, Dionex, Sunnyvale, CA). The column used on the liquid chromatography system was an Acclaim pepmap 100, 15 cm, C18 column, pore size 100 Å. The software used for the 4000 Q-TRAP was Analyst 1.4.1 (Applied Biosystems, Foster City, CA). The software for the liquid chromatography system was Chromeleon 6.50 SP4 Build 1000 (LC Packings, Dionex, Sunnyvale, CA). In all MS analysis, the declustering potential was set at 60.0, the ion source voltage was 2300.0 V, the curtain gas was 20.0, the interface heater temperature was 155.0°C, Q1 was set to low resolution, and Q3 was set to unit resolution. Data were processed using the software package, MultiQuant 2.0 (AB Sciex, Toronto, Canada) using the Summation workflow.

4.2.3 Fluorescence Analysis
Fluorescence analysis was performed on a FLUOstar Omega fluorescence plate reader (BMG Labtech, Buckinghamshire, UK). To read the plate, the UV absorption wavelength was 360 nm, the emission wavelength was 460 nm. As the gain (an instrument parameter) is specific for each 96-well plate, this was determined immediately before analysing each sample. The plate reader was controlled by Omega 1.02 software (BMG) and data were processed using the MARS Omega data analysis software package, version 1.10 (BMG).
4.2.4 AEP SRM Target Substrate Design
Synthetic peptides were designed using the MEROPS database. From this, amino acids most reported at each site ranging from P4 to P4’ were chosen to be in the target peptide. Both the N- and C-termini of each peptide were end-capped with the sequence HPH. These HPH sequences were the D-isoform of each amino acid in order to prevent exopeptidase digestion of the target peptide.

4.2.5 AEP SRM Target Substrate SRM Generation
The Q3 values for all SRM transitions were generated using Protein Prospector. Collision energies were calculated using the following equation:
Collision energy = (m/z * a) + b
Where for a doubly charged peptide, a = 0.044 and b = 3.000
for a triply charged peptide, a = 0.040 and b = 2.000
“a” and “b” are instrument parameters and were optimised by Drs. John Griffiths, Richard Unwin and Duncan Smith in 2004 (unpublished work).
Q1 m/z values for peptides were predicted and used for peptides in both their doubly and triply charged forms.
SRM transitions were generated for the intact form of each synthetic peptide, as well as the N- and C-terminal peptides post-digestion. The dwell time for all transitions was 100.0 ms.
The SRM transitions to detect the intact form of the AEP target peptide, sequence hphFAANDVSKhph, are described in Table 4.1. The dwell time for all transitions was 100.0 ms.
Table 4.1: SRM transitions to detect hphFAANDVSKhph, hphFAAN and DVSKhph.

<table>
<thead>
<tr>
<th>Peptide Sequence</th>
<th>Charge State</th>
<th>Q1 (m/z)</th>
<th>Q3 (m/z)</th>
<th>Collision Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hphFAANDVSKhph</td>
<td>2+</td>
<td>797.4</td>
<td>1456.7</td>
<td>38.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1359.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1222.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1075.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1004.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>933.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>819.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>704.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>605.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>518.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>390.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>253.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>156.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>235.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>372.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>519.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>590.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>661.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>775.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>890.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>989.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1076.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1204.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1341.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1438.7</td>
<td></td>
</tr>
<tr>
<td>hphFAANDVSKhph</td>
<td>3+</td>
<td>531.9</td>
<td>1456.7</td>
<td>23.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1359.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1222.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1075.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1004.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>933.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>819.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>704.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>605.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>518.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>390.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>253.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>156.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>235.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>372.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>519.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>590.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>661.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hphFAAN</td>
<td>2+</td>
<td>397.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>235.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>375.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>372.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>519.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>590.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>656.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>559.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>422.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>275.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>204.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hphFAAN</td>
<td>3+</td>
<td>265.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>235.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>375.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>372.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>519.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>590.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>656.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>559.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>422.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>275.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>204.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hphFAAN</td>
<td>3+</td>
<td>265.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>235.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>375.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>372.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>519.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>590.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>656.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>559.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>422.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>275.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>204.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hphFAAN</td>
<td>2+</td>
<td>410.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>704.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>605.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>518.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>390.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>253.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>156.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>215.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>302.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>430.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>564.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>664.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hphFAAN</td>
<td>3+</td>
<td>273.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>704.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>605.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>518.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>390.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>253.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>156.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>215.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>302.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.2.6 Plasma Preparation for Pilot Experiments
Whole blood was drawn into Vacutainer tubes with anticoagulants, mixed pursuant to BD specifications, allowed to sit at room temperature for 30 minutes, and spun to plasma at 1,303 g for 20 minutes, in a refrigerated centrifuge (5°C). Plasma was then removed by aspiration, ensuring no red blood cell contamination. Blood was collected from stringently screened donors at FDA registered donor centres. The donors must have tested negative for rapid plasma regain (RPR), human immunodeficiency virus (HIV) 1 and 2, Hepatitis B Surface Antigen, hepatitis C virus (HCV) and HIV 1-Ag using approved food and drug administration (FDA) methodology. The donors must have had normal alanine transaminase (ALT) levels as this elevated ALT levels have been linked to a variety of diseases, including hepatitis. Qualifying donors were also subject to a physical examination and verbal screening (Source: Seralab, Haywards Heath, UK).
Plasma (100 µl) was then centrifuged at 16,000 g for 5 minutes at 4°C. The supernatant was removed, diluted four-fold with PBS and spin filtered using a 0.22 µm cut-off filter (Sigma) at 1,500 g for 30 seconds at 4°C. The resulting filtrates were pooled and aliquotted out into 25 µl samples.

4.2.7 SRM Background Levels in Plasma
Plasma was prepared as in section 3.7. A 25 µl aliquot was enriched by acetonitrile precipitation (50 µl) as described previously (Chapter 3) and re-suspended in buffer A (1000 µl). Samples of this (55 µl) were analysed by SRM scanning for the intact as well as the N- and C- termini cleavage products of each synthetic peptide.

4.2.8 Enzyme Activation
Recombinant human AEP (rAEP) (10 µg) (Abingdon, UK) was dissolved in sodium citrate solution (50mM, pH 4.5) containing dithiothreitol (5 mM) (100 µl) and incubated at 37°C for two hours (369).

4.2.9 AEP SRM Target Substrate Digestion
The AEP specific target synthetic peptide, sequence hphFAANDVSKphp (1 µg) was incubated with its activated rAEP (20 ng) at 37°C overnight. Following this, the
sample was dried to completion and re-suspended in buffer A to a final concentration of 1 pmol/µl. Samples were analysed by IDA analysis. To create the HPLC buffers A and B, Switchos buffer and needle wash solution, water was HPLC grade (Rathburn, Walkerburn, UK), acetonitrile was HPLC grade (Fisher Scientific UK Limited, Leicestershire, UK) and formic acid was mass spectrometry grade (Fluka, Sigma-Aldrich, St. Louis, MO).

Buffer A & Switchos buffer consisted of 97.9% water, 2% acetonitrile and 0.1% formic acid.

Buffer B consisted of 79.9% acetonitrile, 20% water and 0.1% formic acid.

The needle wash solution was 100% water.

4.2.10 Target Peptide Detection Investigation

4.2.10.1 LC-IDA of hphFAANDVSKhph
The AEP substrate for SRM assays, hphFAANDVSKhph (1 mg) was dissolved in PBS (1 ml). An aliquot (1 µl) of this was removed and diluted to a concentration of 1 pmol/µl using PBS as the solvent. This was analysed by mass spectrometry in order to find the charge state of the intact precursor and obtain MS/MS data for each peptide.

4.2.10.2 MALDI-ToF MS Analysis of hphFAANDVSKhph
The peptide, hphFAANDVSKhph (10 pmol, 1 µl) was mixed with the MALDI matrix α-hydroxy cinnamic acid dissolved in H₂O: MeCN: TFA (20:80:0.1) (1 µl at a concentration of 2.5 µg matrix/ml) and spotted on a MALDI plate in triplicate. These were analysed by MALDI-ToF to obtain the intact masses of each peptide. Samples were then analysed by MS/MS to obtain MS/MS sequence coverage.

4.2.10.3 Static Infusion analysis of hphFAANDVSKhphon a 4000 Q-TRAP
The AEP target substrate in SRM assays, hphFAANDVSKhph was diluted to a concentration of 10 pmol/µl using the solvent H₂O: MeCN: Formic Acid (50:50:0.1). Each was analysed individually to attempt to find the charge state of each peptide and the collision energy required to fragment it. Intact masses were theoretically calculated for each peptide ranging from a charge state of 1+ to 7+.
4.2.10.4 StageTip Binding
The peptide hphFAANDVSKhph was desalted using a C\textsubscript{18} StageTip using the method described in Chapter 2. The flow through and eluate from each desalting process were kept and dried to completion. Each eluate was re-suspended in H\textsubscript{2}O: MeCN: Formic Acid (98:2:0.1) until a theoretical concentration of 1 pmol/µl was reached for each peptide. These samples were then analysed on the 4000 Q-TRAP in an LC-MS run designed to generate both MS and MS/MS data.
This was repeated at pH 10.5, using 0.1% (v/v) ammonium hydroxide in place of TFA.
All samples were analysed by MALDI–ToF/ToF using a 5800 instrument (AB Sciex).

4.2.10.5 Offline HPLC Run + Fraction Collection
The peptides hphFAANDVSKhph and hphDEVDGAGShph (1 µg each) were pooled together into one sample. This was injected by Dr. Duncan Smith onto a U3000 offline HPLC system (Dionex) with the following setup.
Column: SunFire C\textsubscript{18} 3.5µm, 2.1 x 150 mm column (Waters, Manchester, UK).
Buffer A: 100% H\textsubscript{2}O with 0.1% TFA
Buffer B: 100% Acetonitrile with 0.1% TFA
The gradient conditions are described in Table 4.2.
Table 4.2: Offline LC-UV gradient to detect hphFAANDVSKhph

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Buffer B (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>0.1</td>
</tr>
<tr>
<td>24</td>
<td>40</td>
</tr>
<tr>
<td>26</td>
<td>80</td>
</tr>
<tr>
<td>29</td>
<td>80</td>
</tr>
<tr>
<td>30</td>
<td>0.1</td>
</tr>
<tr>
<td>60</td>
<td>0.1</td>
</tr>
</tbody>
</table>

The flow rate was 0.220 ml/min. No trap/guard column was used; the sample was directly loaded onto the analytical column.

4.2.10.6 Q-STAR Static Analysis of Offline HPLC Fractions

The fractions collected in section 3.11.9 were dried to completion, re-suspended in H2O: MeCN: Formic Acid (50:50:0.1) and diluted to a concentration of 10 pmol/µl. Each was analysed on a Q-STAR XL (Applied Biosystems).

Samples collected in section 4.2.10.5 were re-suspended in H2O with 0.1% TFA and diluted to 1 pmol/µl. Each was then injected onto the Q-TRAP set up to perform an IDA analysis. Each sample was monitored by LC-MS. The loading buffer was changed to 100% water acidified with 0.1% TFA (v/v). In addition, the trapping column used was changed to an Acclaim Pepmap 100, 100 µm x 2 cm NanoViper C18, 5 µm, 100Å trap column (LC Packings). The gradient over the analytical column was changed to the parameters as described in Table 4.3.

Table 4.3: Online LC-UV gradient to detect hphFAAN and DVSKhph.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Buffer A (%)</th>
<th>Buffer B (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>30</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>33</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>38</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>39</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>59</td>
<td>99</td>
<td>1</td>
</tr>
</tbody>
</table>

4.2.10.7 LC-IDA of hphFAANDVSKhph in PBS (nHPLC Changes)

Samples of each peptide were diluted with H2O with 0.1% TFA to 1 pmol/µl. Each was then injected onto the Q-TRAP. Each sample was monitored by LC-MS.
4.2.10.8 Cleaved SRM Substrate MS Parameter Optimisation

hphFAANDVSKhph was digested as described in section 4.2.9 and analysed using the SRM parameters described in Table 4.1. The six most intense SRM transitions used to detect the peptide DVSKhph. This sample was re-analysed and the collision energies for these six transitions were optimised to the nearest 0.5 eV (Table 4.4). The dwell time for all transitions was 100.0 ms.

Table 4.4: Optimised SRM transitions to detect DVSKhph

<table>
<thead>
<tr>
<th>Q1 (m/z)</th>
<th>Q3 (m/z)</th>
<th>Collision Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>410.2</td>
<td>605.3</td>
<td>21.5</td>
</tr>
<tr>
<td>410.2</td>
<td>253.1</td>
<td>26.0</td>
</tr>
<tr>
<td>410.2</td>
<td>215.1</td>
<td>23.0</td>
</tr>
<tr>
<td>410.2</td>
<td>156.1</td>
<td>26.0</td>
</tr>
<tr>
<td>410.2</td>
<td>390.2</td>
<td>22.5</td>
</tr>
<tr>
<td>410.2</td>
<td>664.3</td>
<td>21.0</td>
</tr>
</tbody>
</table>

These transitions were used to monitor the cleavage of the target AEP substrate in all future MS-based experiments.

In order to keep the dwell time for each optimisation to a minimum, this experiment was split into six different LC-MS runs, one to optimise each Q3 m/z. Following this, the collision energy leading to the most intense SRM response was used for all future testing.

4.2.11 Glufibrinopeptide Internal Standard SRM Generation and Optimisation

Glufibrinopeptide (GluFib) (Sigma) was sprayed from a static needle in order to obtain MS/MS data using the following parameters. The ion source voltage was 1600.0 V, the interface heater temperature was 100.0°C, the declustering potential was 60.0, and the curtain gas was 20.0. Q1 resolution was low and focussed on 786.0 with a collision energy of 37.6. MS/MS data was obtained by using Q3 as a linear ion trap and scanning across a mass range of 100.0 – 1500.0 Da. After manually sequencing the peptide in the resulting data file, the six most intense transitions were selected.

Following the selection of the six most intense SRM transitions, the collision energy for each transition was varied ± 5.0 eV in steps of 0.5 eV. The collision energies providing the most intense SRM response were used in all future experiments (Table 4.5). The dwell time for each SRM transition was 100.0 ms.
Table 4.5: Optimised SRM transitions to detect GluFib.

<table>
<thead>
<tr>
<th>Q1 (m/z)</th>
<th>Q3 (m/z)</th>
<th>Collision Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>786.0</td>
<td>333.2</td>
<td>41.8</td>
</tr>
<tr>
<td>786.0</td>
<td>480.3</td>
<td>39.2</td>
</tr>
<tr>
<td>786.0</td>
<td>684.4</td>
<td>39.4</td>
</tr>
<tr>
<td>786.0</td>
<td>813.4</td>
<td>37.6</td>
</tr>
<tr>
<td>786.0</td>
<td>1056.5</td>
<td>36.6</td>
</tr>
<tr>
<td>786.0</td>
<td>1285.6</td>
<td>36.0</td>
</tr>
</tbody>
</table>

4.2.12 GluFib Internal Standard SRM Background Levels in Enriched Plasma

An aliquot of the pooled plasma sample prepared in section 4.2.7 (25 µl) was enriched by acetonitrile precipitation as described in Chapter 3 and re-suspended in water acidified with 0.1% TFA (1,000 µl). An aliquot of this (55 µl) was analysed by LC-SRM using the LC conditions described in section 4.2.11.12 and the optimised GluFib SRM transitions generated in section 4.2.14 in order to check for any response from endogenous peptides present in plasma. Following this, a sample of GluFib (250 fmol) was analysed to ensure it could be detected using these optimised SRM transitions.

4.2.13 Cleaved SRM Substrate Linearity

First, the samples used in section 4.2.12 underwent a serial dilution across three orders of magnitude. The dilution of each sample relative to the most concentrated is shown in Table 4.6.

Table 4.6: The amounts of DVSKhph used in SRM linearity testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of Cleaved Peptide in Sample (fmol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>500</td>
</tr>
<tr>
<td>7</td>
<td>1,000</td>
</tr>
<tr>
<td>8</td>
<td>5,000</td>
</tr>
<tr>
<td>9</td>
<td>10,000</td>
</tr>
</tbody>
</table>
Each sample was analysed in triplicate and analysed by the optimised SRM method. The linearity of this SRM method was assessed by plotting the peak area of the most intense transition against relative sample concentration. This allowed the linearity of SRM transitions monitoring the cleaved and uncleaved forms of each target peptide to be assessed.

4.2.14 **Fluorescent Group Linearity in Plasma**
The fluorescent group NH$_4$Mec (400 pmol) was spiked into well 1 of a 96 well plate containing plasma prepared in section 4.2.7 (50 µl) and serially diluted across the plate (50% dilution each time). This was repeated with plasma diluted ten-fold with AEP assay buffer. Fluorescent response was assessed as in previous experiments and the resulting fluorescent intensity was plotted against NH$_4$Mec concentration.

4.2.15 **AEP Fluorescent Substrate Stability in Plasma.**
The fluorescent AEP target substrate, BazChem-Ala-Ala-Asn-Z-NH$_4$Mec (500 µmol) was added to plasma diluted ten-fold with AEP assay buffer. The fluorescent response was monitored at 0, 1, 2, 4, 8 and 24 hours.

4.2.16 **AEP Activity in Plasma (Fluorescence Analysis)**
Activated rAEP was added to plasma diluted five-fold with AEP assay buffer as described in order to reduce the pH of the sample to 5.8 (Table 4.7). This assay buffer consisted of 39.5 mM citric acid, 121 mM disodium hydrogen phosphate, 1 mM ethylene diamine tetraacetic acid (EDTA), 1 mM DTT and 0.01% CHAPS (v/v). This assay buffer was pH buffered to 5.8 by addition of dilute hydrochloric acid. Following this, the AEP specific substrate BazChem-Ala-Ala-Asn-Z-NH$_4$Mec (500 µmol) dissolved in assay buffer (50 µl) was added to each well. Samples were incubated at 37°C and the fluorescent response of each sample was monitored over a 24 hour period, as described in section 4.2.18.

Table 4.7: The amounts of rAEP added to plasma during proof of concept testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of rAEP in Sample (ng)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.250</td>
</tr>
<tr>
<td>2</td>
<td>3.125</td>
</tr>
<tr>
<td>3</td>
<td>1.563</td>
</tr>
<tr>
<td>4</td>
<td>0.781</td>
</tr>
<tr>
<td>5</td>
<td>0.000</td>
</tr>
</tbody>
</table>
4.2.17 AEP Activity in Plasma pH Buffered to pH 5.8 (SRM Analysis)

This experiment is an exact repeat of the experiment described in section 4.2.19: AEP Activity in Plasma pH Buffered to pH 5.8 (Fluorescence Analysis), except for the following changes. The AEP specific target peptide substrate, sequence hphFAANDVSKhph (2 µg) for MS analysis was added to each sample as opposed to the fluorescence based substrate. Individual samples were prepared for each time point. At each time point, samples were quenched with formic acid (5 µl) and instantly stored at -20°C until ready for analysis. At this point, samples were enriched by acetonitrile precipitation (2 volumes). Samples were then dried to completion and re-suspended in water acidified with 0.1% TFA (100 µl). An aliquot of this (55 µl) was analysed by SRM to detect the cleaved form of the peptide of interest following the addition of the GluFib internal standard (250 fmol). Samples were prepared in triplicate.

4.2.18 SRM Sample Dilution Testing

Aliquots of different amounts a sample containing the most amount of rAEP at the 24 hour time point (Table 4.8) were analysed by SRM. The peak areas for each sample were then processed using MultiQuant software. Each peak area was then examined against the linearity plot of the cleaved peptide shown in section 4.2.16. From this the amount of each sample which needs to be analysed in order to generate reliable quantitative data was found. This amount of sample was analysed from all samples.

Table 4.8: Sample dilution sample amounts.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of Sample Analysed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
</tr>
</tbody>
</table>

4.2.19 Is Effect of pH on the Stability of AEP Reversible?

rAEP (1 µg) was activated as described in 3.9: Enzyme Activation and split into three equal aliquots (0.33 µg each). One aliquot diluted with activity buffer (10 µl), and the other two diluted with PBS (10 µl) and stored at 37°C for 4 hours. Samples were stored at -20°C for 16 hours. Following this, samples were thawed. The rAEP sample originally diluted with activity buffer and one of the samples originally diluted with
PBS were diluted five-fold to a working concentration of 6.5 ng/µl, whilst the remaining sample was diluted to the same extent with PBS. At this point, samples were serially diluted across a 96 well plate as in previous fluorescence experiments, and each well had the fluorescent substrate (500 µmol) added to it. The activity of these samples was monitored by fluorescence over a 24 hour period.

4.2.20 AEP Limit of Detection Testing in Plasma
rAEP (1 µg) was activated as described above. This was then added to plasma diluted five-fold with assay buffer (100 µl). rAEP was then diluted across a 96 well plate and fluorescent substrate (500 mmol) was added to each well. The plate was analysed by fluorescence at the 0 hour and 1 hour time points. Once the limit of detection at the 1 hour time point for the fluorescence assay had been established, samples were prepared at this LOD ± 3 sample dilutions. These samples were analysed by SRM analysis at the 0 and 1 hour time points. From these data, the LOD of AEP activity in plasma for the SRM assay could be determined.

4.2.21 Whole Cell Lysate Sample Preparation
REH and REHM cells were provided by Dr. Seema Alexander. These pellets were lysed by three freeze/thaw cycles with N$_2$(l) in the presence of assay buffer (500 µl). The lysed cells were centrifuged at 10,000 rpm for 10 minutes and the supernatant saved. This underwent a Bradford assay to assess the total protein content in each sample according to the standard protocol supplied by the manufacturer (Sigma).

4.2.22 AEP ELISA Assay
This protocol was developed by Dr. Ashish Masurekar and performed by him (370). Each well of a white opaque Lumintrac 600 96 well microplate (Thermo) was coated with monoclonal mouse anti-human AEP antibody (100 µl) (R & D Systems) that was diluted in coating buffer at a concentration of 2 µg/ml. The plate was then sealed and kept at 4°C overnight. Unbound antibody was aspirated and each well was washed 3 times with washing buffer (200 µl). This aspiration/washing step was repeated each time before successive steps. Blocking was achieved by adding blocking buffer (400 µl per well) and incubating the plate for 2 hours at room temperature. Standard, controls and samples were all diluted in calibrator solution and loaded in duplicates in volumes of 100 µl per well. Using 2-fold serial dilutions in calibrator solution, 7
concentrations of standards were generated ranging from 50 ng/ml to 75 pg/ml. Samples consisted of cell lysates at 1:9 dilutions and plasma at 1:99 dilutions. Cell lysates of AEP+REH cell lines were used as positive control and volunteer plasma was used as a negative control. Next, plate was sealed and kept at 4°C overnight. Detection of bound AEP protein used sequential incubation at room temperature in dark conditions of 100 μl detection antibody at a concentration of 400ng/ml incubated for 2 hours, followed by 100 μl of Streptavidin/HRP solution for 1 hour and lastly 100 μl of chemiluminescent substrate for 5 minutes. The luminescence was measured in relative light units using a luminescent probe on a FLUOstar OMEGA plate reader.

The ELISA coating buffer consisted of sodium bicarbonate (0.795 g) and sodium carbonate (1.465 g) in water (500 ml) which had been filtered through at 0.2 μm filter and pH set to 9.6.

The ELISA washing buffer consisted of 500μl of Tween 20 (500 μl) in PBS (1 litre). This was filtered through a 0.2 μm filter before use.

The ELISA blocking buffer consisted of 1% Bovine Serum Albumin in a solvent of Tween 20 (27.7 μl) in water (50 ml) to 50ml. This was passed through a 0.2 μm filter before use (370).

4.2.23 Whole Cell Lysate AEP Activity Assay
100 μg total protein was diluted to a total volume of 100 μl with assay buffer. This was placed into well 1 of a 96 well plate and serially diluted across the plate (50% dilution each time) with assay buffer until well 11 of the plate. Once well 11 was diluted, 50 μl sample was removed, ensuring each well had 50 μl of analyte present. Well 12 contained 50 μl activity buffer and acted as a blank sample. This was performed in triplicate for each cell line. In addition, a NH₄Mec dilution series was also created on this plate to act as a standard curve. A rAEP dilution series was also created across one row of the plate to act as positive control samples. The solvent used for each dilution was activity buffer and the total volume in each of these wells was 100 μl. Following this, the AEP specific substrate BazChem-Ala-Ala-Asn-Z-NH₄Mec (500 μmol in a total volume of 50 μl) was added to each well in the plate, except for those containing the NH₄Mec fluorescent standard. Fluorescent response was measured at 0, 1, 2, 4, 8 and 24 hours.
### 4.3 Results

#### 4.3.1 AEP SRM Substrate Design

In order to design a substrate most likely to be digested by AEP, the known cleavage sites for all substrates of AEP recorded in the literature were examined via the MEROPS database (353) (Table 4.9).

**Table 4.9: Known AEP cleavages in the MEROPS database (10th July 2010).**

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P1'</th>
<th>P2'</th>
<th>P3'</th>
<th>P4'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Pro</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Ala</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Val</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Leu</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ile</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Met</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Phe</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Tyr</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Trp</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ser</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Thr</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Cys</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Asn</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Gin</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Asp</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Glu</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Lys</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Arg</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>His</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Based on the data in the MEROPS database, it is clear that an asparagine residue is essential at the P1 position. In addition, AEP appears to have a slight preference to digest peptides containing alanine at the P2 and P3 positions, as well as a phenylalanine residue at the P4 position. From these data, the following peptide sequence was used to design an AEP specific peptide target:

**hphFAANDVSKhph.**

AEP cleavage has been designed to take place between the N and D amino acid residues. The peptide has been capped at each end with an HPH amino acid sequence. These amino acids are of the non-natural isoform (D-isoform) and as such will inhibit exoprotease activity, helping target the site of digestion of this peptide to one particular area. Histidine residues were selected as they can sequester positive charge...
in electrospray ionisation (ESI). This should aid in detecting the cleaved forms of this target peptide in SRM analyses. BLAST searching the amino acid sequences HPHFAANDVSKHPH, HPHFAAN and DVSKHPH against the human proteome produced no hits. This indicates that none of these peptides are found endogenously in humans. In addition, the sequence FAAN is the amino acid sequence used in the substrate used to monitor AEP activity using fluorescence based assays.

4.3.2 AEP Substrate SRM Generation
Once a substrate for AEP had been designed, the SRM transitions for the intact peptide (sequence hphFAANDVSKhph) and of its predicted cleavage products (hphFAAN and DVSKhph) were generated \textit{in silico} using the MS Product tool from Protein Prospector (http://prospector.ucsk.edu). Collision energies were calculated using the parameters in Analyst software. From this, the SRM background responses from endogenous plasma peptides were assessed using a pooled plasma sample from ten healthy volunteers.

SRM analysis of a pooled plasma sample from ten healthy volunteers post-enrichment by acetonitrile precipitation showed some background response in plasma when detecting the peptide hphFAAN (Figure 4.4(a)). However, no significant response was seen when using the transitions to detect the peptide DVSKhph. This indicates that any SRM response generated for this peptide would be due to the cleavage of the AEP target substrate, unlike hphFAAN whose SRM transitions may be subject to contamination from endogenous plasma peptides.
Figure 4.4: SRM responses from endogenous plasma peptides when detecting the cleaved AEP SRM substrate. Extracted ion chromatograms (XICs) of a plasma sample enriched by precipitation with two volumes of acetonitrile and analysed by LC-SRM on a 4000 Q-TRAP using the predicted SRM transitions generated by Protein Prospector to monitor (a) The N-terminal form of the cleaved AEP MS substrate, sequence hphFAAN, and (b) The C-terminal form of the cleaved AEP MS substrate, sequence DVSKhph. Any response exceeding the dashed line is deemed significant.
The low response shown in the XIC of DVSKhph shows that there is no interference from endogenous peptides present in plasma at these SRM transitions. However, there are significant responses from several SRM transitions in when monitoring the peptide hphFAAN, including at 4 minutes, 9.5 minutes and 23.3 minutes onwards. Therefore, depending on the elution time of hphFAAN, monitoring this peptide in this assay may be problematic.

4.3.3 Target Peptide Digestion and Detection
Once the expected level of SRM contamination from endogenous plasma had been established, the intact form of the AEP substrate, sequence hphFAANDVSKhph was analysed by LC-MS in an unbiased IDA workflow (Figure 4.5). When MS/MS data from each spectrum was examined, no peptide could be found in any of these samples. This indicated that this peptide could not be detected under these LC-MS conditions.

![Figure 4.5: Total ion chromatogram when detecting hphFAANDVSKhph using LC-MS on a 4000 Q-TRAP. When examining each of these peaks, the peptide hphFAANDVSKhph could not be found.](image)
In order to ensure there was no problem with the peptide received from New England Peptides, a sample of it (1 pmol) was manually spotted on a MALDI plate and analysed by MS (Figure 4.6) and MS/MS (Figure 4.7) analysis using a 5800 MALDI–ToF/ToF instrument (AB Sciex).

![Figure 4.6: MALDI-ToF/ToF MS of hphFAANDVSKhph](image)

**Figure 4.6: MALDI-ToF/ToF MS of hphFAANDVSKhph.** The precursor mass (1593.94) corresponds with the predicted precursor mass of the peptide, along with MS data supplied by the manufacturer.

Whilst there were contaminant peaks in the spectrum (aside from peaks from the matrix of $m/z$ < 880, which are routinely seen, Figure 4.6), indicating some impurities in the the manufacturer. However, the peptide of interest ($m/z$ 1593.9492) is still highly abundant in the sample.
Manually sequencing this spectrum produces the amino acid sequence HPHFAANDVSKHPH. This indicates that the peptide is of the correct amino acid sequence.

MS/MS data from each peptide indicates an amino acid sequence that corresponds with the peptide ordered from New England Peptides. This indicates that the peptide was made correctly and there is no problem with the manufacture of each peptide. Therefore, an inability to detect the peptide previously (Figure 4.5) is not due to the peptide being incorrectly synthesised.

Following this, the peptide hphFAANDVSKhph was analysed by static infusion on the 4000 Q-TRAP. The peptide could not be found in any of these static infusion experiments. However, this may be because the peptide was dissolved in PBS, and it is known that the presence of salt inhibits ionisation in ESI (371). Therefore, samples of this peptide (1 µg) were desalted using a StageTip at either low pH (pH 4) or high pH (pH 10.5). In both of these cases, hphFAANDVSKhph could not be found in the eluate of the sample post-desalting. The flow-through and eluate of these samples post-desalting were then analysed by MALDI MS and MS/MS analysis as before. The results in this experiment (Figure 4.8) indicate that at low and high pH, the peptide is adsorbing to the C18 StageTip but poorly. This could be the cause of a lack of LC-MS data being generated as the peptides are either not adsorbing to the trap/guard column and being washed to waste along with all the salts in the sample, or not adsorbing to the analytical column. However, it is more likely that the former is the case as if the
peptide was not adsorbing to the analytical column, the peptide would expect to be seen, but very early.
Once it had been confirmed that the peptide was synthesised correctly, the LC-UV trace provided by New England Peptides was examined (Figure 4.9). This shows that the peptide hphFAANDVSKhph can bind to a C$_{18}$ column. As a result of this, the peptide hphFAANDVSKhph, along with the peptide hphDEVDGAGShph (used in Chapter 5) were injected onto an offline U3000 HPLC system (Dionex) and fractions collected every 30 seconds. This showed that both peptides can bind to the C$_{18}$ of this column (Figure 4.10). The fractions of interest were then dried to completion and re-suspended in H$_2$O: MeCN: Formic Acid (50:50:0.1) to be analysed in further experiments.

Figure 4.9: New England Peptides LC-UV trace of hphFAANDVSKhph. This shows that the peptide can bind to C$_{18}$ HPLC columns. Therefore, the HPLC conditions being used in this project must be changed.
Figure 4.10: Offline LC-UV trace of hphFAANDVSKph. This indicates that these peptides can bind to C18. The percentage of organic being injected is shown by the dotted purple line.

The peaks found in the UV trace of this offline LC-UV analysis indicate the presence of each peptide. These fractions were automatically collected by the autosampler to be analysed on a Q-STAR XL Q-ToF. After analysing these samples post-LC-UV analysis, fraction 27 was shown to contain the substrate to monitor AEP (Figure 4.11). This was confirmed by fragmenting the peptide and obtaining MS/MS sequence coverage for each peptide.
Figure 4.11: MS/MS data from a Q-STAR XL of hphFAANDVSKhph. Manually sequencing this peptide indicates that this fraction contains the peptide hphFAANDVSKhph.

By making several changes to the nHPLC system couple to the 4000 Q-TRAP, it was possible to get these peptides to stick to the columns used on this system (Figure 4.12). These changes involved completely removing all acetonitrile from the switchos buffer, transport buffer and the solvent used to dissolve/dilute these peptides. Additionally, the 0.1% formic acid normally used needed to be changed to 0.1% TFA in these solvents. This also required the starting percentage of buffer B (high organic) on the analytical column to be lowered from 8% to 1%. As a result of these changes, these three peptides can now be detected by mass spectrometry using an LC-MS setup. These peptides elute at far lower organic concentrations than described by New England Peptides. However, this is more an observation than a concern.
Figure 4.12: MS/MS of hphFAANDVSKhph from LC-IDA analysis on a 4000 Q-TRAP. This shows that the changes to the nHPLC system described above have enabled the intact forms of each peptide to be detected by mass spectrometry.

Following these LC changes, it was also possible to detect the peptide hphFAANDVSKhph when the peptide is dissolved in PBS without the need to go through an offline HPLC fractionation (Figure 4.13).

Figure 4.13: MS/MS of hphFAANDVSKhph dissolved in PBS from LC-IDA analysis on a 4000 Q-TRAP.

Following changes to the LC system, the cleaved forms of the AEP target peptide, sequence hphFAAN and DVSKhph could be found (Figure 4.14).
Figure (a) shows the mass spectrometry peak for hphFAAN at 10.97 minutes.

Figure (b) displays the SRM transitions for hphFAAN. The left side shows the doubly charged transitions, while the right side shows the triply charged transitions. The y-axis represents intensity in cps (counts per second), and the x-axis represents the mass-to-charge ratio (m/z) values.
Figure 4.14: LC-SRM analysis of hphFAAN and DVSKhph. Extracted ion chromatograms for the potential SRM responses for (a) the N-terminal cleavage product of the AEP SRM substrate, sequence hphFAAN, (b) SRM transitions monitoring hphFAAN at 10.97 minutes, (c) the C-terminal cleavage product of the
The SRM responses for hphFAAN DVSKhph show that both these peptides are exclusively in a 2+ charge state as a minimal response is observed for transitions detecting the peptide in a 3+ charge state when analysing these peptides on a 4000 Q-TRAP.

However, there is the possibility of deamidation at the asparagine residue of the N-terminal cleaved form of the peptide, sequence hphFAAN, which would result in a 1 Da mass shift. In addition, whilst there is no contamination from endogenous peptides at the elution time of the peptide hphFAAN (Figure 4.4), small shifts in the retention time of this peptide or contaminating peptides in plasma may lead to the presence of contaminants appearing in the SRM response for this target peptide. These factors in combination have the potential to lead to erroneous results. Therefore, it was decided that only the C-terminal form of the cleaved AEP synthetic substrate, sequence DVSKhph, would be screened for in future AEP activity SRM-based assays.

It was observed that the peptide DVSKhph could only be seen in its doubly charged (precursor m/z 410.2) state, and not its triply charged (precursor m/z 273.8) state (Figure 4.14(d)). To detect the C-terminal cleavage product of the AEP specific peptide, sequence DVSKhph, the six transitions with the most intense response were used. These were optimised further to give the optimum collision energy for each SRM response (Table 4.2).

### 4.3.4 GluFib Internal Standard Detection

Internal standards have multiple uses. Firstly, they can be used to aid in the absolute quantitation of an analyte of interest. Secondly, they can be used as a quality control to ensure that the instrument is operating correctly. This can be done by comparing the instrument’s response to the internal standard over time as different samples containing the same internal standard are analysed. Internal standards can also be used to compensate for different matrix effects from different samples. As GluFib is used to calibrate the linear ion trap on a weekly basis, its suitability as an internal standard was investigated using SRM transitions generated from calibration data. SRM analysis of a pooled plasma sample (ten healthy volunteers) enriched with acetonitrile precipitation showed no significant response (Figure 4.15 (a)). However, significant
SRM responses were seen when a sample of GluFib (250 fmol) was analysed by these transitions (Figure 4.15 (b) and (c)). This indicates that GluFib has the potential to act as an internal standard in this assay.
Figure 4.15: (a) SRM responses from endogenous peptides, (b) when detecting GluFib and (c) SRM detection of GluFib at 250 fmol.

The lack of response in Figure 4.15 (a) and the presence of a significant response in (b) and (c) indicate that GluFib can be used as an internal standard in this assay. A significant response is deemed to be an SRM response exceeding 100 counts per second.

4.3.5 Cleaved AEP Target Substrate Linearity

The linearity of the fluorescence and SRM assays were tested by analysing a dilution series of the cleaved form of their respective AEP substrates spanning four orders of magnitude. Fluorescence data indicates that the cleaved form of the AEP fluorescent substrate can be detected in a linear fashion between 1.5 and 200 pmol, with a limit of quantitation of 3.125 pmol. SRM data indicate that the C-terminal form of the cleaved AEP target peptide developed to monitor AEP activity in this assay has a limit of detection of 1 fmol, a limit of quantitation of 5 fmol and a linear range spanning three orders of magnitude (Figure 4.16). These data illustrate that the SRM assay has a lower limit of detection and quantitation than the fluorescence assay in terms of...
detecting the cleaved form of the AEP target substrate. The SRM assay also has a greater linear range than the fluorescence assay.
Figure 4.16: Cleaved AEP substrate linearity. (a) The fluorescent group NH$_4$Mec used in fluorescence analysis, and (b) C-terminal product of the AEP target peptide used in SRM analysis, sequence DVSKhph. This shows that both assays demonstrate a linear response, with the SRM assay providing a linear response over a larger concentration range ($r^2$ values of 0.9914 and 0.9981 for the fluorescence and SRM assays respectively). Error bars represent the standard deviation of each sample population ($n = 3$).
In the fluorescence assay, a linear response is seen over two orders of magnitude with a limit of detection of 1.563 pmol, a lower limit of quantitation of 3.125 pmol, and an upper limit of quantitation of 200 pmol, with an average variance of 2.8% across the linear range. In the SRM assay, a linear response is seen across four orders of magnitude, with a limit of detection of 1 fmol, a lower limit of quantitation of 5 fmol and an upper limit of quantitation of 5 pmol with an average variance of 28.0%.

When SRM linearity data (Figure 4.16(b)) is normalised against a GluFib internal standard present in every sample at the same amount, the ratio of DVSKhph SRM response and GluFib SRM response (Figure 4.17) produce a linear trend similar to that shown in data not normalised against GluFib (Figure 4.16(b)). The variance between data sets is comparable, with the non-normalised data having an average variance across the linear range of 30.9% and the normalised data set having an average variance of 27.7% across the same concentration range.

![Figure 4.17](image-url)  
**Figure 4.17:** DVSKhph SRM linearity data normalised against the GluFib internal standard. *This shows a linear response ($r^2 = 0.9948$). Similar data were found using non-normalised DVSKhph SRM data (Figure 4.16(b)). Error bars represent the standard deviation of each sample population normalised against the GluFib SRM response in each replicate ($n = 3$).*
4.3.6 AEP Fluorescence Substrate Stability in Plasma.
To ensure that the fluorescence assay is specific and stable over an extended time course the substrate BazChem-Ala-Ala-Asn-Z-NH$_4$Mec was incubated in plasma over a 24 hour period. No increase in fluorescence was observed over this period (Figure 4.18).

![Fluorescent substrate stability in plasma over a 24 hour period.](image)

The lack of an increase in fluorescent response indicates that no endogenous proteases are breaking down the AEP target fluorescent substrate and releasing the fluorescent group. This indicates that the substrate is only cleaved by AEP, highlighting the selectivity of this assay.

The substrate BazChem-Ala-Ala-Asn-Z-NH$_4$Mec is stable in both neat and ten-fold diluted plasma over a 24 hour period. This indicates that no naturally occurring enzyme present in plasma cleaves this substrate as no increase in fluorescent response (and therefore increase in cleaved substrate levels) was shown. This indicates that this fluorescence assay is selectively monitoring only AEP activity in the plasma sample.

4.3.7 AEP Activity in Plasma
The detection of AEP activity in plasma was tested by the addition of differing levels of AEP and the monitoring of fluorescence or SRM response. As AEP has been shown to be inactive at the pH of native plasma (Figure 4.4), the activity of the protease was monitored at pH 5.8 (Figure 4.19) Both fluorescence and SRM showed
AEP activity at this pH higher than observed at the natural pH of plasma. Therefore, plasma must be pH buffered to 5.8 to assess the activity of AEP in this sample type.
Figure 4.19: AEP activity in plasma at pH 5.8. (a) Fluorescence and (b) SRM responses showing the proteolysis of the AEP substrate (1 µg) in plasma spiked with various amounts of AEP at pH 5.8. This shows a clear response, showing that AEP is active in plasma when the pH of the sample is reduced to pH 5.8. Error bars in both graphs represent the standard deviation of each sample population (n = 3).
Whilst there is increased variance in the SRM assay, the profiles of these two graphs are similar and show good correlation. This indicates that the data produced by the SRM assay are valid as they compare well to the data generated by the fully validated fluorescence-based method.

Statistical analysis of these SRM data (single factor ANOVA testing) showed a significant change in the production of cleaved substrate (p < 0.05) at all time points (Table 4.10).

Table 4.10: Single factor ANOVA analysis of the data produced in Figure 4.19 (b) at each time point.

<table>
<thead>
<tr>
<th>Timepoint (hrs)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.23x10^{-8}</td>
</tr>
<tr>
<td>2</td>
<td>3.39x10^{-4}</td>
</tr>
<tr>
<td>4</td>
<td>1.33x10^{-4}</td>
</tr>
<tr>
<td>8</td>
<td>9.61x10^{-5}</td>
</tr>
<tr>
<td>24</td>
<td>6.03x10^{-6}</td>
</tr>
</tbody>
</table>

When assessing the ability of each method to produce rate of reaction data and the ability to quantify AEP activity in plasma in a linear fashion, the SRM data is capable of quantifying AEP activity in plasma linearly for up to eight hours (Figure 4.20(b)). This outperforms the fluorescence assay as this is only able to produce a linear rate of reaction up to one hour (Figure 4.20(a)). After one hour, the fluorescence detector does not produce a linear response. This is because the amount of cleaved AEP target substrate is sufficient to saturate the detector. As such, AEP activity cannot be quantified beyond the 1 hour time point.

SRM results correlate well with fluorescence data, showing that the greater the amount of active rAEP in the plasma samples at pH 5.8, the greater the SRM response. In addition, the data produced from samples with no AEP present show that there is no increase in SRM response for these transitions. This indicates that the substrate is stable in this environment and is not being broken down by other proteases to give an SRM response. Therefore, the increase in SRM/fluorescence response is due to the AEP-specific substrate being broken down by its target protease, AEP.

At this point, it would be worthwhile to calculate a value for $K_M$ for this substrate and compare it to the known $K_M$ value for the fluorescence substrate. However, to
calculate a value for $K_M$, an experiment involving keeping rAEP concentration constant and varying the amount of substrate present needs to be performed. This has not been done in this project as this study involves the development of a method which could be used in a clinical assay, where the amount of substrate added to the sample is kept constant, and the amount of AEP in the sample would vary depending on a patient’s disease progression. However, a value for the rate of cleaved peptide production ($k_2$) in the SRM assay can be calculated.

Based on equation 1.13, The maximum rate of reaction can be described as $V_{max} = k_2[E]_t$. Additionally, $[E]_t$ is defined as the total amount of free enzyme, plus the total amount of enzyme bound to its substrate (equation 1.3). As the amount of AEP substrate added to the sample is approximately orders of magnitude greater than the amount of rAEP present in the sample, it can be assumed that all rAEP is being used to digest this substrate. Therefore, $[E]_t = [ES] = [E]$.

As the amount of rAEP added to the sample is known (Table 4.7) and the rate of reaction ($V_{max}$) has been determined experimentally (Figure 4.20(b)), it is possible to predict the value of $k_2$ by re-arranging equation 1.13 (Equation 4.1).

$$k_2 = \frac{V_{max}}{[E]} \quad \text{(Equation 4.1)}$$

Based on the data generated in Figure 4.20(b) and shown in Table 4.7 the value for $k_2$ at each time point can calculated (Table 4.11).

**Table 4.11: Calculated $k_2$ values for each time point when determining rAEP activity in plasma.**

<table>
<thead>
<tr>
<th>Time Point (hrs)</th>
<th>$k_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$3.20 \times 10^3$</td>
</tr>
<tr>
<td>2</td>
<td>$2.87 \times 10^3$</td>
</tr>
<tr>
<td>4</td>
<td>$2.73 \times 10^3$</td>
</tr>
<tr>
<td>8</td>
<td>$2.20 \times 10^3$</td>
</tr>
<tr>
<td>24</td>
<td>$6.42 \times 10^3$</td>
</tr>
</tbody>
</table>

It is possible that the decrease in the value of $k_2$ over time is due to further proteolysis of the cleaved AEP SRM substrate. In this assay, the peptide DVSKhph is monitored by SRM. However, when the intact substrate is digested by rAEP to product DVSKhph, only the three C-terminal amino acids are the D-isoform and not subject to proteolytic degradation. The remaining amino acid sequence DVSK is subject to proteolysis by proteases naturally occurring in the sample. If the peptide DVSKhph is cleaved by these proteases, the amount of this peptide the MS will detect will be
reduced. This in turn will reduce the value for $k_2$. This is a potential source of variation which cannot be avoided. However, if all samples are incubated for the same amount of time (e.g. one hour), this effect will be comparable across all samples and an assay can still be developed to monitor AEP activity in plasma using SRM MS.
Figure 4.20: Rate of reaction of AEP in plasma at pH 5.8 when being analysed by (a) fluorescence and (b) SRM assays generated from the data produced in Figure 4.16. Error bars in both graphs represent the standard deviation of each sample population (n = 3).
Rate of reaction analysis shows that whilst the fluorescence assay has significantly less variance than the SRM method, the fluorescence assay can detect AEP activity in plasma in a linear fashion at up to one hour of incubation with its target substrate; whereas the SRM assay can quantify AEP activity in plasma for up to eight hours if necessary (Figure 4.20).

When using GluFib as an internal standard in linearity samples containing no complex background, a similar trend is seen between these normalised data to the same data analysed using the peak areas of DVSKhph only. However, when the AEP activity in plasma SRM data (Figure 4.20(b)) is normalised against a GluFib internal standard present in every sample at the same amount, the ratio of DVSKhph SRM response and GluFib SRM response (Figure 4.21) produce a linear trend similar to that shown in data not normalised against GluFib (Figure 4.20(b)), although a discrepancy at 4 hours is seen. In this case, the SRM response for plasma containing 3.125 ng rAEP is greater than plasma containing 6.25 ng rAEP. In addition, the variance for this normalised data is significantly larger (74.7%) when compared to non-normalised data (42.4%). It is suspected that the reason behind this increased variance in data is due to the large variance in internal standard response (64.9%). Examples of this are shown in Table 4.1. This may also be the source of error in the data at 4 hours. As a result of this, GluFib is not suitable as an internal standard to normalise data against when assaying AEP activity in plasma samples. As such, GluFib can be used as a quality control (QC) standard in plasma samples, but not as an internal standard which AEP activity in plasma samples can be normalised against.
Figure 4.21: AEP activity in plasma at pH 5.8 normalised against GluFib. These data show a similar trend to non-normalised DVSKhph SRM data (Figure 4.16(b)), but with significantly higher variance. Error bars represent the standard deviation of each sample population (n = 3).

Table 4.12: GluFib SRM responses when monitoring AEP activity in plasma.

<table>
<thead>
<tr>
<th>Amount of rAEP in Sample (ng)</th>
<th>Time Point (hrs)</th>
<th>DVSKhph SRM Response</th>
<th>GluFib SRM Response</th>
<th>DVSKhph:GluFib Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.125</td>
<td>24</td>
<td>9.25x10^4</td>
<td>4.21x10^4</td>
<td>2.20x10^1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.29x10^4</td>
<td>2.59x10^4</td>
<td>3.21x10^1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.63x10^4</td>
<td>1.19x10^4</td>
<td>5.55x10^0</td>
</tr>
<tr>
<td>1.5625</td>
<td>24</td>
<td>1.79x10^4</td>
<td>1.08x10^4</td>
<td>1.65x10^1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.30x10^4</td>
<td>1.56x10^5</td>
<td>2.11x10^1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.91x10^4</td>
<td>5.47x10^3</td>
<td>7.15x10^-2</td>
</tr>
</tbody>
</table>

In order to assess if this assay has potential to monitor AEP activity in plasma when it has been incubated in plasma for a long period of time, as would be the case when analysing patient samples, AEP was incubated at pH 7.4. At fixed periods, the pH of samples incubated in this way were reduced to pH 5.8 in order to assess whether AEP can be restored by reducing the pH of the same, or whether AEP had been irreversibly degraded as a result of incubation at this higher pH. When testing if AEP activity can
be recovered in plasma by reducing the pH of the sample, fluorescence data shows that there is potential to measure AEP activity in plasma (Figure 4.22). However, the incubation time of this assay must be greatly increased. This is shown by the difference in activity between the rAEP positive control, and all other samples in this assay. There also appears to be no significant correlation between the amounts of time samples have been incubated in PBS, but it does appear that the addition of activity buffer immediately helps restore AEP activity, albeit to a lower level than the positive control. As a result, there is a potential to monitor AEP activity in plasma whose pH has been reduced to pH 5.8. However, the time samples need to be incubated to get a significant response when performing this assay needs to be greatly increased. This is shown by the difference in fluorescent response between all samples incubated at pH 7.4 originally, and the AEP positive control. The best example of this is the fluorescent response for the samples incubated at pH 7.4 eight hours into this assay is of a similar level to the AEP positive control at 15 minutes.
Figure 4.22: Restoring AEP activity in plasma by altering pH. *Error bars represent the standard deviation of each sample population (n = 3, except for rAEP control which was n = 1).*

This shows that AEP does have some activity restored after incubation with activity buffer, but it does not fully recover to the levels of the rAEP positive control. Therefore, there is potential to monitoring AEP activity in plasma once the pH of the sample has been reduced to 5.8 by dilution with assay buffer. However, the incubation time for this assay must be greatly increased.
4.3.8 AEP Activity in Whole Cell Lysates

ELISA assays are routinely used to quantify the total amount of a protein of interest. However, these assays do not provide any information about the activity of the protease they are monitoring. Additionally, if a protein is mutated, there is the potential for the antibody used in the ELISA assay not to recognise the mutated form of this protein. The ELISA used to quantify the total amount of AEP in samples provided a linear response from 0 – 50 ng rAEP present in 1 ml sample (Figure 4.23), as performed by Dr. Ashish Masurekar.

![Figure 4.23: AEP ELISA assay linearity curve. This shows a linear response from 0 – 50 ng/ml AEP ($r^2 = 0.9970$). This allows samples containing AEP in this range to be accurately quantified (performed by Dr. Ashish Masurekar). Error bars represent the standard deviation of each sample population (n = 3).](image)
When testing the REH and REHM WCL samples, AEP in the sample was found to be at a concentration of 0.98 and 0.67 ng/µg total AEP respectively (ELISA assay performed by Dr. Ashish Masurekar). If this H150A mutation at the active site of AEP had no effect on AEP activity, it would be suspected that the activity of the REHM WCL would be approximately 2/3 that of the REH WCL, when samples are normalised to total protein content using a Bradford assay.

WCL analysis shows that missense mutations at the active site of AEP have a profound effect on the activity of AEP (Figure 4.24) with little activity measurable in REHM. This provides more biological insight than simply quantifying the amount of AEP present in a sample, in the form of an ELISA assay (Figure 4.23).
Figure 4.24: AEP activity in WCL samples. (a) Fluorescence and (b) SRM-based data showing the activity of AEP in each cell line over a 24 hour period. REH WCL samples show an increase in AEP activity compared to the REHM samples containing a H150A missense mutation in the active site of AEP, which shows no significant fluorescent response. Error bars represent the standard deviation of each sample population (n = 3).
Figure 4.24 highlights the importance of assessing the activity of AEP, rather than its total quantity. This is because the activity assay offers a more accurate representation of what is occurring in the biological system of interest, rather than simply quantifying the amount of AEP present in the sample with no information about its activity in biological systems.

Statistical analysis of these SRM data (single factor ANOVA testing) showed no significant change in the production of cleaved substrate ($p < 0.05$) at any time point except 2 hours when analysing REH WCL samples (Table 4.13).

**Table 4.13: Single factor ANOVA analysis of the data produced in Figure 4.24 (b) at each time point.**

<table>
<thead>
<tr>
<th>Timepoint (hrs)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.02x10^{-2}</td>
</tr>
<tr>
<td>2</td>
<td>0.41</td>
</tr>
<tr>
<td>4</td>
<td>2.57x10^{-3}</td>
</tr>
<tr>
<td>8</td>
<td>5.59x10^{-4}</td>
</tr>
<tr>
<td>24</td>
<td>2.84x10^{-3}</td>
</tr>
</tbody>
</table>

It is possible that the outlying p value at 2 hours is due to an error in sample preparation. In addition, this experiment was performed using three replicates for each sample at each time point. This experiment needs to be repeated using higher numbers of replicates in order to perform more robust statistical analysis.

It appears that the SRM assay provides a qualitative measure of AEP activity in plasma, as when a missense mutation is present at the active site of AEP, the amount of cleaved SRM substrate produced is significantly reduced (Figure 4.24(a)). However, when examining the rate of proteolysis of the SRM substrate, a linear response across the range of concentrations is not seen at all time points (Figure 4.24(b)). It appears that a linear trend over the first 2 hours is seen for up to 25 µg of REH WCL, but this trend is lost when more than 25 µg of REH WCL is assayed. After 2 hours the linear range of this assay is reduced (Figure 4.25(b)). However, the fluorescence assay appears to provide a linear response when assaying REH activity at similar rates until 8 hours (Figure 4.25(a)). From this point onwards, the rate of reaction is reduced. It is possible that proteolysis of AEP takes place, reducing the amount of cleaved substrate produced and reducing the rate of AEP fluorescent substrate proteolysis.
Figure 4.25: Rate of reaction of AEP in WCL samples for the (a) fluorescence and (b) SRM assays. Error bars represent the standard deviation of each sample population ($n = 3$).
Whilst the fluorescence assay provides a similar linear response across the entire concentration range for all time points up until 8 hours, the SRM assay’s substrate appears to be further degraded at all time points past 25 ng of REH WCL. At later time points, this is reduced to 12.5 ng.
4.4 Discussion

4.4.1 AEP Activity in Plasma Samples

This chapter shows that the activity of the clinically relevant protease, AEP can be quantified in an assay ranging from one to eight hours (Figure 4.19), with AEP activity being detected at picogram per millilitre levels using the SRM assay developed in this project. AEP activity assays are commonly performed using fluorescence based approaches. However, we show in this chapter that the SRM assay developed in this project outperforms the corresponding fluorescence assay in terms of AEP activity limit of detection, dynamic range and cleaved substrate limit of detection and quantitation with similar profiles being seen in the SRM and fluorescence assays, indicating that this data is valid.

AEP is not active at the pH of native plasma. However, by reducing the pH of the sample to 5.8, AEP activity can be monitored using the SRM assay in a linear fashion (Figure 4.19).

This chapter also shows that the fluorescent group BazChem-Ala-Ala-Asn-Z-NH₄Mec produces a linear response in plasma. This offers the potential of monitoring AEP activity in plasma using fluorescence based approaches. The NH₄Mec group has potential to be used to monitor the activities of other proteases as the cleaved form of a fluorescent substrate. However, much work involving enzyme specificity and substrate needs to be done in order to create this method. Results generated in this chapter indicate that AEP is not active in plasma. It has been shown that this is due to a pH effect. This is based on the results generated in sections 5.3.5 and 5.3.6 and work performed by Chen et al., described above.

As discussed in Chapter 3, the number of SRM transitions used in this assay exceeds the requirements suggested by Sherman et al. (262, 332). This, along with the known retention time of the peptide, allows the cleaved peptide substrate to be confidently identified, despite the highly complex background this peptide is in.

However in order to make this method fully amenable to a clinical setting, further testing needs to be performed. This includes testing this assay on a fully optimised non-nanoflow HPLC system and drastically reducing the sample analysis time in order to produce valid, robust and reliable results as quickly as possible. In addition, other criteria such as inter- and intra-lab reproducibility need to be assessed before monitoring AEP activity in five-fold diluted plasma as a prognostic indicator can become a routine clinical test.

It is suspected that the reason AEP activity in plasma cannot be assayed beyond 8 hours using the SRM assay is due to further proteolysis of the cleaved AEP target substrate. Once this...
substrate has been cleaved by AEP, producing the peptide DVSKhph, this peptide is only
capped at one end by D-amino acids. Therefore, other proteases can act upon this newly
cleaved substrate. As this SRM assay is only detecting the peptide DVSKhph, this further
proteolysis of DVSKhph lowers the amount of this peptide present in the sample, which can
account for the lower response seen after eight hours in this SRM assay. However, this
requires further investigation. Statistical analysis (single factor ANOVA testing) showed a
significant increase in cleaved substrate production with respect to increased levels of rAEP
(p < 0.05). However, due to limitations in instrument time, this experiment was performed in
triplicate at all time points and with all enzyme: substrate ratios. Repeating this experiment
with increased numbers of replicates would enable more powerful and robust statistical
testing to be performed using ANOVA, which would in turn provide a more accurate
understanding into the significance of the change in cleaved substrate production in relation
to the amount of rAEP in plasma samples.

4.4.2 Fluorescence/MS Comparison
The presence of a fluorescence-based method (351) to monitor AEP activity in whole cell
lysate samples offers an opportunity to compare this fluorescence method to the SRM-based
method being developed in this project. This allows several important comparisons to be
made between the two methods. A comparison of the results between the two methods can be
made in terms of how well the results correlate to each other. As the two methods involve
two different detectors, parameters such as absolute detector response (i.e. fluorescent
response, SRM response) cannot be compared. However, things such as shape of the profiles
of the two methods can be compared and assessed. If the SRM-based method correlates well
with the previously validated fluorescence-based method, this indicates that this SRM method
provides valid results. This was shown to be the case in both plasma (Figure 4.19) and WCL
samples (Figure 4.25), although the SRM WCL data did not provide a linear response for
REH data and can only be used for a qualitative assay. Additionally, other parameters can be
compared between the two methods such as dynamic range, variation and, limits of detection
and quantitation. From these data, it can be assessed whether this SRM-based method has
potential for use in a clinical assay to monitor AEP activity in terms of a potential prognostic
indicator. Before work in this chapter took place, the fluorescence method has only been used
to monitor AEP activity in whole cell lysate samples. No work has been done on plasma at
this point. However, comparisons between whole cell lysate samples when monitoring the
same samples by SRM and fluorescence can take place as this method has been validated in house (351).

This chapter shows that the SRM assay has a lower limit of detection for the proteolytic product than the fluorescence assay (1 fmol as opposed to 1.5 pmol in the fluorescence assay), with a greater linear range (3 orders of magnitude for the SRM method as opposed to 2 orders of magnitude for the fluorescence assay). In addition, the SRM assay can detect AEP activity at significantly lower levels of AEP in plasma than the fluorescence method. This overcomes the primary problem with the fluorescence method.

However, unlike the SRM assay, the fluorescence assay does not require additional sample handling steps such as peptide enrichment. This additional sample handling step offers an explanation into the greater variance in the SRM assay when compared to the fluorescence assay. Additionally, because fluorescence analysis is a non-destructive technique, the same sample can be analysed multiple times. This is not the case with the SRM assay as separate samples need to be prepared at every time point (Table 4.1). However, after pilot experiments have been performed, the analyst knows what time points to take samples. This will in turn reduce the cost per assay to those described in Table 4.1.

Table 4.1: A comparison of the fluorescence and SRM based assays used to quantify AEP activity in plasma.

<table>
<thead>
<tr>
<th></th>
<th>Fluorescence</th>
<th>Mass Spectrometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Cumulative Variance (%)*</td>
<td>2.8</td>
<td>28.0</td>
</tr>
<tr>
<td>Linear Range (Orders of Magnitude)</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>AEP Limit of Detection in Plasma (ng/ml)</td>
<td>1</td>
<td>&lt; 0.13</td>
</tr>
<tr>
<td>Cleaved AEP Substrate Lower Limit of Detection (pmol)</td>
<td>1.563</td>
<td>0.001</td>
</tr>
<tr>
<td>Cleaved AEP Substrate Lower Limit of Quantitation (pmol)</td>
<td>3.125</td>
<td>0.005</td>
</tr>
<tr>
<td>Cleaved AEP Substrate Upper Limit of Detection (pmol)</td>
<td>200</td>
<td>10</td>
</tr>
<tr>
<td>Cleaved AEP Substrate Upper Limit of Quantitation (pmol)</td>
<td>200</td>
<td>5</td>
</tr>
<tr>
<td>Analyse Sample Multiple Times?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cost per Sample ($)</td>
<td>1.50</td>
<td>1.75</td>
</tr>
<tr>
<td>Cost per Assay ($)†</td>
<td>4.50</td>
<td>10.50</td>
</tr>
<tr>
<td>Sample Processing Required?</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

* Average variance across linear range of assay
† Two time points, triplicate analysis.
Table 4.11 shows that the fluorescence method has some considerable advantages over the MS method in areas such as average cumulative variance and the time taken to analyse samples. However, the MS method also easily outperforms the fluorescence method in other areas, such as limits of detection and quantitation, as well as the dynamic range of the assay. These qualities make the MS method an attractive choice to assay AEP activity in plasma, especially when low responses are predicted.

In the fluorescence and SRM assays, the amount of substrate added to each sample was approximately ten times greater than the $K_M$ value for AEP proteolysis of the fluorescent substrate (372). This ensures that AEP is saturated by substrate, and the rate of substrate proteolysis is at $V_{MAX}$. As the SRM substrate has been specifically designed for this assay, there is no $K_M$ data available for AEP proteolysis of this peptide. However, as the amino acids at the P1 – P4 positions are identical to that of the fluorescent substrate, the rate of proteolysis of this SRM substrate is suspected to be similar to that of the fluorescent substrate.

4.4.3 AEP Activity in Whole Cell Lysate Samples

The results produced in this project in the whole cell lysate project answer many important questions in this project. WCL sample testing indicates that if a missense mutation exists in the active site of AEP, the activity of the protease is greatly reduced. Both the fluorescence and SRM assay confirmed this with good correlation. This shows the importance of measuring AEP activity as opposed to quantifying the amount of AEP in an ELISA assay. ELISA data shows the REHM cell line having 67% of the amount of AEP present than REH cell line. However, the activity of the REHM is far less than 67% of the activity of the REH cell line, as the REHM cell line shows no significant fluorescent response. This offers greater insight into the activity of AEP in the biological system of concern, information which ELISA assays do not provide. In addition, as ELISA assays are antibody based, mutations in the protein of interest (in this case, AEP) may not necessarily be detected by the antibody used in the ELISA assay. Whilst the SRM data for REH whole cell lysate samples do not produce a linear response, there is a clear difference in the amount of AEP target digested when comparing these REH whole cell lysates to the REHM whole cell lysates. Therefore, a qualitative assay has been developed to determine the presence of active AEP in WCL samples. Although every possible step has been taken to prevent proteolysis of the target AEP peptide by other proteases, very little can be done to inhibit the activity of other proteases on the newly cleaved AEP substrate. Therefore, we suspect that further proteolysis
of the peptide DVSKhph takes place after the original substrate has been digested by AEP. Because the SRM assay is only detecting DVSKhph, and not any of its products following additional proteolysis, this can explain the non-linear response for this assay in WCL samples. Statistical analysis (single factor ANOVA testing) showed a significant increase in cleaved substrate production with respect to increased levels of rAEP (p < 0.05) at all time points apart from 2 hours. It is possible that the lack of significance at this time point was from an error in sample preparation. However, due to limitations in instrument time, this experiment was performed in triplicate at all time points and with all enzyme: substrate ratios. Repeating this experiment with increased numbers of replicates would enable more powerful and robust statistical testing to be performed using ANOVA, which would in turn provide a more accurate understanding into the significance of the change in cleaved substrate production in relation to the amount of rAEP in WCL samples.

Work described in this chapter indicates the potential to use this MS based method to qualitatively monitor protease activity in samples other than plasma, such as WCL samples.

4.4.4 Changes to the HPLC Setup for the SRM Assay

The peptide to detect AEP activity in the SRM assay was designed in a very specific way. The four amino acids either side of the desired cleavage site could not be changed, as this is the consensus sequence that AEP recognises (353). The three amino acids capping each end of the peptide, however, could be any amino acid at all. In order to assist the detection of the target peptide, two histidine residues were used at this end cap, as these can sequester positive charge, which will help the peptide become ionised. If it is assumed that acidic amino acids can hold a 1+ charge, basic amino acids lose a 1+ charge, and neutral amino acids have no ability to either gain or lose charge when ionised, then each form of the cleaved AEP target (FAAN and DVSK) should, in theory have charges of 0 and 0 respectively (373). Therefore, in order to aid ionisation, amino acids which can hold positive charge needed to be used in the end caps of the AEP target peptide for this SRM assay. However, adding these hph end caps increases the hydrophilicity of these peptides considerably (374).

When trying to detect the intact form of hphFAANDVSKhph at the outset of this work, no data could be generated. It was suspected that this could be due to one of three things; 1) A problem with the sample itself, 2) A problem with the peptides becoming ionized at the source of the mass spectrometer, or 3) A problem with the peptides at the HPLC level, either them not sticking to the C18 column or adsorbing permanently. However, due to no pressure increases in the HPLC-MS runs immediately following the analysis of these peptides, the
peptides sticking to the column permanently was ruled out. If these peptides were adsorbing irreversibly to the surface of the column, the pressure inside the column in following runs would increase. This was not the case. Therefore, this was ruled out. MALDI analysis at the MS level proved that each peptide was of present in a high amount and MALDI-MS/MS analysis proved each peptide was of the correct amino sequence. Therefore, there was no problem with the samples received from the manufacturer, New England Peptides. As these peptides were dissolved in PBS, they needed to be desalted before being able to be analysed by static infusion. This is because the salts present in PBS suppress the ionization of peptides. A static infusion experiment was sought in order to assess whether these peptides could be ionized and, if so, what the charge state of these peptides would be. However, when desalting according to the StageTip protocol developed in-house, no peptides were detected in the eluate following desalting. This was confirmed by MALDI analysis. Analysis by MALDI of the flow-through of each desalting procedure showed that the peptides of interest were in the flow-through amongst all the PBS salts, rather than in the eluate which should contain a minimal amount of salt. This indicates that the peptides did not bind to the C₁₈ of the StageTip. This offers a reason why these peptides could not be seen in an LC-MS run on the 4000 Q-TRAP. However, the manufacturer of these peptides shows clear evidence of each peptide sticking to the C₁₈ of their HPLC columns. After running these peptides on an offline HPLC system, we were able to show that these peptides can indeed bind to this C₁₈ column. Following this, several changes to the HPLC system up front of the 4000 Q-TRAP were made. These included substituting formic acid to TFA in the solvent the sample is dissolved in, changing the loading buffer to 100% water with TFA as opposed to formic acid, and changing the flow rates and gradient on the analytical column of the HPLC system. These changes in combination appeared to lead the peptides to indeed stick to the column. Therefore, the HPLC setup in front of the Q-TRAP needed to be changed. As a result, the formic acid used when trapping and washing samples was changed to the more efficient ion pairing agent, TFA. However, TFA is known to suppress ionisation in electrospray ionisation (375) so only a small amount was used and this was only in the loading buffer, with formic acid employed for the other gradient buffers. In addition, the trapping column was changed to the more efficient nanoViper columns which have recently been released by Dionex. These changes allowed the intact form of the AEP target peptide to be detected, as well as both halves of the cleaved AEP target post-digestion with AEP. This allowed the SRM assay to be developed.
4.4.5 Internal Standards

SRM linearity data of DVSKhph normalised against an internal standard (GluFib) added to each sample in a fixed amount (1 pmol) shows good correlation with these data without any internal standard normalisation (Figures 4.7 and 4.16(b) respectively). When normalising AEP plasma data against the GluFib internal standard based on these linearity data, a significant increase in variability is seen in the normalised data (Table 4.10). This indicates GluFib cannot be used as an internal standard to assay AEP activity in plasma samples.

When testing patient samples, the internal standard must be changed to one which elutes at the same time as the peptide DVSKhph. As all plasma samples analysed in this proof of concept testing are from one stock of plasma, the matrix effects from other species in these plasma samples will be the same. However the ionisation and spray stability can vary during an LC run in an unpredictable manner and so a standard is required which co-elutes with the peptide of interest and so is affected equally by these factors (376). The most common way of doing this is by using a heavy peptide incorporating heavy isotopes of carbon, nitrogen and/or oxygen (255). This will provide an increase in mass, and as such m/z, but the internal standard will have the same elution time as the cleaved AEP substrate, and as such will be subjected to the same matrix effects as this cleaved substrate. This type of internal standard is typically used in SRM workflows when assaying different samples (252). Therefore, GluFib can be used as a QC standard at best and not an internal standard. However, the presence of a heavy labelled internal standard of the same sequence as the target analyte replaces the need for GluFib acting as a QC standard.

4.5 Conclusion

The data presented in this chapter offer a proof of concept that the method developed in Chapter 3 has the potential to quantify AEP activity in plasma samples. However, further testing with patient samples with known clinical outcome in terms of patient response to L-asparaginase must be done in order to confirm that AEP activity in plasma can be used as a predictive marker for L-Asparaginase treatment response. This SRM based assay shows good correlation with the results produced by the fluorescence assay previously validated by Holland et. al. (351) in plasma samples, with a greater linear range and lower limits of AEP activity being able to detected be than when using the fluorescence method. This SRM assay overcomes the sensitivity and limited dynamic range problems associated with the fluorescence based assay. Additionally, this chapter shows that the method developed in Chapter 3 can be used as a qualitative assay to monitor AEP activity in WCL samples. Now
an assay has been produced which can quantify AEP activity in plasma orders of magnitude lower than current fluorescence methods, AEP’s potential as a predictive marker for L-Asparaginase treatment response must be assessed. This involves monitoring AEP activity in plasma samples from children who respond well to L-Asparaginase treatment and those who respond poorly. If a significant difference in AEP activity is seen between these patient populations, AEP may have potential as a predictive marker for L-Asparaginase treatment response, providing a clinician with information about whether a child will respond well to L-Asparaginase treatment before it has been administered.
Chapter 5: Assaying the Activity of the Protease, Caspase-3 using SRM MS

5.1 Introduction

5.1.1 Caspase-3

Caspase-3 is a cysteine dependent, aspartic acid specific protease involved in apoptosis as an executioner protease (377, 378). Its pro-enzyme (zymogen) form becomes cleaved by initiator caspases in the caspase cycle (TCC) and once activated by this caspase-8 cleavage, goes on to perform apoptosis in cells (158, 377, 379-382). Upon cleavage, the 32 kDa zymogen gets cleaved into a 17 kDa and a 12 kDa subunit, termed p17 and p12 respectively. These go on to form the active protease in the form of a heterotetramer, constituting two of each subunit. The crystal structure of caspase-3 has been described by Thornberry et. al. (383) (Figure 5.1).

Figure 5.1: Caspase-3 crystal structure when bound to a tetrapeptide aldehyde inhibitor (shown in red). The active (mature) caspase-3 consists of two p17 (purple) and two p12 subunits (grey). (taken from Thornberry et. al., reference (383)).

Caspase-3 is stable over a wide pH range, from pH 3 to pH 9, with its optimal activity at ~ pH 7.5 (Figure 5.2) (384). This allows caspase-3 to be active under both normal and apoptotic conditions (385). As the enzymes optimum activity range falls within the pH of native plasma (384), any caspase-3 present in plasma which has been activated previously should remain active as long as it has not been subject to further proteolysis or irreversible inhibition.

200
Caspase-3 has a high degree of substrate specificity and will almost exclusively cleave proteins and peptides after an aspartic acid residue (386-388). It has also been observed that caspase-3 has a high preference for the amino acid sequence Asp-x-x-Asp, at the P4’ to P1’ positions, cleaving at the C-terminus of this sequence (386). Additionally, caspase-3 has shown a preference for Gly or Ser at the C-terminal side of the cleavage site (387). Due to the highly specific nature of caspase-3 cleavage as well as the variety of diseases associated with this protease, inhibitors and drugs have been designed and developed to inhibit caspase-3 activity (389-391). All of these inhibitors contain the amino acid sequence DEVD, providing further evidence of the necessity of aspartic acid at specific positions of molecule in order for it to be recognised by caspase-3.

Caspase-3 is encoded by the CASP3 gene (392) and has been linked to cancer (including drug resistance) (392-399) and Alzheimer’s disease (400-402), with potential links to other diseases including myocardial infarction (403). In addition, caspase-3 has been linked to the differentiation of embryonic stem cells (404).
Expanding on this link between caspase-3 and cancer, down-regulation of caspase-3 has been linked to resistance to chemotherapeutic drugs. These include cis-platin, doxorubicin and etoposide when used to treat breast cancer (405), as well as cis-platin treatment in ovarian cancer (406). It has been hypothesised that this is due to caspase-3 playing such an essential role in apoptosis (see Chapter 1). A down-regulation in caspase-3 (assuming no mutations causing an alteration of the activity of the protease) would infer that cells are less susceptible to undergoing apoptosis under normal conditions. Therefore, tumour cells that have lower levels of caspase-3 are less able to undergo apoptosis, and as such will continue to proliferate. This would cause an increase in the size and progression of the tumour.

However, a link has been shown between up-regulation of caspase-3 and a recurrence of tumours which have been treated by radiotherapy (350). Huang et. al. proposed a mechanism of tumour regrowth based on downstream caspase-3 targets. In this mechanism, they use the fact that tumours being treated by radiotherapy undergo a great deal of cell death. These dying cells are then removed by macrophages and similar cells. However, any remaining tumour cells can undergo accelerated proliferation which can cause the tumour to re-grow. Huang et. al. have shown (using human xenografts) that tumour cells deficient in caspase-3 proliferate at a slower rate than tumour cells with higher amounts of caspase-3. It is thought that several downstream targets of caspase-3 are growth signalling molecules, including the lipid prostaglandin E$_2$. As these signalling molecules are regulated by caspase-3, the increased amounts of caspase-3 in these surviving tumour cells, cause the regrowth of the tumour (Figure 5.3) (350).
Figure 5.3: Caspase-3 causing tumour regrowth. A proposed mechanism for the up-regulation of caspase-3 leading to the proliferation of tumours previously treated by radiotherapy of chemotherapy (taken from Huang et. al., reference (350)). In this schematic, a tumour is treated by chemotherapy or radiotherapy. This kills the majority of the tumour cells, but not all of them. During the death of the majority of these tumour cells, caspase-3 is activated. This goes on to affect a number of downstream targets, including cytosolic calcium-independent phospholipase A2 (iPLA2). This activated iPLA2 then goes on to activate prostaglandin E2 (PGE2), a growth stimulating molecule. This can then cause the remaining tumour cells to grow at an increased rate than usual (350).

As described in this chapter, abnormal levels of caspase-3 (either up or down regulated) have been linked to a variety of cancers, including the likelihood of relapse. This and the fact that apoptotic bodies can undergo secondary necrosis, releasing their contents into the blood (207) mean that caspase-3 may be active in peripheral blood. This, along with the pH range in which caspase-3 is active (384) make caspase-3 an ideal target protease to monitor as a prognostic indicator in peripheral blood.
5.1.2 Aim of This Chapter
The aim of this chapter is to develop a method to quantitatively measure caspase-3 activity in plasma using Selected Reaction Monitoring (SRM) mass spectrometry. The known links between caspase-3 and tumour recurrence, as well as its optimal activity being at pH 7.4 indicates that caspase-3 activity could be quantified in plasma. If a link between caspase-3 activity in plasma and tumour regrowth in patients following cancer treatment was found, this assay would have a clear use in a clinical environment.
5.2 Methods and Materials

5.2.1 Chemicals
In all experiments, the water used was HPLC grade (Rathburn, Walkerburn, UK), acetonitrile was HPLC grade (Fisher Scientific UK Limited, Leicestershire, UK) and formic acid was mass spectrometry grade (Fluka, Sigma-Aldrich, St. Louis, MO). TFA was mass spectrometry grade (Sigma-Aldrich, St. Louis, MO). Caspase-3 was human recombinant caspase-3, ≥ 90% (SDS-PAGE), expressed in Escherichia coli (C-terminal histidine-tagged), buffered aqueous glycerol solution, (Sigma). 4-aminophenylmercuric acetate (APMA) was ≥ 90%. The caspase-3 specific target peptide (95% purity) was supplied by New England Peptides (Gardner, MA).

5.2.2 Mass spectrometry Analysis
All samples were analysed on a 4000 Q-TRAP (Applied Biosystems, Foster City, CA) coupled to a liquid chromatography system consisting of a Famos autosampler, a Switchos trap system and an Ultimate Plus dual pump (LC Packings, Dionex, Sunnyvale, CA). The column used on the liquid chromatography system was an Acclaim pepmap 100, 15 cm, C18 column, pore size 100 Å. The software used for the 4000 Q-TRAP was Analyst 1.4.1 (Applied Biosystems, Foster City, CA). The software for the liquid chromatography system was Chromeleon 6.50 SP4 Build 1000 (LC Packings, Dionex, Sunnyvale, CA). Data was processed using Multiquant 2.0 software (AB Sciex, Toronto, Canada) with the following parameters;

5.2.3 Synthetic Peptide Design
Synthetic peptides were designed using the MEROPS database. From this, amino acids most reported at each site from P4 to P4’ were chosen to be in the target peptide. Both the N- and C-termini of each peptide were end-capped with the sequence HPH. These HPH sequences were the D isoform of each amino acid in order to prevent exopeptidase digestion of the target peptide. The substrate used in this assay was of the amino acid sequence hphDEVDGAGShph, with amino acids in lower case correspond to the D isoform of that amino acid. Upon caspase-3 cleavage, this was predicted to yield two peptides – hphDEVD and GAGShph.
5.2.4 **Synthetic Peptide SRM Generation**

SRM transitions were generated using Protein Prospector. Collision energies were calculated as described in Chapter 4.

SRM transitions were generated for the intact form of each synthetic peptide, as well as the N- and C-terminal peptides post-digestion.

The SRM transitions to detect the intact form of the caspase-3 target peptide are described in Table 5.1.

Table 5.1: SRM transitions to detect hphDEVDGAGShph, hphDEVD and GAGShph.

<table>
<thead>
<tr>
<th>Peptide Sequence</th>
<th>Charge State</th>
<th>Q1 (m/z)</th>
<th>Q3 (m/z)</th>
<th>Collision Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hphDEVDGAGShph</td>
<td>2+</td>
<td>746.3</td>
<td>1354.6</td>
<td>35.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1257.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1120.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1005.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>876.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>777.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>662.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>605.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>534.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>477.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>390.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>253.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>156.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>235.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>372.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>487.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>616.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>715.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>830.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>887.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>958.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1015.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1102.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1239.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1336.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hphDEVDGAGShph</td>
<td>3+</td>
<td>497.9</td>
<td>1354.6</td>
<td>21.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1257.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1120.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1005.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>876.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>777.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>662.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>605.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>534.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2+</td>
<td>424.7</td>
<td>711.3</td>
<td>21.7</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>hphDEVD</td>
<td>2+</td>
<td>424.7</td>
<td>711.3</td>
<td>21.7</td>
</tr>
<tr>
<td></td>
<td>3+</td>
<td>283.5</td>
<td>711.3</td>
<td>13.3</td>
</tr>
<tr>
<td>GAGShph</td>
<td>2+</td>
<td>331.7</td>
<td>605.3</td>
<td>17.6</td>
</tr>
<tr>
<td></td>
<td>3+</td>
<td>283.5</td>
<td>711.3</td>
<td>13.3</td>
</tr>
</tbody>
</table>
5.2.5 SRM Background Levels in Plasma
Plasma prepared in from the same stock used in chapter 4 (25 µl) was enriched by acetonitrile precipitation (50 µl) as described in Chapter 3 and re-suspended in buffer A (1000 µl). Samples of this (55 µl) were analysed by SRM scanning for the intact as well as the N- and C- termini cleavage products of each synthetic peptide.

5.2.6 Target Peptide Digestion, Detection and SRM Optimisation
Recombinant human caspase-3 was (10 µg) dissolved in phosphate buffered saline pH 7.4 (PBS) and diluted to its required amount. As this protease is already in its active form, no further activation step was necessary (407).
The caspase-3 specific target synthetic peptide (1 µg) was incubated with activated caspase-3 (20 ng) at 37°C overnight in PBS at pH 7.4. Following this, the sample was dried to completion and re-suspended in buffer A to a final concentration of 1 pmol/µl. samples were analysed by IDA analysis. To create the HPLC buffers A and B, Switchos buffer and needle wash solution, water was HPLC grade (Rathburn, Walkerburn, UK), acetonitrile was HPLC grade (Fisher Scientific UK Limited, Leicestershire, UK) and formic acid was mass spectrometry grade (Fluka, Sigma-Aldrich, St. Louis, MO).
The loading buffer was 100% water acidified with 0.1% TFA (v/v). The trapping column was to an Acclaim Pepmap 100, 100 µm x 2 cm NanoViper C18, 5 µm, 100Å trap column (Dionex). The gradient over the analytical column is described in Table 5.2.
Buffer A & Switchos buffer consisted of 97.9% water, 2% acetonitrile and 0.1% formic acid (v/v/v).
Buffer B consisted of 79.9% acetonitrile, 20% water and 0.1% formic acid (v/v/v).
The needle wash solution was 100% water.
Table 5.2: Gradient conditions to detect hphDEVD and GAGShph.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Buffer A (%)</th>
<th>Buffer B (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>30</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>33</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>38</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>39</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>59</td>
<td>99</td>
<td>1</td>
</tr>
</tbody>
</table>

Samples were analysed using the SRM transitions described in Table 5.1. The six most intense SRM transitions were selected, and their collision energies optimised to the nearest 0.5 eV. These optimised transitions were used in future experiments (Table 5.3). The dwell time for all SRM transitions was 100.0 ms.

Table 5.3: Optimised SRM transitions to detect hphDEVD.

<table>
<thead>
<tr>
<th>Q1 (m/z)</th>
<th>Q3 (m/z)</th>
<th>Collision Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>424.7</td>
<td>134.0</td>
<td>22.7</td>
</tr>
<tr>
<td>424.7</td>
<td>616.2</td>
<td>20.2</td>
</tr>
<tr>
<td>424.7</td>
<td>233.1</td>
<td>23.2</td>
</tr>
<tr>
<td>424.7</td>
<td>235.1</td>
<td>24.7</td>
</tr>
<tr>
<td>424.7</td>
<td>715.3</td>
<td>19.2</td>
</tr>
<tr>
<td>424.7</td>
<td>614.2</td>
<td>17.7</td>
</tr>
</tbody>
</table>

5.2.7 Cleaved Peptide SRM Linearity
First, the samples used in section 5.2.9 underwent a serial dilution across four orders of magnitude (Table 5.4). In each sample, GluFib (1 pmol) was used as an internal standard using the optimised SRM transitions developed in Chapter 4.

Table 5.4: The amounts of hphDEVD used in linearity testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of Cleaved Peptide in Sample (fmol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>500</td>
</tr>
<tr>
<td>7</td>
<td>1,000</td>
</tr>
<tr>
<td>8</td>
<td>5,000</td>
</tr>
<tr>
<td>9</td>
<td>10,000</td>
</tr>
</tbody>
</table>
Each sample was analysed in triplicate by the optimised SRM method. The linearity of this SRM method was assessed by plotting the peak area of the most intense transition against relative sample concentration. This allowed the linearity of SRM transitions monitoring the cleaved and uncleaved forms of each target peptide to be assessed.

5.2.8 Caspase-3 Activity in Plasma
Caspase-3 (1 µg) was serially diluted with PBS to the amounts described in Table 5.5.

Table 5.5: The amounts of caspase-3 added to plasma in proof of concept testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of Caspase-3 Added to Plasma (pg/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12,000</td>
</tr>
<tr>
<td>2</td>
<td>6,000</td>
</tr>
<tr>
<td>3</td>
<td>3,000</td>
</tr>
<tr>
<td>4</td>
<td>1,200</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>PBS</td>
</tr>
</tbody>
</table>

Following this, the caspase-3 specific target peptide substrate (2 µg) was added to each sample. Samples were incubated at 37°C for 0, 1, 2, 4, 8, and 24 hours. At each time point, samples were quenched with formic acid (5 µl) and stored at -20°C until they were to be analysed. At this point, samples were enriched by acetonitrile precipitation (2 volumes) as described in Chapter 3. Samples were then dried to completion and re-suspended in buffer A (1 ml). This was then centrifuged at 21,000 g for 30 seconds. An aliquot of this (55 µl) was analysed by SRM to detect the cleaved form of the caspase-3 target peptide. In each sample, GluFib (1 pmol) was used as an internal standard using the optimised SRM transitions developed in Chapter 4. Samples were prepared in triplicate.

5.2.9 Sample Dilution Testing
A dilution series of a sample taken at the 24 hour time point containing the most caspase-3 was analysed by SRM (Table 5.6). The peak areas for each sample were then processed using MultiQuant software. Each peak area was then examined against the linearity plot of the cleaved peptide shown in section 5.2.10. From this the amount of each sample which needs to be analysed in order to generate reliable quantitative data was found. This amount of each sample was analysed by SRM.
Table 5.6: Sample dilution sample amounts.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of Sample Analysed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
</tr>
</tbody>
</table>

5.2.10 Caspase-3 Assay Failure Investigation

5.2.10.1 Has the Caspase-3 Target Substrate Degraded Over Time?
The caspase-3 substrate (10 µg) was dissolved in water:acetonitrile (50:50 v/v) acidified with 0.1% TFA and made up to concentration of 10 pmol/µl. An aliquot of this (2 µl) was mixed with an equal volume of a saturated α-cyano-4-hydroxycinnamic acid (CHCA) solution made in the same solvent as above. This was manually spotted onto a MALDI plate and analysed by MS & MS/MS using a 5800 MALDI–ToF/ToF (AB Sciex) looking for the cleaved and uncleaved forms of the caspase-3 target peptide.

5.2.10.2 The Effect of Acid in the nHPLC Buffers on Substrate Stability
Aliquots of the caspase-3 target substrate were stored at room temperature overnight under the conditions described in Table 5.7.

Table 5.7: Overnight substrate storage conditions to investigate the effect of acid on hphDEVGDAGShph stability.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Storage Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PBS</td>
</tr>
<tr>
<td>2</td>
<td>0.1% Formic Acid</td>
</tr>
<tr>
<td>3</td>
<td>10% Formic Acid</td>
</tr>
<tr>
<td>4</td>
<td>0.1% TFA</td>
</tr>
<tr>
<td>5</td>
<td>10% TFA</td>
</tr>
</tbody>
</table>

After overnight incubation, samples were dried to completion, re-suspended in water and acetonitrile (50:50 v/v) acidified with 0.1% TFA to a concentration of 10 pmol, and analysed by MALDI–ToF/ToF as described in section 3.15.1 – “Has the Caspase-3 Substrate Degraded Over Time?” to detect the cleaved and uncleaved forms of this target substrate.
5.2.10.3 Investigating Carry Over of hphDEVD in nHPLC-SRM Analysis

10 pmol of the cleaved caspase-3 substrate was analysed by LC-SRM. Following this, three blank injections containing 0.1% TFA in water were performed. From this, SRM response in these three injections were assessed in order to determine the carry-over of this cleaved caspase-3 peptide in further injections. This experiment was repeated using a sample containing the highest level of caspase-3 taken at the 24 hour time point in the section 5.2.11”, as opposed to 10 pmol of cleaved caspase-3 substrate with no plasma background.

5.2.10.4 The Effect of Storing Samples Post-Enrichment on the Stability of the Caspase-3 Substrate

Samples were prepared according to Table 5.8. These samples were then dried to completion, re-suspended in 0.1% TFA in water (v/v) and analysed by SRM analysis to detect the cleaved form of the caspase-3 substrate.

Table 5.8: Samples prepared post-acetonitrile precipitation to investigate hphDEVDGAGShph stability.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Substrate (2 µg) in PBS (20 µl), then…</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2 volumes of acetonitrile added, instant vortex (10 seconds), centrifugation then supernatant removal</td>
</tr>
<tr>
<td>2</td>
<td>2 volumes of acetonitrile added, left for 60 seconds on bench then vortexed, centrifugation and supernatant removal</td>
</tr>
<tr>
<td>3</td>
<td>2 volumes of acetonitrile added, instant vortex, then left for 60 seconds on bench before centrifugation and supernatant removal</td>
</tr>
<tr>
<td>4</td>
<td>2 volumes of acetonitrile added, left for 15 minutes on bench then vortexed, centrifugation and supernatant removal</td>
</tr>
<tr>
<td>5</td>
<td>2 volumes of acetonitrile added, instant vortex, then left for 15 minutes on bench before centrifugation and supernatant removal</td>
</tr>
<tr>
<td>6</td>
<td>2 volumes of acetonitrile added, left for 1 hour on bench then vortexed, centrifugation and supernatant removal</td>
</tr>
<tr>
<td>7</td>
<td>2 volumes of acetonitrile added, instant vortex, then left for 1 hour on bench before centrifugation and supernatant removal</td>
</tr>
</tbody>
</table>

5.2.10.5 The Effect of Each Stage of the Sample Preparation Method on the Stability of the Substrate

Samples were prepared according to Table 5.9. These samples were then dried to completion, re-suspended in 0.1% TFA in water (v/v) and analysed by MALDI–ToF/ToF analysis to detect the intact and cleaved forms of the caspase-3 substrate.
Table 5.9: Samples prepared to investigate each stage of the sample preparation method on the stability of hphDEVDGAGShph.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Substrate (2 µg) in PBS (20 µl), then…</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No Incubation</td>
</tr>
<tr>
<td>2</td>
<td>24 hour incubation at 37°C</td>
</tr>
<tr>
<td>3</td>
<td>As Sample 2 plus acetonitrile precipitation</td>
</tr>
<tr>
<td>4</td>
<td>As Sample 3 plus drying to completion and re-suspension in 0.1% TFA in water (v/v)</td>
</tr>
</tbody>
</table>

5.2.10.6 Previous Samples Analysed by SRM Now Analysed by MALDI-MS

Aliquots of samples (10 µl) previously analysed by SRM in section 3.13 – “Caspase-3 Activity in Plasma” were analysed by MALDI–ToF/ToF as described in previous sections (Table 5.10).

Table 5.10: The samples previously analysed by LC-SRM analysed by MALDI-ToF/ToF.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of Caspase-3 in Sample (pg/ml)</th>
<th>Time Point (hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>12,000</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>12,000</td>
<td>24</td>
</tr>
</tbody>
</table>

5.2.11 Caspase-3 Activity with Higher Levels of Caspase-3 in Plasma and/or Shorter Incubation Times

Caspase-3 (1 µg) was serially diluted with PBS to the amounts described in Table 5.11. Following this, the caspase-3 specific target peptide substrate (2 µg) was added to each sample. Samples were incubated at 37°C for 0, 4 and 8 hours. At each time point, samples immediately precipitated with acetonitrile (60 µl), vortexed for 10 seconds and centrifuged at 21,000 g for 30 seconds. The supernatant was then removed and dried to completion. Samples were then re-suspended in buffer A (1 ml). An aliquot of this (55 µl) was analysed by SRM to detect the cleaved form of the caspase-3 target peptide. Samples were prepared in triplicate.
Table 5.11: The amount of caspase-3 added to plasma in adjusted proof of concept testing.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of Caspase-3 Added to Plasma (pg/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24,000</td>
</tr>
<tr>
<td>2</td>
<td>12,000</td>
</tr>
<tr>
<td>3</td>
<td>6,000</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>PBS*</td>
</tr>
</tbody>
</table>

*PBS samples were taken at 0 and 8 hours only.

Samples were then re-prepared as described above, except for the following changes. Samples were incubated at 37°C for 0, 10, 20, 30, 60 and 120 minutes. Caspase-3 was added to plasma samples in the amounts described in Table 5.12.

Table 5.12: The amounts of caspase-3 added to plasma in proof of concept testing using increased levels of caspase-3.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Amount of Caspase-3 Added to Plasma (pg/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60,000</td>
</tr>
<tr>
<td>2</td>
<td>30,000</td>
</tr>
<tr>
<td>3</td>
<td>6,000</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>
5.3 Results

5.3.1 Caspase-3 Synthetic Peptide Target Design and SRM Transition Generation

The known cleavage sites described in the MEROPS database (353) for caspase-3 are shown in Table 5.13. From this, a substrate to assay caspase-3 activity in plasma was designed in the same way as described in Chapter 4.

Table 5.13: Known caspase-3 cleavages in the MEROPS database (10th July 2010).

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P1'</th>
<th>P2'</th>
<th>P3'</th>
<th>P4'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Giv</td>
<td>15</td>
<td>19</td>
<td>9</td>
<td>0</td>
<td>172</td>
<td>61</td>
<td>55</td>
<td>53</td>
</tr>
<tr>
<td>Pro</td>
<td>6</td>
<td>6</td>
<td>57</td>
<td>0</td>
<td>5</td>
<td>40</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>Ala</td>
<td>19</td>
<td>27</td>
<td>26</td>
<td>0</td>
<td>73</td>
<td>75</td>
<td>47</td>
<td>45</td>
</tr>
<tr>
<td>Val</td>
<td>22</td>
<td>38</td>
<td>142</td>
<td>0</td>
<td>12</td>
<td>41</td>
<td>43</td>
<td>31</td>
</tr>
<tr>
<td>Leu</td>
<td>17</td>
<td>51</td>
<td>67</td>
<td>0</td>
<td>20</td>
<td>45</td>
<td>52</td>
<td>39</td>
</tr>
<tr>
<td>Ile</td>
<td>7</td>
<td>15</td>
<td>34</td>
<td>0</td>
<td>9</td>
<td>14</td>
<td>25</td>
<td>35</td>
</tr>
<tr>
<td>Met</td>
<td>11</td>
<td>19</td>
<td>13</td>
<td>0</td>
<td>6</td>
<td>14</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Phe</td>
<td>5</td>
<td>14</td>
<td>13</td>
<td>0</td>
<td>14</td>
<td>9</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Tyr</td>
<td>6</td>
<td>20</td>
<td>13</td>
<td>0</td>
<td>17</td>
<td>20</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>Thr</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Ser</td>
<td>54</td>
<td>51</td>
<td>21</td>
<td>0</td>
<td>119</td>
<td>57</td>
<td>49</td>
<td>45</td>
</tr>
<tr>
<td>Thr</td>
<td>27</td>
<td>25</td>
<td>83</td>
<td>0</td>
<td>20</td>
<td>17</td>
<td>29</td>
<td>40</td>
</tr>
<tr>
<td>Cys</td>
<td>9</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>10</td>
<td>2</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Asn</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>25</td>
<td>19</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>Gin</td>
<td>6</td>
<td>30</td>
<td>10</td>
<td>0</td>
<td>4</td>
<td>35</td>
<td>20</td>
<td>31</td>
</tr>
<tr>
<td>Asp</td>
<td>270</td>
<td>38</td>
<td>8</td>
<td>1</td>
<td>561</td>
<td>15</td>
<td>18</td>
<td>28</td>
</tr>
<tr>
<td>Gli</td>
<td>56</td>
<td>156</td>
<td>17</td>
<td>1</td>
<td>8</td>
<td>18</td>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td>Lys</td>
<td>3</td>
<td>10</td>
<td>7</td>
<td>0</td>
<td>5</td>
<td>39</td>
<td>40</td>
<td>32</td>
</tr>
<tr>
<td>Arg</td>
<td>2</td>
<td>7</td>
<td>17</td>
<td>0</td>
<td>12</td>
<td>18</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>His</td>
<td>3</td>
<td>12</td>
<td>9</td>
<td>0</td>
<td>9</td>
<td>11</td>
<td>18</td>
<td>10</td>
</tr>
</tbody>
</table>

The following peptide sequence was used as a caspase-3 specific peptide target:

hphDEVDGAGShph

Caspase-3 cleavage has been designed to take place between the D and G amino acid residues. The peptide has been capped at each end with a HPH amino acid sequence. These amino acids are of the non-natural isoform (D-isoform) and as such will inhibit exoprotease activity, helping target the site of digestion of this peptide to one particular area.

In order to ensure that the intact and predicted cleaved forms of this peptide are not found endogenously in humans, the sequences HPHDEVDGAGSHPH, HPHDEV and GAGSHPH were BLAST searched against the human proteome. All three peptides showed no proteins containing these amino acid sequences. This indicates that these peptide
sequences are not endogenously found in humans and as such are suitable for use in this assay.

Following this, SRM transitions were generated using the MS Product feature in Protein Prospector to detect the intact form of the caspase-3 specific peptide, sequence hphDEVDGAGShph as well as the cleaved forms of this substrate, sequences hphDEVD and GAGShph (Table 5.1). As the charge state for this peptide in ESI was not known, SRM transitions were designed to detect these peptides in both their 2+ and 3+ charge states.

Following this, SRM analysis of a pooled plasma sample from ten healthy volunteers post-enrichment by acetonitrile precipitation showed no background response in plasma (Figure 5.4).
Figure 5.4: SRM responses from endogenous plasma peptides when detecting (a) the N-terminal and (b) the C-terminal form of the cleaved caspase-3 substrate. The low response in these XICs shows that there is no interference from endogenous peptides present in plasma enriched by acetonitrile precipitation at these SRM transitions. Any response greater than 100 counts per second is deemed significant, as illustrated by the dashed line.
The cleaved forms of the caspase-3 substrate, sequences hphDEVD and GAGShph were detected by SRM (Figure 5.5). However, due to poor chromatography, with significant SRM responses being seen between 7.5 and 24 minutes and a mixture of charge states, the C-terminal form of the cleaved target caspase-3 peptide, sequence GAGShph, will not be used to monitor caspase-3 activity in plasma for the remainder of this project (Figure 5.5 (c) and (d)).
Figure 5.5: Cleaved caspase-3 substrate SRM detection. Extracted ion chromatograms for the potential SRM responses for (a) the N-terminal cleavage product of the caspase-3 target peptide, sequence hphDEVD, (b) SRM transitions at 8.51 minutes, (c) the C-terminal cleavage product of the caspase-3 target peptide, sequence GAGShph, and (d) SRM transitions at 8.37 minutes when analysed on a 4000 Q-TRAP operating in SRM mode.
As only a minimal response is observed for transitions detecting the peptide in a 3+ charge state, the peptides hphDEVD and GAGShph are typically doubly charged when ionised by ESI. However, there is a significant response for transitions monitoring the 3+ charge state of the C-terminal peptide. Due to poor chromatography for the C-terminal peptide, along with the peptide being in both the 2+ and 3+ charge states, this peptide will not be monitored in this assay. Therefore, only the N-terminal form of this cleaved substrate will be used to quantify caspase-3 activity in plasma samples in all remaining.

The six most intense transitions for the peptide hphDEVD were selected and their collision energies optimised to the nearest 0.5 eV. These optimised transitions were used in all remaining experiments to detect caspase-3 activity in plasma using SRM analysis.

5.3.2 Cleaved Peptide SRM Linearity
When analysing the N-terminal form of the cleaved caspase-3 target peptide, sequence hphDEVD, by SRM analysis, a linear response across four orders of magnitude, with a limit of detection of 1 fmol was seen (Figure 5.6).

![Figure 5.6: Cleaved caspase-3 substrate linearity. SRM analysis of hphDEVD showed a linear response spanning four orders of magnitude \((r^2 = 0.9998)\) with a limit of detection of 1 fmol and a limit of quantitation of 5 fmol. Error bars represent the standard deviation of each sample population \((n = 3)\).](image-url)
Therefore, samples producing a response with a peak area in the range of $2.42 \times 10^3$ – $3.82 \times 10^6$ can be quantified in this assay. This corresponds to a concentration range of 5 fmol – 10 pmol.

### 5.3.3 Caspase-3 Activity in Plasma

The ability of the assay to measure caspase activity was tested by spiking differing known amounts of the protease into plasma. When these samples were analysed in a random order, there was no correlation between the amount of caspase-3 in the sample and the observed SRM area. Additionally, these results show no correlation between the time samples have been incubated before quenching the reaction and the amount of cleaved target substrate produced. This indicates that this assay cannot be performed in its current form (Table 5.14).

**Table 5.14: SRM responses of ten randomly analysed samples in proof of concept testing.**

<table>
<thead>
<tr>
<th>Caspase-3 Concentration (pg/ml)</th>
<th>Time Point (hrs)</th>
<th>Cleaved Peptide Peak Area</th>
<th>Internal Standard Peak Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS</td>
<td>4</td>
<td>$2.69 \times 10^5$</td>
<td>$2.37 \times 10^5$</td>
</tr>
<tr>
<td>$0.00 \times 10^9$</td>
<td>0</td>
<td>$3.68 \times 10^5$</td>
<td>$1.61 \times 10^5$</td>
</tr>
<tr>
<td>$0.00 \times 10^7$</td>
<td>1</td>
<td>$3.81 \times 10^5$</td>
<td>$1.55 \times 10^5$</td>
</tr>
<tr>
<td>$0.00 \times 10^7$</td>
<td>2</td>
<td>$3.09 \times 10^5$</td>
<td>$1.89 \times 10^5$</td>
</tr>
<tr>
<td>$1.50 \times 10^3$</td>
<td>2</td>
<td>$3.36 \times 10^5$</td>
<td>$4.14 \times 10^5$</td>
</tr>
<tr>
<td>$3.00 \times 10^3$</td>
<td>4</td>
<td>$2.64 \times 10^5$</td>
<td>$3.79 \times 10^5$</td>
</tr>
<tr>
<td>$6.00 \times 10^3$</td>
<td>1</td>
<td>$2.27 \times 10^5$</td>
<td>$3.51 \times 10^5$</td>
</tr>
<tr>
<td>$6.00 \times 10^4$</td>
<td>24</td>
<td>$3.11 \times 10^5$</td>
<td>$5.35 \times 10^5$</td>
</tr>
<tr>
<td>$1.20 \times 10^6$</td>
<td>0</td>
<td>$2.51 \times 10^5$</td>
<td>$3.87 \times 10^5$</td>
</tr>
<tr>
<td>$1.20 \times 10^6$</td>
<td>8</td>
<td>$2.88 \times 10^5$</td>
<td>$1.64 \times 10^5$</td>
</tr>
</tbody>
</table>

When examining the LC-SRM peaks for the peptide hphDEVD, peak widths at half maximum height were less than 30 seconds wide. Additionally, no side peaks were observed for any samples (Figure 5.7).
Figure 5.7: LC-SRM of cleaved caspase-3 substrate when assaying caspase-3 activity in plasma for (a) a plasma sample containing caspase-3 at a concentration of $1.2 \times 10^4$ pg/ml after incubation at 37°C for 8 hours, and (b) a plasma sample with no additional caspase-3 and no incubation at 37°C. These show similar peak widths and intensities. Spectra (c) and (d) show SRM responses for all six transitions for samples (a) and (b) respectively.
5.3.4 Caspase-3 Assay Investigation
Following the lack of a significant change in SRM response when assaying caspase-3 activity in plasma, it was suspected that either caspase-3 is not active in plasma due to endogenous caspase-3 inhibitors in plasma, degradation of the caspase-3 substrate, or the sample preparation method itself is the cause of this. As it is not possible to measure the effect of endogenous caspase-3 inhibitors on this assay, the caspase-3 substrate was examined to determine whether it had degraded over time. In addition, each stage of the sample preparation protocol was examined in turn in an attempt to find the origin of these consistent SRM responses, regardless of the amount of caspase-3 in a plasma sample, or the sample’s incubation time (Figure 5.8).
1. Thaw Caspase-3 Target Substrate
2. Add Caspase-3 Target Substrate to Plasma (1 μg)
3. Incubate Sample for x Minutes at 37°C
4. Add Acetonitrile to Sample (two volumes)
5. Vortex Sample for Ten Seconds
6. Centrifuge Sample for 30 Seconds at 9,300 g
7. Remove Supernatant and Place into Clean Eppendorf Tube
8. Dry Supernatant to Completion
9. Resuspend Dried Supernatant in 0.1% TFA in Water (100 μl)
10. Place 10% of Sample in HPLC Vial
11. Add GluFib Internal Standard (1 pmol)
12. Analyse by LC-SRM
5.3.4.1 Has the Caspase-3 Substrate, Sequence hphDEVDGAGShph Degraded over Time?

As the caspase-3 substrate peptide had been stored at -20°C for up to twelve months, it was suspected that the substrate may have degraded over time under these conditions. In order to investigate this, a sample of the intact caspase-3 substrate was analysed in order to determine whether the substrate itself had degraded over time. MALDI-MS analysis shows that the only form of the caspase-3 substrate which could be detected was the intact form of the peptide, sequence hphDEVDGAGShph (Figure 5.9). This shows that the caspase-3 target substrate has not completely degraded over time.
Figure 5.9: MALDI-ToF/ToF data showing the caspase-3 substrate has not degraded over time. (a) MS generated on a 5800 MALDI-ToF/ToF instrument showing no cleaved caspase-3 target substrate (m/z 848.35 or 662.35) and (b) MS/MS data from the same instrument sequencing the peptide of m/z 1491.4249. These data indicate that the substrate has not degraded over time.
5.3.4.2 Using MALDI-MS to Determine the Effect of Each Stage of the Sample Preparation Method on the Spontaneous Production of hphDEVD

The effect of each stage of the sample preparation method on the stability of the caspase-3 substrate was tested by MALDI-MS. This showed no increase in the cleaved forms of the caspase-3 substrate, indicating that the process of preparing samples for SRM analysis does not cause the degradation of the caspase-3 substrate (Figure 5.10).
Figure 5.10: MALDI-ToF/ToF data showing the caspase-3 substrate at each stage of sample preparation. MS spectra generated on a 5800 MALDI-ToF/ToF instrument of a sample of hphDEVDGAGShph showing the lack of a peak at m/z 848.35 following (a) no incubation of the peptide at 37°C (b) incubation at 37°C, (c) incubation at 37°C followed by acetonitrile precipitation, (d) incubation at 37°C, followed by acetonitrile precipitation then drying to completion and re-suspension in 0.1% TFA in water. The absence of a peak at
5.3.4.3 The Effect of Formic Acid on the Degradation of the Caspase-3 Target Peptide?

Whilst it has been reported previously that formic acid can cleave peptides at the C-terminus of aspartic residues (408), no evidence of this was seen when incubating peptides overnight in 10% formic acid (Figure 5.11). This indicates that the caspase-3 substrate should not be degraded by formic acid during the time it is passing through the HPLC system.

Figure 5.11: MALDI-ToF/ToF data showing the effect of formic acid on the caspase-3 substrate. No predicted degradation products as a result of formic acid hydrolysis after overnight incubation in 10% formic acid.

The absence of peaks at m/z 505.22, (corresponding to the peptide hphD) and m/z 848.35 (corresponding to the peptide hphDEVD) indicate that acid hydrolysis has not taken place after overnight incubation of the peptide in 10% formic acid. Therefore, hydrolysis of the caspase-3 substrate is not expected to take place during the 8.5 minutes the caspase-3 substrate is exposed to 0.1% formic acid in the HPLC system.
5.3.4.4 Investigating Carry Over of hphDEVD in nHPLC-SRM Analysis

Now it has been proven that the time the caspase-3 substrate is exposed to formic acid is not sufficient to cause proteolysis and the generation of hphDEVD in significant amounts, the issue of carry-over in this assay was investigated. In the injections immediately following the injection of samples containing the cleaved form of the caspase-3 target peptide, no SRM response was seen. This was the case for samples previously analysed in linearity testing, and samples with caspase-3 added to plasma and incubated for 24 hours at 37°C. This indicates that there is no problem of carry-over in this assay (Figure 5.12).
Figure 5.12: LC-SRM data investigating carry over in this assay. (a) the injection of a blank sample immediately following the analysis of 10 pmol of the peptide hphDEVD and (b) the injection of a sample containing caspase-3 at a concentration of 12,000 pg/ml after 24 hour incubation at 37°C

The lack of SRM response (SRM response < 100 cps) in these spectra indicates that carry-over is not an issue in this assay and as such cannot explain the SRM responses seen in Table 5.14.

5.3.4.5 Can Caspase-3 Interact with its Target Substrate Post-Acetonitrile Precipitation?

Potentially precipitation of protein may not inactivate proteases. Therefore the effect of the precipitation on caspase-3 activity was tested. No cleaved substrate (hphDEVD) was seen by MALDI-MS. This shows that there is no detectable interaction between precipitated caspase-3 and its target substrate post-acetonitrile precipitation (Figure 5.13). This shows that leaving the supernatant of plasma on the bench for one hour post-acetonitrile precipitation does not cause the degradation of hphDEVGDAGShph. Therefore, this is not the cause of the responses seen in Table 5.12.

Figure 5.13: MALDI-ToF/ToF data investigating the interaction of caspase-3 and its substrate post-acetonitrile precipitation.
The absence of a peak at m/z 848.35 in Figure 5.13 shows the lack of the peptide hphDEVD.

5.3.4.6 Using MALDI-ToF/ToF MS to Analyse Samples Previously Assayed by SRM

Since the investigation into the cause of large SRM responses has not yielded a cause for these results, samples previously analysed by SRM were analysed by MALDI-MS in order to detect the cleaved form of the caspase-3 substrate as a result of caspase-3 proteolysis. When analysing samples previously analysed by SRM, MALDI-MS analysis showed no cleaved peptide (Figure 5.14). Therefore, the response seen in section 4.3 may be an artefact, rather than a genuine response from hphDEVD. However, this needs investigating by analysing a pure form of the peptide hphDEVD by MALDI-ToF/ToF MS and MS/MS in order to determine whether this peptide can be ionised by MALDI ionisation.

Due to the lack of an SRM response when analysing a pooled plasma sample containing no caspase-3 substrate (Figure 5.4(a)), the source of this artefact is unknown. The presence of a high amount of contamination from endogenous plasma peptides when attempting to detect the other half of the cleaved caspase-3 substrate, sequence GAGShph (Figure 5.5(c)), means that this peptide cannot be used to assay caspase-3 activity in plasma samples.
Figure 5.14: MALDI-ToF/ToF data of samples previously analysed by LC-SRM. Samples containing hphDEVDGAGShph and (a) caspase-3 at a concentration of 24,000 pg/ml after no incubation, (b) caspase-3 at a concentration of 24,000 pg/ml after 24 hour incubation, (c) no caspase-3 after no incubation, (d) no caspase-3 after 24 hour incubation which have previously been analysed by LC-SRM on a 4000 Q-TRAP operating in SRM mode.

5.3.5 Caspase-3 Activity in Plasma Using Higher Levels of Caspase-3 in Samples

It was suspected that longer incubation times cause further proteolysis of the newly cleaved caspase-3 substrate, sequence hphDEVD. This is because this newly formed peptide is only end-capped with D-amino acids at the N-terminus. So as to overcome this issue leading to the lack of a significant change in SRM response for the peptide hphDEVD with different amounts of caspase-3 in plasma, the assay described in section 4.3 – “Caspase-3 Activity in Plasma” was repeated using higher amounts of caspase-3 but with shorter incubation times. Ten samples analysed in a random order show no correlation between the amount of caspase-3 in the sample and the amount of cleaved target substrate produced. Additionally, these results show no correlation between the time samples have been incubated before quenching the reaction and the amount of cleaved target substrate produced. This indicates that this assay cannot be performed in its current form (Table 5.15).
Table 5.15: SRM responses of ten randomly analysed samples in adjusted proof of concept testing.

<table>
<thead>
<tr>
<th>Caspase-3 Concentration (pg/ml)</th>
<th>Time Point (hrs)</th>
<th>Cleaved Peptide Peak Area</th>
<th>Internal Standard Peak Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS</td>
<td>0</td>
<td>3.52x10^5</td>
<td>1.43x10^5</td>
</tr>
<tr>
<td>PBS</td>
<td>8</td>
<td>2.26x10^5</td>
<td>1.44x10^5</td>
</tr>
<tr>
<td>0.00x10^9</td>
<td>4</td>
<td>1.37x10^5</td>
<td>1.55x10^5</td>
</tr>
<tr>
<td>0.00x10^9</td>
<td>8</td>
<td>1.86x10^5</td>
<td>1.39x10^5</td>
</tr>
<tr>
<td>6.00x10^3</td>
<td>0</td>
<td>1.47x10^5</td>
<td>1.33x10^5</td>
</tr>
<tr>
<td>6.00x10^3</td>
<td>4</td>
<td>2.19x10^5</td>
<td>1.61x10^5</td>
</tr>
<tr>
<td>1.20x10^4</td>
<td>4</td>
<td>3.69x10^5</td>
<td>6.00x10^5</td>
</tr>
<tr>
<td>2.40x10^4</td>
<td>0</td>
<td>1.70x10^5</td>
<td>1.25x10^5</td>
</tr>
<tr>
<td>2.40x10^4</td>
<td>4</td>
<td>3.17x10^5</td>
<td>1.04x10^5</td>
</tr>
<tr>
<td>2.40x10^4</td>
<td>8</td>
<td>3.99x10^5</td>
<td>1.45x10^5</td>
</tr>
</tbody>
</table>

5.3.6 Caspase-3 Activity in Plasma Using Shorter Incubation Times

It was suspected that the cause of no significant change in SRM response could be due to the cleavage of hphDEVDGAGShph, forming hphDEVD; this peptide could then be subject to further proteolysis, as was seen when investigating AEP activity in plasma and whole cell lysate samples in Chapter 4. Therefore, samples were recreated using greater amounts of caspase-3 and shorter incubation times. SRM analysis of ten samples analysed in a random order show no correlation between the amount of caspase-3 in the sample and the amount of cleaved target substrate produced. Additionally, these results show no correlation between the time samples have been incubated before quenching the reaction and the amount of cleaved target substrate produced. This indicates that this assay cannot be performed in its current form (Table 5.16).
Table 5.16: SRM responses of ten randomly analysed samples in proof of concept testing using increased levels of caspase-3.

<table>
<thead>
<tr>
<th>Caspase-3 Concentration (pg/ml)</th>
<th>Time Point (mins)</th>
<th>Cleaved Peptide Peak Area</th>
<th>Internal Standard Peak Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00x10⁷</td>
<td>60</td>
<td>3.37x10⁵</td>
<td>1.14x10⁵</td>
</tr>
<tr>
<td>6.00x10⁴</td>
<td>0</td>
<td>2.87x10⁵</td>
<td>1.85x10⁵</td>
</tr>
<tr>
<td>6.00x10⁴</td>
<td>10</td>
<td>3.08x10⁵</td>
<td>1.33x10⁵</td>
</tr>
<tr>
<td>6.00x10⁴</td>
<td>30</td>
<td>4.05x10⁵</td>
<td>1.63x10⁵</td>
</tr>
<tr>
<td>6.00x10⁴</td>
<td>60</td>
<td>3.65x10⁵</td>
<td>1.33x10⁵</td>
</tr>
<tr>
<td>3.00x10⁴</td>
<td>120</td>
<td>9.13x10⁵</td>
<td>1.28x10⁵</td>
</tr>
<tr>
<td>6.00x10⁴</td>
<td>0</td>
<td>3.72x10⁵</td>
<td>1.33x10⁵</td>
</tr>
<tr>
<td>6.00x10⁴</td>
<td>0</td>
<td>4.00x10⁵</td>
<td>1.92x10⁵</td>
</tr>
<tr>
<td>6.00x10⁴</td>
<td>20</td>
<td>6.94x10⁵</td>
<td>1.24x10⁵</td>
</tr>
<tr>
<td>6.00x10⁴</td>
<td>120</td>
<td>5.23x10⁵</td>
<td>1.47x10⁵</td>
</tr>
</tbody>
</table>
5.4 Discussion
The results in this chapter show an assay that potentially offered a linear range spanning four orders of magnitude, with a limit of detection of 1 fmol of cleaved synthetic peptide substrate (Figure 5.6). The sensitivity and wide dynamic range of this assay would make it ideal for a clinical setting, should caspase-3 become used as a biomarker or prognostic indicator. As discussed in Chapters 3 and 4, the number of SRM transitions used to in this assay exceed the requirements suggested by Sherman et. al. (262, 332). This, along with the known retention time of the peptide, allows the cleaved peptide substrate to be confidently identified, despite the highly complex background this peptide is in.

However, there is no correlation between the amount of active caspase-3 present in a plasma sample, and the SRM response for the cleaved form of the caspase-3 specific target synthetic peptide (Table 5.21). This is shown by adding different amounts of active recombinant caspase-3 to plasma.

5.4.1 Caspase-3 Assay Investigation
When attempting to quantify caspase-3 activity in plasma, in the first instance, no significant change in SRM response was seen. MALDI-MS analysis showed that the caspase-3 target substrate, sequence hphDEVDGAGShph had not degraded over time and was not the source of this SRM response.

It has previously been reported that formic acid can cleave peptides with a high specificity at aspartyl residues (408). As there are two aspartyl residues in the caspase-3 target substrate, it was suspected that the peptide could be cleaved by the formic acid used in the HPLC buffers. This is especially problematic as this could lead to the production of the peptides hphD and hphDEVD. hphDEVD is the amino acid sequence of the cleaved form of the target substrate being used to monitor caspase-3 activity in this assay. Therefore, in order to determine whether formic acid hydrolysis of this peptide was taking place, samples of hphDEVDGAGShph were incubated overnight in both 0.1% and 10% formic acid. MALDI-MS data showed no formic acid hydrolysis of these peptides. Therefore, the presence of an SRM response cannot be attributed to formic acid hydrolysis of this peptide.

It was also shown that carry-over was not a problem in this assay as injections of 0.1% TFA in water immediately following SRM analysis of samples known to contain hphDEVD showed no SRM response. Therefore, the similar SRM response seen in all time course samples could not be attributed to carry-over of hphDEVD from injection to injection in this assay.
MALDI-MS analysis also showed no degradation of hphDEVDGAGShph as a result of interaction of the supernatant and the precipitant of samples post-acetonitrile enrichment, nor from any stage of the sample preparation method used in this study. MALDI-MS analysis of samples previously analysed by SRM showed no peak at m/z 848.35. This indicates that the caspase-3 substrate has not been cleaved by caspase-3 in any of the samples analysed by MALDI-MS and that the SRM response seen was not a result of proteolysis of hphDEVDGAGShph by caspase-3.

As a result of these tests, the origin of this significant SRM response cannot be established. However, as this investigation has not shown that the cause of this consistent SRM response is due to substrate degradation or the sample preparation method, it is possible that any caspase-3 present in the sample is inactivated by caspase-3 inhibitors naturally present in plasma. These include members of the inhibitors of apoptosis (IAP) family, such as XIAP, c-IAP1 and c-IAP2 which have been shown to directly bind to caspase-3 and inhibit its activity (409). However, the effects of these inhibitors on this assay cannot be directly determined. Whilst caspase-3 activity in rat plasma has been assayed previously (410, 411), this plasma is typically diluted with PBS and significant activity increases in caspase-3 are only seen up to twelve hours after injection with the drug thioacetamide, a drug that stimulates apoptosis in rat liver (411). Whilst caspase-3 levels have been shown to change response to a disease or treatment (350, 405, 412, 413), no reports have shown a change in caspase-3 activity in response to a disease or treatment. This may be because caspase-3 is inactive in plasma due to endogenous caspase inhibitors, despite its optimal activity being at the pH of native plasma.

### 5.5 Conclusion

In conclusion, the data presented in this chapter indicate that assessing caspase-3 activity in plasma using an LC-SRM based assay in its current form cannot potentially be used as a prognostic indicator for predicting cancer patient response to treatment in the future. This is because there is no clear correlation between the amount of cleaved caspase-3 target peptide produced and either amount of caspase-3 present in the sample, nor the amount of time the sample has been incubated. MALDI-MS data does not confirm the response seen in SRM analysis, despite the responses seen in six SRM transitions which is typically enough to capture 80% of the total ion count of a peptide (414). However, the lack of SRM response when analysing an enriched plasma sample containing no hphDEVD peptide (Figure 5.4(a))
show no response from endogenous plasma peptides. Therefore, the positive SRM results seen in time course testing cannot be attributed to these endogenous peptides. This problem was not circumvented by increased levels of caspase-3, nor shorter incubation times. Additionally, an investigation into the stability of the caspase-3 substrate throughout the various stages of the sample preparation method did not uncover a source of hphDEVDGAGShph degradation. Therefore, the origin of the SRM responses seen in these time course assays cannot be attributed to endogenous plasma peptides or the sample preparation workflow. As a result, the source of these SRM responses cannot be found. A potential way forward is recommended: the generation of a novel caspase-3 substrate should be considered.

When re-designing this peptide, it is recommended that the sequence DEVDGAGS remain unchanged. This is because the MEROPS database describes this amino acid sequence as the most likely to be cleaved by caspase-3. In addition, the substrate for a fluorescence assay supplied by the Cayman Chemical Company (Ann Arbor, MI) contains the amino acid sequence DEVD, with proteolysis taking place at the C-terminus of this sequence. This provides more evidence that this amino acid sequence is ideal for caspase-3 cleavage. However, the end-capping D-amino acid sequences can be altered. Based on the findings in Chapter 4, using less hydrophilic amino acid residues may be better suited to aid the binding to C<sub>18</sub> HPLC columns and the detection of the intact and cleaved forms of target synthetic peptide by HPLC-SRM. Therefore, an alternative caspase-3 substrate has been proposed. This substrate has the amino acid sequence:

**wkvDEVDGAGSaly**

This incorporates one polar amino acid at the each end cap which can sequester positive charge in ESI, aiding ionisation, and two non-polar, neutral amino acids which will reduce the hydrophilicity of the intact and cleaved forms of this substrate, aiding in their detection by HPLC-SRM. Additionally, BLAST searching the amino acid sequences, WKVDEVGDAGSALY, WKVDEV and GAGSALY against the human proteome result in no protein hits. This indicates that these peptides are all non-endogenous in humans, an essential requirement in the design of this substrate to assay caspase-3 activity in plasma samples.

If this substrate shows no significant change in the amount of degradation with different amounts of caspase-3 present in plasma, this may be due to naturally occurring caspase-3 inhibitors present in plasma, including members of the inhibition of apoptosis (IAP) family.
These bind irreversibly to caspase-3, rendering it inactive. This may be the cause of the failure of this assay (415).

In order to determine the cause of the failure of this assay, further work needs to be performed. At present, the source of the positive SRM response, but negative MALDI-ToF/ToF response for the cleaved form of the substrate hphDEVDGAGShph is not known. One possible explanation is that the caspase-3 used in this study is inactive, and the SRM responses seen are due to contamination. However, this seems unlikely as 21 out of 22 possible SRM transitions to detect the cleaved forms of the substrate hphDEVD and GAGShph (Figure 5.5) showed a positive response. One possible explanation for this is that these peptides can be ionised via ESI but not MALDI ionisation, but further work using techniques such as nuclear magnetic resonance (NMR) spectroscopy may be required to confirm that this is the case.

An additional study which must be performed involves the development of the caspase-3 fluorescence assay (Cayman Chemical Company). Once this assay has been developed, the caspase-3 fluorescent substrate can be incubated with the activated caspase-3 used in this study both in PBS and in plasma. If the caspase-3 is active, it will cleave the fluorescent substrate, and a fluorescent response will be seen. If no increase in fluorescence is detected in the PBS samples, it can be assumed that the caspase-3 used in this study is inactive. If a fluorescent response is seen in the PBS samples, but not the plasma samples, it can be assumed that the caspase-3 used in this study is active, but becomes inactivated in plasma, either by other proteases naturally occurring in plasma, or by IAPs.
6 Chapter 6: Final Conclusions and Future Work

6.1 Final Conclusions

This thesis described the development of a mass spectrometry based assay to monitor protease activity in plasma. We have determined that due to the lower limit of detection and limited number of targets in this assay, the optimal way to monitor these proteolytic products is by using an MRM-based approach, as opposed to the novel SBITS methodology.

We also generated a series of criteria to be used when designing synthetic peptides to monitor protease activity in this project. These address aspects such as targeting digestion of the target to one particular area, ensuring the peptide is not endogenous in humans and can be detected by ESI MS. This study also demonstrates the utility of using D-isoforms in target synthetic peptides and its ability to inhibit proteases, using trypsin as a model system. This highlights the ability to focus digestion to take place in one area of the target peptide. This increases the selectivity and specificity of the assay and increases the signal:noise ratio, producing a more sensitive assay.

This study also included a comprehensive study into which of the many published methods is best at preparing plasma samples for MS analysis. This study consisted of determining which method is best at removing intact protein from a plasma sample, which could block nHPLC columns, whilst retaining peptides. This in depth study concluded that precipitation with two volumes of acetonitrile was the method best at removing intact protein, whilst retaining the highest amount of peptide in the most reproducible manner. These results were incorporated into this assay and acetonitrile precipitation was used to prepare all plasma samples in this project for SRM MS analysis.

Following this, we showed that this method shows a clear correlation between the amount of active protease in the plasma sample and the amount of cleaved synthetic peptide produced over a 24 hour period using trypsin as a model protease. Additionally, this proof of concept study showed that the entire process of incubating a synthetic peptide in a plasma sample with different amounts of active trypsin, quenching the reaction at various time points with formic acid, enriching the sample with acetonitrile precipitation, drying the sample to completion, re-suspending it in MS-compatible buffers, and monitoring the cleaved peptide target by MRM was successful.

Following this, two clinically relevant proteases were assessed for their potential as prognostic indicators. The first of these was asparaginyl endopeptidase (AEP), which has links to childhood acute lymphoblastic leukaemia (ALL). The MS data produced in this study
showed good correlation between the data produced in the fully validated fluorescence based approach developed by Dr. Mark Holland in the Saha group when quantifying AEP activity in plasma. The MS assay offered a lower limit of detection than the fluorescence assay (1 fmol for MS as opposed to 1563 fmol for fluorescence) with comparable variance and a greater linear range (3 orders of magnitude for MS as opposed to 2 orders of magnitude for fluorescence). This study shows that the more active AEP present in the sample, the greater the fluorescence and MS responses. However, at the natural pH of plasma, AEP is not active. Therefore, the pH of the sample needs to be lowered in order for AEP to become active. This was done by diluting the plasma sample five-fold with AEP activity buffer to lower the pH to 5.8. At this pH level, AEP is active and this assay can be performed. This study also shows that AEP activity can be restored in plasma samples by reducing the pH of the sample when AEP has been incubated in plasma before the pH of the sample has been reduced, although the assay must be performed for longer. In addition, we show that a H150A missense mutation at the active site of AEP leads to a dramatic change in the protease’s activity. This would have dramatic consequences in biological systems such as cells, and provide a more accurate picture of AEP activity in the cell, rather than quantifying the amount of AEP in the sample, as would be the case in an ELISA assay. This work shows that this assay can be used in both plasma and WCL samples, extending the range of samples this assay can be performed on, although this assay can only be used for qualitative purposes in WCL samples.

The second clinically relevant protease assessed in this study was caspase-3, which has links to lung cancer. The caspase-3 MRM assay showed limits of detection of 1 fmol, and a linear dynamic range spanning four orders of magnitude. Proof of concept testing by spiking different amounts of active caspase-3 into plasma samples at levels around that seen in sepsis patients has shown no significant change in the amount of cleaved caspase-3 substrate produced. It is suspected that naturally occurring caspase-3 inhibitors, such as members of the inhibitor of apoptosis (IAP) family bind irreversibly to caspase-3, rendering it inactive. As such, caspase-3 activity in plasma cannot be used as a prognostic indicator.

### 6.2 Future Work

The method developed in this study shows that AEP activity can be quantified using SRM MS. Due to the links between AEP activity and poor chemotherapeutic response, AEP activity in plasma samples may have potential as a prognostic indicator. Therefore, samples taken from patients responding well and poorly to L-Asparaginase treatment need to be examined in order to determine whether a significant difference in AEP activity exists.
between these two patient populations. If a significant difference is seen, then monitoring AEP activity in plasma may have potential as a prognostic indicator to predict whether a patient will respond well or poorly to L-Asparaginase treatment.

However, at present, each sample takes around 75 minutes to analyse. As this assay would require triplicate analysis of samples generated from at least two time points, each patient would require over eleven hours of time for this LC-SRM assay to provide the information the clinician needs to decide whether L-Asparaginase treatment is a suitable option for the patient. Therefore, the time taken to analyse samples generated in this assay must be significantly reduced in order for this SRM assay to have potential in a clinical environment.

In addition, it was shown that the variance in the SRM assay is significantly higher than that of the fluorescence assay (Table 4.11). The variance of this assay may be improved by the use of an internal standard of identical amino acid sequence to the cleaved form of the AEP substrate (DVSKhph) which contains stable heavy isotopes. This internal standard would have the same chemical properties as the cleaved substrate of the same amino acid sequence, including its LC retention time, its ability to become ionised by ESI and its losses via the acetonitrile precipitation, drying and re-suspension in LC-MS buffer process. Therefore, any source of variance due to this process could potentially be removed by normalising the cleaved AEP substrate response to the internal standard response. However, as this internal standard contains four L-amino acids, it is still subject to proteolysis by proteases naturally present in plasma and WCL samples. Therefore, this standard must be added after these proteases have become inactivated, but before acetonitrile precipitation takes place. An amendment to the method developed in this project could involve the addition of protease inhibitors such as ethylene diamine tetraacetic acid (EDTA), E64, 4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF), Bestatin, Pepstatin A, E-64, Leupeptin, and 1,10-Phenanthroline. This would inhibit a broad range of proteases including serine, cysteine, aspartic and metalloproteases as well as aminopeptidases, helping to ensure that the internal standard was not subject to proteolysis. Following the addition of these inhibitors and internal standard, samples could then undergo acetonitrile precipitation and the remainder of this LC-SRM workflow.

Additionally, a further study to determine the $K_M$ value for the AEP specific SRM substrate needs to be performed. This involves performing a series of incubations at 37°C using samples containing the sample amount of activated rAEP, but varying the amount of substrate present in the sample. From this, a Lineweaver-Burk plot can be generated and the $K_M$ value for this substrate can be determined. This $K_M$ value can then be compared to that of the AEP
substrate used in the fluorescence assays to determine whether the amino acid sequence of this SRM substrate needs to be altered in order to further optimise this assay.

The caspase-3 assay developed in this project showed no significant increase in the amount of cleaved substrate produced as caspase-3 levels in samples increased. The cause for the failure of this assay must be investigated. The discrepancy between significant SRM responses and no MALDI responses when attempting to detect the cleaved form of the caspase-3 substrate must be investigated by using other techniques such as NMR spectroscopy after the substrate has been incubated in PBS. This can provide structural information about the substrate and determine if it has been cleaved by caspase-3, or if caspase-3 is inactive. In addition, the SRM assay could be compared to the commercially available caspase-3 fluorescence assay to determine whether caspase-3 is active in PBS and in plasma. This would assist in determining the cause of the failure of this assay. If it is shown that caspase-3 is active in plasma based on fluorescent data, then the caspase-3 SRM substrate must be re-designed. If caspase-3 is shown to be inactive in both SRM and fluorescence analysis, then monitoring caspase-3 activity in plasma may not be suitable as a clinical assay, and other methods designed to quantify the amount of caspase-3 in plasma, such as ELISA assays, may potentially be used as an alternative.

Both proof of concept testing using trypsin and AEP testing show that the method developed in this project can be used to quantify protease activity in plasma. Therefore, other proteases can be assessed to determine whether their activities in plasma are linked to clinical outcome. Examples of these proteases include members of the matrix metalloprotease (MMP) and ADAM families, which have been linked to metastasis of lung cancer and tumour progression of pancreatic cancer respectively.
7 Chapter 7: References


251


95. Burke W, Daly M, Garber J, Botkin J, Kahn MJ, Lynch P, et al. Recommendations for follow-up care of individuals with an inherited predisposition to cancer. II. BRCA1 and


256


315. Rehemtulla A, Barr PJ, Rhodes CJ, Kaufman RJ. PACE4 is a member of the mammalian propeptidase family that has overlapping but not identical substrate specificity to PACE. Biochemistry. 1993 11/01;32(43):11586-90.


333. Domon B, editor Advances in quantitative proteomics: Towards targeted strategies. ProteoMMX: Strictly Quantitative; 2010; Chester, UK.


Chen JM, Rawlings ND, Stevens RA, Barrett AJ. Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases. FEBS letters. 1998;441(3):361-5.


8 Chapter 8: Publications and Conference Proceedings

8.1 Publications

David N. Potier, John R. Griffiths, Richard D. Unwin, Michael J. Walker, Emma Carrick, Andrew J. Williamson, Anthony D. Whetton.

An assessment of peptide enrichment methods employing mTRAQ quantification approaches.

*Analytical Chemistry*, 2012, **84** (13), 5604 – 5610.

François Griaud, Andrew J. Williamson, Samuel Taylor, David N. Potier, Elaine Spooncer, Andrew Pierce, Anthony D. Whetton.

BCR/ABL modulates protein phosphorylation associated with the etoposide-induced DNA damage response.


David N. Potier, Mark Holland, Michael J. Walker, Duncan L. Smith, Richard D. Unwin, Vaskar Saha, Anthony D. Whetton.


David N. Potier*, Andrew D. Williamson*, Anthony D. Whetton.

A Comparative Study between Selected Reaction Monitoring (SRM) on 4000 Q-TRAP Triple Quadrupole and 5600 Triple ToF Mass Spectrometers and Pseudo-SRM on an Orbitrap Velos Instrument.

8.2 Conference Proceedings
Royal Society of Chemistry Analytical Research Forum 2012, Durham, UK (Oral):
David N. Potier*, John R. Griffiths, Michael J. Walker, Richard D. Unwin, Duncan Smith, Yvonne Connolly, Ralf Hoffmann, Anthony D. Whetton.
Developing a Mass Spectrometry Based Assay to Monitor Protease Activity in Plasma to Predict a Patient’s Response to Cancer Treatment.

Royal Society of Chemistry Analytical Research Forum 2011, Manchester, UK (Poster):
David N. Potier*, Andrew J. Williamson* and Anthony D. Whetton
A Comparative Study between Selected Reaction Monitoring (SRM) on 4000 Q-TRAP and 5600 Triple ToF Mass Spectrometers and pseudo-SRM on an Orbitrap Velos Instrument.

59th American Society of Mass Spectrometry Conference 2011, Denver, CO (Poster):
David N. Potier*, John R. Griffiths, Richard D. Unwin, Ralf Hoffmann, Anthony D. Whetton.
The Use of Mass Spectrometry Combined with mTRAQ-Labelled Surrogate Standards to Compare Enrichment Strategies for Target Peptides in Plasma.

Paterson Institute for Cancer Research Colloquium 2010, Ambleside, UK (Poster):
David N. Potier*, John R. Griffiths, Ralf Hoffmann, Anthony D. Whetton.
Investigating Protease Activity Using Mass Spectrometry.

Royal Society of Chemistry Analytical Research Forum 2010, Loughborough, UK (Poster):
David N. Potier*, John R. Griffiths, Ralf Hoffmann, Anthony D. Whetton.
A Comparative Study of Multiple Reaction Monitoring (MRM) and Signal Boosting & Ion Trap Scanning (SBITS) to Detect Low Abundance Analytes.